A UK Public Pension System

The UK public pension is complicated and has been reformed many times over the past fifty years. For the cohort we are concerned with, those born in the 1940s, the public pension consists of two components:

1. Entitlement to the first and largest component, known as the 'Basic State Pension' depends only on the number of years an individual has been in work (or been doing other 'creditable' activities – such as caring for young child). Those with a full working life between ages 16 and 64 (or those with 30 or more years for those born after 1945) would be entitled to a fixed amount worth around £3,800 per year in 2002 while those with a less complete work history would receive an amount that was pro-rated by their period in work.

2. The second component, introduced in 1978 and reformed slightly on two occasions subsequently, provided an earnings related aspect. This was known variously as Graduated Retirement Benefit, the State Earnings Related Pension and the Second State Pension. This gave individuals approximately 20% of earnings between a lower threshold (at approximately the 8th percentile of positive earnings) and an upper threshold (at approximately the 80th percentile of positive earnings).

The second (earnings-related) component accounts for approximately 20% of entitlements. The first (closer to flat-rate) components accounts for approximately 80% of entitlements.

The median entitlement for our sample of couples is £7,800 with an the interquartile range stretching from £6,330 – £9,100 in 2002. To compare to Social Security in the US, we can adjusting for prices to 2014 and converting to US dollars using the average exchange rate for that year yields for these percentiles ($15,100 – $18,600 – $21,700). This compares to values (calculated using the Health and Retirement Study) of household Security Society income for a similarly selected sample in the US of ($18,800 – $26,360 – $33,400). Payments under the UK public pension system are lower on average, and dispersed to a lesser extent that are those in the US.

For further information on the UK public pension system see Bozio, Crawford & Tetlow (2010).

B Sensitivity to alternative treatment of housing wealth

In this section we compare our results with an alternative treatment of housing wealth. Following Cagetti (2003), we deduct all housing wealth from the measure of wealth which we use for estimation. We also deduct from net income an age-specific share which we estimate as being spent on mortgage interest and principal. We estimate this share using data on couples in the same cohort as that studies in this paper from the UK's Household Budget Survey: The Living Costs and Food Survey. Figure 7 shows this estimated share.

Figure 7: Mortgage share of income

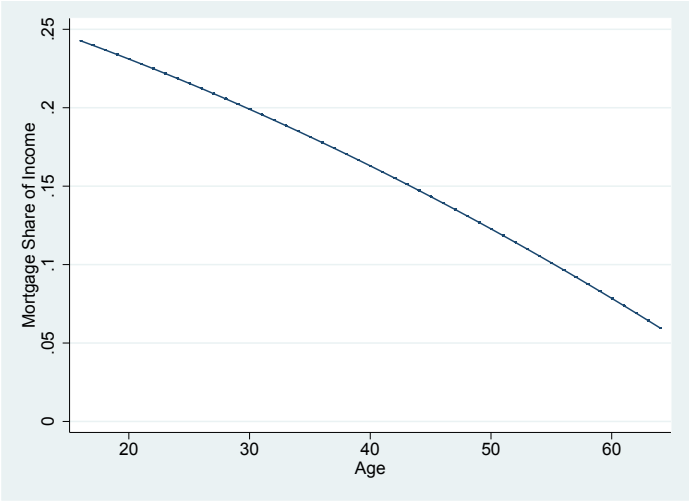


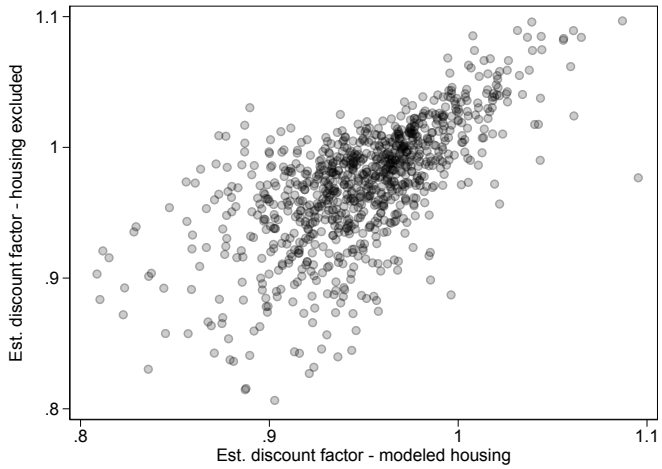
Table 10 summarizes the distribution of discount factors under our baseline and with the alternative treatment of housing; Figure 8 illustrates how the estimated household discount factors correlate with each other at a household level – these two distributions have a correlation coefficient of 0.65.

Table 10: Distribution of discount factors - comparison to alternative housing treatment

Model	Mean	p10	p25	Median	p75	p90
modeled housing	0.947	0.894	0.923	0.950	0.975	1.002
data excludes housing	0.969	0.901	0.944	0.977	1.002	1.026

Figure 9 below compares the share of wealth held in pension wealth in three different versions of the estimated model. Each graph shows, for each decile, the ratio of mean wealth held in private pensions to the mean wealth. Graph a) shows this ratio for a model where the returns on housing are ignored in the calibration of the return on non-pension wealth (this is the model that is the basis for the 'no housing treatment' estimates of discount factors given in the second

Figure 8: Comparison of estimated discount factors – baseline (modeled housing) and housing excluded

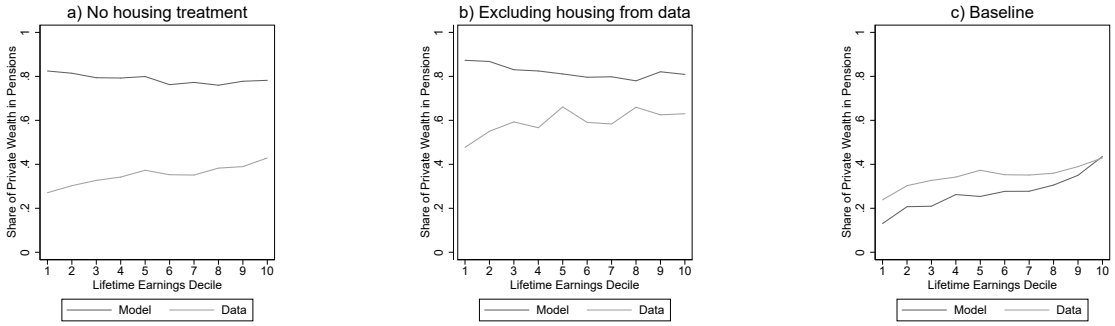


row of Table 4). In this case the model suggests that households will hold most of their wealth in pensions, while in the data most of wealth held in non-pension wealth. Graph b) shows this ratio for our model with the alternative specification outlined in this Appendix where we deduct housing wealth in the data. The model here is the same as that in graph a) (though the estimated discount factors will differ given the difference in the measure of wealth used for estimation). In this version, the model continues to over-predict the share of wealth held in pensions, but to a lesser extent (as in removing housing wealth from non-pension wealth, the latter's share in the data falls).

Graph c) shows the share for our baseline. The data figure here is the same as that for graph a). However, by accounting for the returns on the wealth held in that part of non-pension wealth assumed to be held in housing, the model much more closely matches the share of wealth held in pensions. This is our preferred approach to housing and is the basis for the main results in the paper.

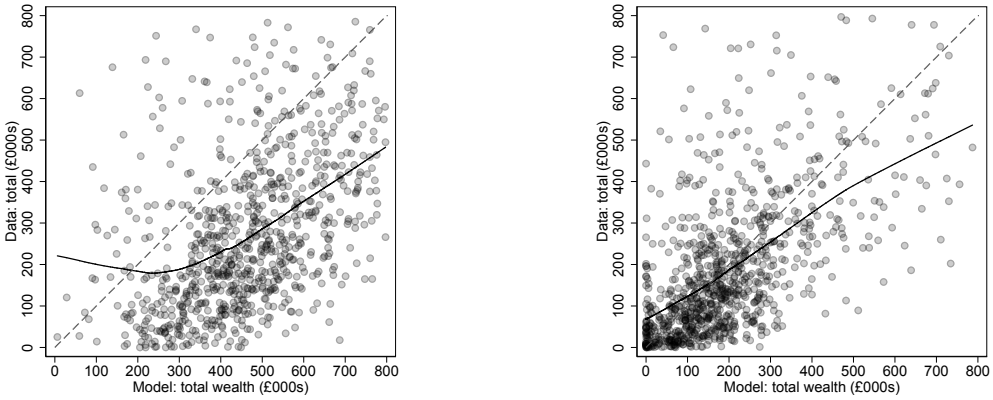
We turn now to the implications of this alternative model for our comparison of our results of those of Scholz et al. (2006) in Section 5.2.2. The left panel of Figure 10 reproduces the graph in Figure 5 which illustrates the fact that, once housing is modeled, fewer than 20% of households have 'oversaved' relative to a benchmark where the discount rate is set equal to the interest rate. The right hand panel shows the equivalent figure when, instead of modeling the return to housing, housing is deducted. Wealth levels are lower in the data here (due to the exclusion of housing); they are also lower in the model due to the return on non-pension wealth being considerably lower and due to our deduction from income of mortgage expenses. In this

Figure 9: Portfolio Shares - Comparing Models



case, the proportion oversaving is 41.7%, not as low in the case of the housing return being modeled, but (as in our baseline) substantially lower than then the 65.4% who oversave when no adjustment is made for housing. Table 11 adds to Table 8 from the paper an additional row which summarizes the oversaving/undersaving results for the alternative housing model.

Figure 10: Comparison of Scholz et al. (2006) style results when housing return is modeled and when housing wealth is excluded



C Sensitivity to alternative treatment of DB wealth

Our model does not include a Defined Benefit pension wealth. Where DB wealth is observed in the data, we model them as having being accumulated as Defined Contribution wealth. In this section, we assess the sensitivity of our results to an alternative treatment of Defined Contribution wealth.

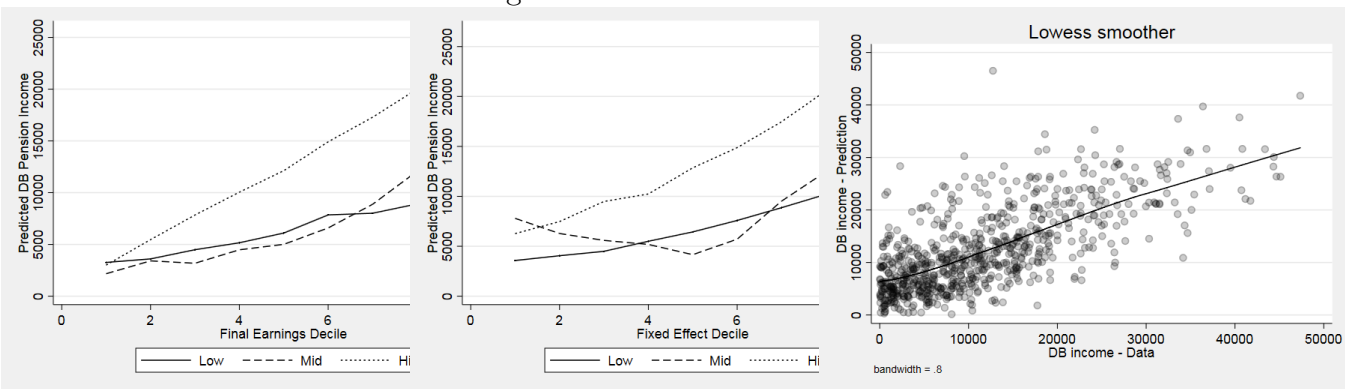
We let Defined Benefit income be a function of a subset of our state variables. We let Defined Benefit pension income be a quadratic function of the earnings fixed effect and 'final earnings'

Table 11: Comparison to Scholz et al. (2006)

Spec	Median wealth		Percent	Median	Median	
	Observed	Optimal	oversaving	surplus	deficit	R^2
One asset	280.7	117.9	87.8	185.0	31.5	0.366
Simple two asset	280.7	203.8	65.4	133.2	65.6	0.398
Alternative housing treatment	148.3	171.6	41.7	77.0	73.3	0.346
Modeled housing (baseline)	280.7	482.9	18.1	198.0	238.2	0.306

(where our measure of final earnings in the data is taken to be the average decile in the final five years in which the individuals are observed). Figure 11 shows the relationship between DB pension income and each of fixed effect (conditional on mean 'final earnings'), final earnings (conditional on mean fixed effect), as well as a scatter of data and predicted Defined Benefit pension income.

Figure 11: Defined Benefit Process



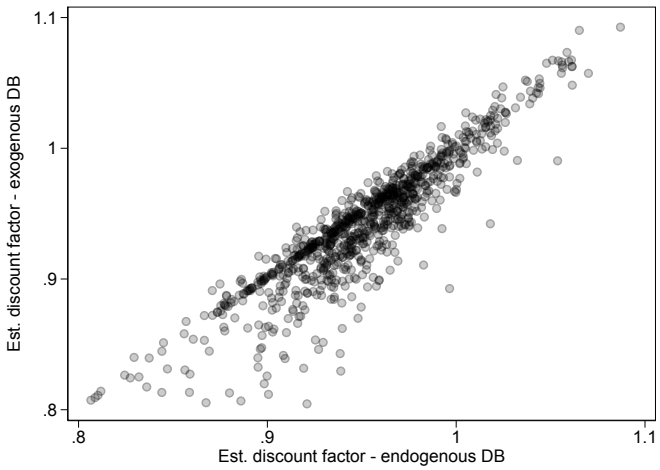
We summarize the differences between our baseline specification and this suggested approach below. In Table 12 we compare the distribution of discount factors under our baseline and under the exogenous-DB treatment showing a close correspondence between the two distributions. Figure 12 shows the comparison of discount rates at an individual level – and shows that the there is a tight link between the estimated degree of patience of each household – the correlation coefficient between these is 0.9.

Turning to the implications of this alternative model for our comparison with Scholz et al. (2006) in Section 5.2.2. Figure 13 below shows a version of Figure 5 from the paper, which shows the scatter of wealth in the data and 'optimal' wealth, where the latter is determined by setting the discount rate equal to the interest rate. The three figures show, from left to right: i) the one-asset model in which all wealth accrues a return equal to the discount rate, ii) a

Table 12: Distribution of discount factors – comparison to exogenous DB specification

Model	Mean	p10	p25	Median	p75	p90
Baseline	0.947	0.894	0.923	0.950	0.975	1.002
... with exogenous DB	0.939	0.881	0.912	0.942	0.969	0.999

Figure 12: Comparison of estimated discount factors – baseline and exogenous DB



two-asset model which adds the tax-advantaged private pension, and iii) our preferred model in which the return on non-pension wealth is set to take into account the high (and typically leveraged) returns that housing has accrued. Unlike the versions in the body of the paper, in these versions of the graph, Defined Benefit wealth is paid out as income in retirement according to the function discussed above.

The figures are very similar to those in our baseline – as are the proportions 'over-saving'. For each of these three models (with those from the baseline model reported in Table 8 in the paper given in parentheses) are 82.0% (87.8%), 65.5% (65.4%), 16.5% (18.1%) for, respectively the one-asset model, the two-asset model not accounting for housing wealth and the the two-asset model which does account for housing wealth.

D Data Appendix

D.1 Earnings data

The national insurance (NI) data are the administrative record of individuals' national insurance contributions, and are the data that is used by the UK government to establish individuals' rights

Figure 13: Comparison of Modeled and Optimal Wealth with Homogeneous Discount Rate set equal to safe rate of return

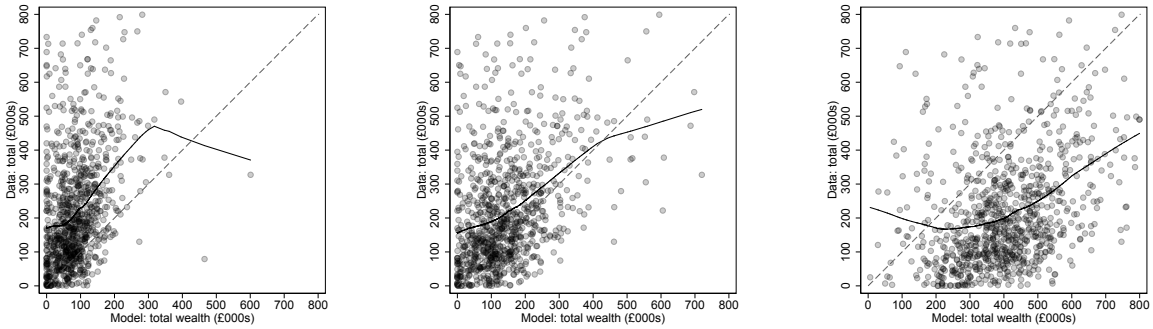
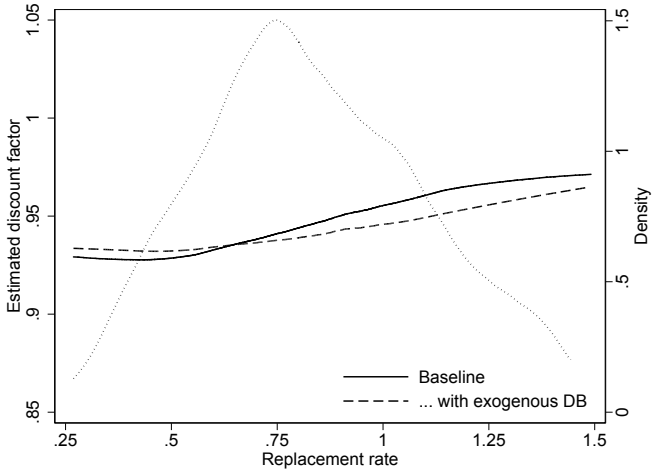


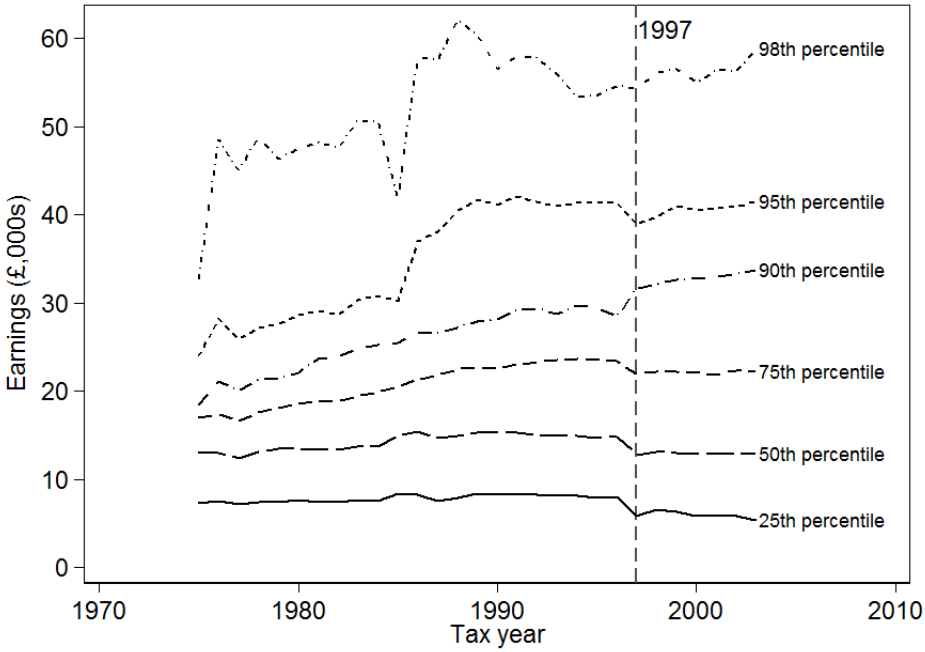
Figure 14: Comparison of relationship between Replacement Rates and Discount Factor – baseline and exogenous DB



to claim contributory benefits such as the state pension. We use this data to estimate ELSA respondents history of earnings. The NI records cover the years 1948 to 2003, though there are different levels of of information for each of three sub-periods: 1948-1974, 1975-1996 and 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. For the middle period - years between 1975 and 1996 - the NI records contain data on employee National Insurance contributions. National Insurance contributions 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

Figure 15: Selected quantiles of earnings



right-censoring at the UEL and conditional on there being some earnings above the LEL. Prior to 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 state benefits or the state pension.)

To predict censored earnings in the years 1975 to 1996, we estimate 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). We use these coefficients to predict earnings for those who are affected by the censoring. The fixed-effect Tobit, with a fixed panel length, yields inconsistent results due to the incidental parameters problem (see Neyman & Scott (1948)). However Greene (2004) investigates its properties, using a Monte Carlo approach, and finds that parameters of the fixed effects Tobit model are little affected by this problem even with panel of lengths substantially shorter than our panel (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.

To simulate earnings before 1975 we follow broadly the methodology used by Bozio et al. (2017). Using the NI data, we calculate an individual's mean earnings over the years 1975 to

2003 in which they are observed working, and then estimate potential previous years' earnings by adjusting for average economy-wide earnings growth and individual level earnings growth given their age, sex and education level. Having obtained this measure of potential earnings in each year, we then need 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. For men we assume they 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, we assume 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 our earnings estimates for years prior to 1975.

Household earnings are calculated by summing in each year the earnings for each individual in the household.

The discussion above relates only to earnings in employment and not income earned in self-employment. National insurance payments 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. Our measure of earnings therefore excludes income from self-employment. However, we have confirmed that our results are not affected by the exclusion of the 13% of households with more than 5 years of self-employment income.

E Additional Details on Parameterization/Estimation

Figure 16 shows the estimated share of non-pension wealth held in net housing wealth: $(s(a))$. Figure 17 shows the estimated leverage ration: $lev(t)$.

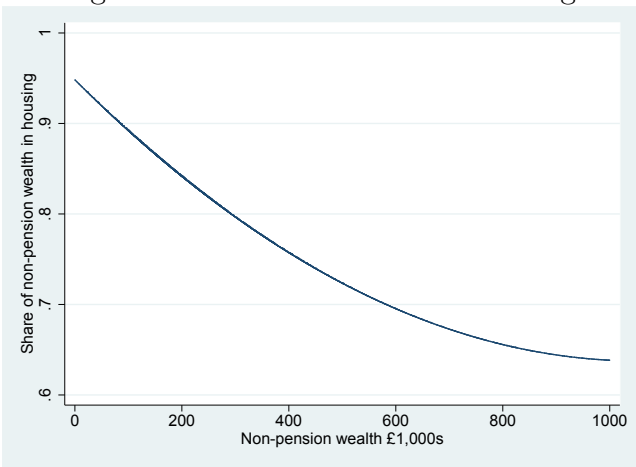
F Taxes and Transfer Function

Net of tax income (y) is given by a function τ:

$$y_t = \tau(e_t, a_t, a_t^c, pp_t, sp_t, h_t, k_t, dc_t, t)$$

that depends on household earnings (e_t), non-pension wealth pre-retirement (a_t), cash post-retirement (a_t^c), private pension payments (pp), public (or 'state') pension payments (sp), number

Figure 16: Share of wealth in housing



Notes: Coefficients that generate liquidity share (i.e. one minus the quantity shown in this figure) on the constant and the linear and quadratic terms on assets (in £000s) are respectively 0.0517, .000589, -2.79e-07.

of adults still alive (h_t), number of dependent children (k_t), chosen contributions to the pension fund (dc_t) (since those attract tax relief) and finally, on the age of the household (t) (the UK tax system taxes the elderly to a lesser extent than those of working age).

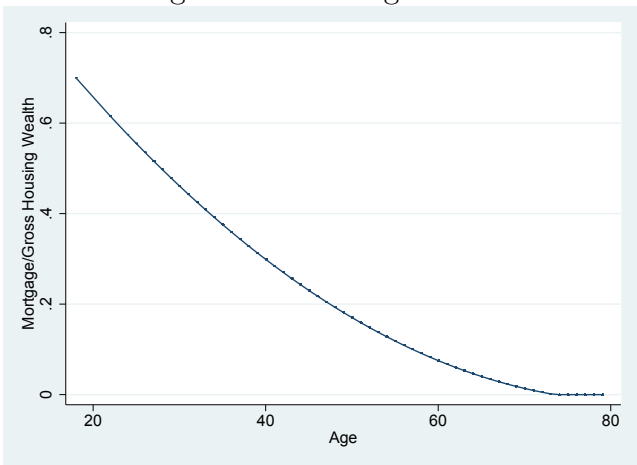
Income tax

Income tax is levied in the UK on a definition of income which includes earnings, private and state pensions and income from capital (excluding the return on home-ownership). income tax system used in the model.

$$y^{gr} = e + pp + sp + r^c a^c$$

In 2002/03, income was taxed in four bands, the smallest was exempt from tax, the second attracted tax at 10%, the third at 22% and the largest at 40%. The thresholds that define the bands vary with age, with a more generous treatment of older individuals. The equations below, together with Table 13, give the income tax function:

Figure 17: Leverage Ratio



Notes: Coefficients that generate the leverage ratio on the constant and the linear and quadratic terms on assets (in £000s) are respectively 1.16, -.028 and .00017.

$$
\begin{aligned}
it(e, pp, sp, a^c, t) &= 0 && \text{if } \; y^{gr} \leq \kappa_1 \\
&= 0.1(y^{gr} - \kappa_1) && \text{if } \; \kappa_1 < y^{gr} \leq \kappa_2 \\
&= 0.1(\kappa_2 - \kappa_1) + 0.22(y^{gr} - \kappa_2) && \text{if } \; \kappa_2 < y^{gr} \leq \kappa_3 \\
&= 0.1(\kappa_2 - \kappa_1) + 0.22(\kappa_3 - \kappa_2) + 0.4(y^{gr} - \kappa_3) && \text{if } \; \kappa_3 > y^{gr}
\end{aligned}
$$

Table 13: Income tax thresholds

	Age		
	< 65	$\geq 65, < 75$	≥ 75
κ_1	4,615	6,100	6,370
κ_2	6,535	8,020	8,290
κ_3	36,435	37,920	38,190

National Insurance

National Insurance contributions are levied on earnings (not on pension income, capital income or other forms of income) and only on those aged less than the state pension age (65). In 2002/03

it was levied at a rate of 10% of income between the 'Lower Earnings Limit' (LEL - £3,900) and the 'Upper Earnings Limit' (UEL £30,420).

$$\begin{aligned} ni(e,t) &= 0.1(max(0,(min(uel,e)-lel))) \text{ if } t<65 \\ &= 0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{ if } t \geq 65 \end{aligned}$$

Jobseekers' Allowance

Jobseekers' Allowance is paid to unemployed households under the age of 60 at a rate which depends on the number of adults and children in the household. In 2002/03 an unemployed couple were entitled to £4,401.80 with an additional payment of £1,924 for each dependent child.

$$\begin{aligned} jsa(e,h,k,t) &= 4401.8 + 1,924k \text{ if } t<60 \text{ and } e=0 \\ &= 0 \qquad\qquad\qquad\qquad \text{ if } t \geq 60 \text{ or } e>0 \end{aligned}$$

Child Benefit

Child Benefit is paid on the basis of the number of dependent children that a household has. It is paid at a more generous rate (£834.6 per year) for first children than subsequent children (£559).

$$childben(k) = 834(1(k \geq 0)) + 559(max((k-1),0)$$

Minimum Income Guarantee

Households aged over 60 are entitled to a means-tested transfer (the Minimum Income Guarantee) that aims to ensure no older household faces destitution. Entitlement to the MIG is based on current circumstances only and does not depend on a household's history of tax or national insurance contributions. The MIG simply tops *net* income up to a minimum level (f, which was £5,184 per year for singles and £7,790 for couples in 2002/03). Define net income (before payment of any MIG) as:

$$y^{preMig} = e + r^c a^c + sp + pp - it - ni + childben + jsa$$

MIG is then:

$$\begin{aligned} mig(e, ra, sp, pp, h, k, t) &= max(0, f - y^{preMig}) \text{ if } t \geq 60 \\ &= 0 \qquad\qquad\qquad\qquad\qquad \text{ if } t < 60 \end{aligned}$$

Net income

We can now summarize the model's net income function. Income before taxes and transfers is given by:

$$\begin{aligned} y^{pre} &= e \qquad\qquad\qquad\qquad\qquad\qquad\qquad \text{ if } t \leq 64 \\ &= sp + pp + r^{h_r} gh_t - r^{mort} lev(t) gh_t \text{ if } t \geq 65 \end{aligned}$$

Pre-retirement y^{pre} is simply earnings, post-retirement this is state and private pension income plus return on gross housing less mortgage interest[20]

$$\begin{aligned} \tau(e, ra, pp, sp, h, k, dc, t) &= y^{pre} - it(e, pp, sp, a^c, t) - ni(e, t) \\ &\quad + jsa(e, h, k, t) + childben(k) + mig(e, ra, sp, pp, h, k, t) \end{aligned}$$

G Computational Appendix

G.1 Value functions

Section 3.2 gives the optimization problems in both retirement and working life faced by households in which neither spouse has died. Here we give the corresponding value function for households where the male has died (the only two differences between these and the case where the wife has died are that the value function in the latter case are conditional on $h_t = 2$ and the survival probabilities are those relating to the male $s_{t+1}^{j,m}$).

[20]Pre-retirement, these last two terms are included in the rate of return on the overall return on housing rather than explicitly measured as in income.

Retired household's problem

$$V_t(a_t^c, gh_t, pp_t, \tilde{e}_{64}, h_t = 3; \theta_i) = \max_{c_t} \Big(u(c_t) \quad + \quad \beta_i s_{t+1}^{j,f} V_{t+1}(a_{t+1}^c, gh_{t+1}, pp_{t+1}, \tilde{e}_{64}, h_{t+1} = 3; \theta_\mathbf{i}) \Big)$$
$$s.t. \quad y_t \quad = \tau(e_t, a_t, a_t^c, gh_t, pp_t, sp_t, h_t, k_t, dc_t, t)$$
$$\text{and} \quad \text{intertemporal budget constraints (5) to (8)}$$

G.2 Model solution and simulation of optimal behavior

In this section we outline how we a) solve the households' maximization problem to obtain decision rules (functions which give, as a function of the state variables, optimal consumption and optimal pension saving) and b) use these decision rules, along with our data to simulate the optimal saving behavior of the households in our sample.

a) Solution

There is no analytical solution to the maximization problem outlined. Decision rules are obtained numerically by iterating on the value function from the final period of life. Let us rewrite the value function in (10) as:

$$V_{100}(\mathbf{X}_{100}; \theta_i) = \max_{c_{100}} u(c_{100}) + \beta E[V_{101}(\mathbf{X}_{101}; \theta_i) | \mathbf{X}_{100}] \tag{16}$$

where the vector $\mathbf{X}$ contains the state variables of the problem and the expectation operator is over survival past the age of 100. For years before retirement, the expectation will additionally be over employment offers, earnings draws and returns on the DC fund. Our assumption that death in the next period is certain for those still alive at the age of 100 ($s_{101}^{j,m} = s_{101}^{j,f} = 0$), combined with the assumption on the absence of bequest motives means that the expectation in equation (16) evaluates to 0. At any particular point in $\mathbf{X}$, the maximization is therefore possible and we can obtain $c_{100}(\mathbf{X}_{100}; \theta_i)$, the consumption function, and $V_{100}(\mathbf{X}_{100}; \theta_i)$, the associated value function at those points (we discuss below our procedure for maximization). The knowledge of $V_{100}(\mathbf{X}_{100}; \theta_i)$ at a subset of points in $\mathbf{X}$, combined with approximation methods (also discussed below), yields an approximation of $V_{100}(\mathbf{X}_{100}; \theta_i)$ ($\hat{V}_{100}(\mathbf{X}_{100}; \theta_i)$) at each point in $\mathbf{X}$.

With an approximation to $V_{100}(\mathbf{X}_{100})$ so obtained, we can solve for approximations to the true consumption function ($\hat{c}_{100}(\mathbf{X}_{100}; \theta_i)$) and value function ($\hat{V}_{100}(\mathbf{X}_{100}; \theta_i)$) for the particular household i at age 99, again at a subset of points in the state space in that period, by solving the following functional equation:

$$\hat{V}_{99}(\mathbf{X}_{99};\theta_i) = \max_{c_{99},} u(c_{99}) + \beta E[\hat{V}_{100}(\mathbf{X}_{100};\theta_i)|\mathbf{X}_{99}] \tag{17}$$

and obtain $\hat{c}_{99}(\mathbf{X}_{99};\theta_i)$ and $(\hat{V}_{99}(\mathbf{X}_{99};\theta_i))$. This iterative process is repeated until we get to the beginning of working life (at age 20). For periods before retirement, a second decision rule - the quantity paid into the pension fund $(\hat{dc}_t(\mathbf{X}_t;\theta_i))$ is also calculated and stored.

Four particular features of the solution procedure will be detailed in the following discussion. These are the i) the discretization of the continuous variables, ii) the process by which the integral in the functional equation is evaluated, iii) the manner in which the value function is approximated at points outside the discretized state space and iv) how the optimization is carried out.

Discretization of state and control variables We have four continuous state variables that need to be discretized. These are earnings, cash assets, pension wealth and pension income. Earnings are placed on a grid (that has 33 elements) using a procedure suggested by Tauchen (1986). Assets, DC stocks and pensions are discretized 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 realizations of these states. 15 discrete points for each of cash assets (and other continuous state variables). Our method for approximating the value function at points not on this grid is discussed in the next sub-section.

There are two choice variables in the model - consumption and the contribution to the DC pension fund. Consumption is not placed on a grid - households can choose any feasible consumption value. To avoid the computational burdens associated with having two continuous control variables, the proportion of earnings that is contributed to the DC pension fund is restricted to take on one of 8 values. That is, households can contribute 0%, 5%, 10%, 15%, 25%, 40%, 75% or 90% of their earnings to the pension fund.

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 discretized state space. To approximate we implement a method in Blundell, Costa Dias, Meghir & Shaw (2016), in the spirit of Carroll (2006). From the perspective of solving for period t consumption, we know marginal utility at every point in the state space of $t+1$: $u'\big(\hat{c}_{t+1}(a_{t+1},.)\big)$. Of the state variables we only make explicit non-pension wealth a_{t+1} here, but the consumption choice does, of course, depend on

the values of the rest of the state variables (indicated by a period). We also know expected marginal utility, conditional on the point in the state space at time t: $E_{\mathbf{X}_t}\big[u'\big(\hat{c}_{t+1}(a_{t+1},.)\big)\big]$. Our objective is to find the root of the Euler equation by finding $\hat{c}_t(a_t,.)$ such that:

$$u'\big(\hat{c}_t(a_t,.)\big) = E_{\mathbf{X}_t}\left[u'\left(\hat{c}_{t+1}\big(a_{t+1}(\hat{c}_t(a_t)),.\big)\right)\right] \tag{18}$$

where we omit, for expositional purposes the discount factor and the interest rate. Note that assets tomorrow ($a_{t+1}(\hat{c}_t(a_t))$ depend on consumption chosen today (through the intertemporal budget constraint), and so we will need to evaluate the right hand side of equation (18) at points off our grid. We could use linear interpolation (which we indicate the object that is approximated with an overbar):

$$u'\big(\hat{c}_t(a_t,.)\big) = \overline{E_{\mathbf{X}_t}\left[u'\left(\hat{c}_{t+1}\big(a_{t+1}(\hat{c}_t(a_t)),.\big)\right)\right]}$$

However, following Carroll (2006) and Blundell, Costa Dias, Meghir & Shaw (2016), we can improve the quality of the approximation by, before interpolating applying the analytical inverse of the marginal utility function, applying the approximation, and then applying the marginal utility function once again.

$$u'\big(\hat{c}_t(a_t,.)\big) = u'\left(\overline{u'^{-1}\left(E_{\mathbf{X}_t}\left[u'\left(\hat{c}_{t+1}\big(a_{t+1}(\hat{c}_t(a_t)),.\big)\right)\right]\right)}\right)$$

The 'quasi-linearized' expected marginal utility function is closer to linear than the marginal utility function, this allows a better approximation, and the use of a lower number of grid points. For a comprehensive exposition of this method see Blundell, Costa Dias, Meghir & Shaw (2016) Appendix pages 7 and 8.

Integration Evaluation of the expectations in the households' problem involves integration of the value function over four stochastic variables. These are unemployment, productivity, survival and the return on DC funds. Realizations of survival and earnings take one of a number of discrete outcomes – the former as it is naturally discrete, the latter as the procedure we apply (Tauchen (1986)) allows earnings to take only a discrete subset of outcomes. Integration over the possible realizations of earnings and survival is therefore carried out by taking a weighted average of the value function realized at each possible outcome with the weights equal to the probability of that outcome. Realizations 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.

Optimization In retirement households face a single choice each year - how much to consume (with the rest of their resources saved in a safe asset). Each optimization is carried out by finding the root the Euler equation. This will successfully find a maximum as our approximated Value Function in retirement is quasiconcave. In working life households face two choices - how much to consume and how much to pay into a pension (again, with the rest of their resources saved in the safe asset). Here we solve (again by finding the root of the Euler equation) for optimal consumption at each of the permitted rates of contribution to the DC fund. The optimal rate of contribution to the DC fund is that which, of these, maximizes utility.

The discretization of the pension contribution choice implies that the approximated value function may be not be quasiconcave and therefore a local optimization routine, like finding the root of the Euler equation, finds the global optimum. The reasons for this are discussed in Appendix A of Low et al. (2009) where a similar issue arises. Those authors suggest that in problems where there is a lot of uncertainty (as there is in ours) local optimizers are likely to obtain the global solution. Their approach is to use the local estimator while estimating their parameters (which involves many solutions of the value function), and then at the set of parameters to check their result using a global optimizer. Our approach is similar. The results presented in the paper (on both the model set out and the sensitivity analyses) use the local optimizer. We then check at our baseline estimate of discount factors that the predicted level of wealth does not materially change when we take a different approach to optimization - one that is robust to departures from quasiconcavity of the value function. This involves restricting consumption to be on a grid of 100 values (so that, in each period, households can choose to consume 1% of their resources, 2%, 3% etc.) and selecting (from the discrete subset of permissible selections) the levels of consumption and pension contribution that maximize utility. The results from this check support the use of the local optimizer - the distribution of wealth is very similar in both cases.

b) Simulation

Once decision rules for pension saving ($\hat{dc}_t(\mathbf{X}_t;\theta_i)$) and consumption ($\hat{c}_t(\mathbf{X}_t;\theta_i)$) are obtained we can simulate the behavior that a household member would exhibit if they followed those rules. The procedure is as follows:

1. Set initial values for state variables at the beginning of working life (age 20). The state variables that are relevant at the start of working life are cash, pension fund value, earnings and household composition. We set cash and pension fund value to zero. We set earnings

to the value on the grid that is closest to actual observed earnings at the age of 20. For household composition we assume both members of the couple are alive and in a couple at that age.

2. Using these values for the state variables and our knowledge of the household's type (θ_i), and the decision rules ($\hat{c}_t(\mathbf{X}_t;\theta_i)$ and $\hat{dc}_t(\mathbf{X}_t;\theta_i)$) we obtain optimal consumption and optimal payments into the pension fund in period 20 (c^i_{20}, dc^i_{20}).

3. Obtain the new state variables for period 21. These are obtained as follows:

   (a) Non-pension wealth in period 21 will follow from the consumption and saving decisions in period 20 along with equation (4) - the intertemporal budget constraint of the working age household.

   (b) Pension wealth in period 21 will be the sum of the stock of pension wealth in period 20, the flow into the pension wealth and the assumed growth rate of pension funds between ages 20 and 21 (equation (19)). That growth rate is assumed to be equal to the growth rate (from our time series of pension fund growth rates) in the year that this household turns 21.

$$DC^i_{21} \quad = (1+\phi_{21})\big(DC^i_{20} + dc^i_{20}\big) \tag{19}$$

   (c) Earnings in period 21 will be that point on the earnings grid that is closest to actual earnings observed at the age of 21.

   (d) Household composition will remain set equal to $h_{t+1} = 1$, that is both members of the couple are still alive. This is as we only retain sample members where nobody has died by the time they are observed in the data.

4. Repeat steps 2 and 3 to obtain optimal consumption and pension saving at each age up to the age at which the (male in the household) is observed in the data in 2002 (we call this age τ). None of these men will have reached their state pension age before this period and therefore the decision rules of retired households are not needed in the simulations.[21] This will allow a time series of the value held in both assets from the age of 20 to age

[21] Though of course the decision rules for working age households could not have been calculated without first solving the retired households' problem.

τ :$\{(x_t^i, DC_t^i)\}_{t=20}^{\tau}$. Our central results involve used simulated optimal wealth at age τ (x_τ^i, DC_τ^i) and that observed in the data at that age for estimation.

H Estimating DC fund return

The DCisions index is an index of total fund return that reflects the asset allocation decisions made by leading DC pension plans in their default investment strategies. Over the period 1994 - 2010 the DCisions index exhibited slightly greater growth than that of the FTSE all-share index (an index representing the performance of the majority of companies listed on the London Stock Exchange). Across financial years where the FTSE all-share index grew in nominal terms, the median ratio of the growth in the DCisions index to the growth in the FTSE all-share index was 1.17, while across financial years where the FTSE all-share index fell in nominal terms, the median ratio was 0.89. This is the result of including re-investment of dividends (the DCisions index is a total return index while the FTSE all-share is an asset price index), slightly offset by the average DC pension plan being diversified into a portfolio with slightly lower return (but also lower risk) than the equities included in the FTSE all-share.

For years 1994 to 2010, therefore, ϕ_t (the model's rate of growth of funds in pension wealth) is assumed to be the real growth in the annualized DCisions index. For years prior to 1994 in which the FTSE all-share index increased in nominal terms, ϕ_t is assumed to be 1.17 times the growth in the FTSE all-share index; for years prior to 1994 in which the FTSE all-share index fell in nominal terms, ϕ_t is assumed to be 0.89 times the decline in the FTSE all-share index. The FTSE index is assumed to have grown by 4% per year in nominal terms in years before data on the FTSE all-share index are available.

I Supplementary Tables and Figures

Section 4.1 gives our method for estimating education-specific survival curves, as well as our values for the interest rate and administrative load. Table 14 gives the modeled annuity rates for each of the nine couple types. These are actuarially fair up to the administrative load. The formula is given in Section 3 of the Supplementary online material accompanying this paper.

The annuity rates vary from the highest rate of 5.47% for a couple where both members of the couple are in a low-educated group to 4.73% where both members are in a high-educated group.

Table 14: Annuity Rates, by education level

		Husband's Education Level		
		Low	Middle	High
Wife's Education	Low	5.47%	5.34%	5.02%
	Middle	5.33%	5.22%	4.91%
	High	5.13%	5.02%	4.73%

Table 15 give the estimates of the parameters of the earnings process for each of three education groups.

Table 15: Estimates of earnings process parameters

	Education group		
	Low	Middle	High
ρ	0.914	0.897	0.875
	(0.0149)	(0.0023)	(0.0831)
σ_ξ^2	0.0496	0.0453	0.0531
	(0.0035)	(0.0022)	(0.0021)
σ_m^2	0.0063	0.0060	0.0066
	(0.0021)	(0.0019)	(0.0020)