Nest Insight

How much is enough?

Technical note to the Nest Insight report on retirement savings adequacy

26 September 2024

nestinsight.org.uk

Authored by

Armine Ghazaryan, Nest Insight

About Nest Insight

Nest Insight is a public-benefit research and innovation centre. Our mission is to find ways to support people to be financially secure, both today and into retirement. We conduct rigorous, cutting-edge research, working collaboratively with industry and academic partners to understand the financial challenges facing low- and moderate-income households. We use these data-driven insights to identify and test practical, real-world solutions. Our findings are shared widely and freely so that people around the world can benefit from our work. For more information visit: **nestinsight.org.uk**

About our strategic partner

BlackRock is a global investment manager serving the UK market for more than 30 years with a purpose to help more and more people experience financial wellbeing. BlackRock's Emergency Savings Initiative is made possible through philanthropic support from the BlackRock Foundation and the BlackRock Charitable Gift Fund. The initiative brings together partner companies and non-profit financial health experts to make saving easier and more accessible for low- to moderate-income people across the US and UK, ultimately helping more people to establish an important financial safety net. For more information, visit **blackrock.com/corporate/about-us/social-impact**

About our programme partner

Phoenix Insights is a think tank set up by Phoenix Group to transform the way society responds to the possibilities of longer lives. They use research to lead fresh debate and inspire the action needed to make better longer lives a reality for all of us. The core of their work is focused on financial security, work, and learning and skills. Reimagining longer lives means making changes in all these areas. At the heart of all of Phoenix Insight's work, they are committed to reducing inequalities and building a society that enables all of us, not just the fortunate few, to live better longer lives. For more information, visit: **thephoenixgroup.com/phoenix-insights**

Acknowledgements

We would like to thank the expert interviewees who helped inform the focus of our research and gave us a broader perspective on the challenge of retirement adequacy: Joe Ahern, WPI Economics; Renny Biggins of TISA; Vivien Burrows of the ILC; Paul Cullum of Frontier Economics; Nathan Long of Hargreaves Lansdowne and Ben Skelton of Oxford Economics; Shelley Morris of the Living Wage Foundation and Mike Brewer of the Resolution Foundation; Tim Pike and John Upton of the PPI; Simon Sarkar of the PLSA; Alice Duggan and Jake Blampied of Nest Corporation and Matt Padley of Loughborough University – as well as those who provided further input through our roundtable sessions.

About this report

This report contributes to Nest Insight's larger research programme on supporting financial resilience and wellbeing through joined-up savings solutions. For more information, visit: **nestinsight.org.uk/research-projects/pensions-adequacy-and-the-household-balance-sheet/**

Contents

Introduction

The Nest Insight report *How much is enough* aims to give a new perspective on retirement savings adequacy. It does this by modelling the lifetime experiences of 30 modelled ‘saver personas’ and their households.

The main report provides a detailed account of the findings from this modelling exercise, illustrated by many output data from the persona modelling work. The separate annex to the report provides still more detail on the lifetime experiences and pensions savings priorities of all 30 of our personas. However, both documents deliberately avoid going into to much depth or technical detail about the statistical and analytic methods used to general our realistic saver personas. This technical note provides that detail.

It's important to say up front that the selection of the saver personas is fundamentally arbitrary, as we have exercised judgement to choose combinations of factors that characterise people reflecting different circumstances experienced by real individuals in the USoc panel. This is a deliberate strategy. The purpose of this selective approach is that it enables us to understand the experiences and priorities of people who diverge from the norms, averages or medians of the workplace pensions population.

As a consequence, our findings illustrate the financial realities experience by *some* people working in the UK today, but they don’t attempt to set out the experiences of *all* people. Nor can they quantify the subgroups or segments represented by our selected personas.

Summary of our approach

We use the following structure to create the adequacy model:

- We use data from UK Household Longitudinal Study to model a number of income and likelihood outcomes for individuals.
- We then use same data to define 30 individuals based on various socioeconomic individual and household characteristics.
- We use the model results to predict the income and likelihood outcomes for personas.
- We use the household inputs to make income predictions for personas’ partners.
- We use NEST internal pensions projections model to make personas’ pensions projections based on predicted income and a number of assumptions regarding contribution start age, retirement age, prior contributions patterns, using inputs of baseline inflation for all projections and projected state pension income.
- Using personas’ partners’ income predictions, we make projections of partners’ retirement income in similar manner.
- We use personas’ working life individual and household income predictions, including income from labour and other sources, and housing costs, against retirement income projections to optimise personas pensions adequacy approach.

The following section provides detail on each of these steps.

Methodology

Data

We use data from waves 1-13 of the UK Household Longitudinal Study, also known as Understanding Society (hereinafter, USoc). We define the *main sample* to include individuals of ages 22 to 65, who are in paid employment and earn a monthly gross income above £833 and below £6,000. We look at this subset of Understanding Society sample in order to capture individuals of working age that are eligible for auto-enrolment, and who are representative of low- and middle-income individuals.

Additionally, we define the *self-employed sample* to include individuals of ages 18 and 65, who are self-employed and earn a monthly gross income above £0 and below £6,000. We also define the *sample of individuals below auto-enrolment* (AE) *threshold*, which includes individuals of ages 18 to 21, who are in paid employment and earn a monthly gross income above £833 and below £6,000, or individuals of ages 22 to 65, who are in paid employment and earn a monthly gross income below £833.

We identify and define a number of outcome variables for the model, including:

- **Income variables:**
  - individual gross labour income (***fimnlabgrs_dv***)
  - net labour income (***fimnlabnet_dv***)
  - household gross labour income (***fihhmnlabgrs_dv***)
  - household net income from all sources (***fihhmnnet1_dv***)
  - housing cost (***houscost1_dv***)
- **Binary variables:**
  - likelihood of being behind with mortgage/rent payments (based on **xphsdb**)
  - likelihood of saving (**save**)
  - likelihood of being a member of a pension scheme (derived, based on **pppex, penmex, jbpenm**)
- **Categorical variables:**
  - levels of savings -- Monthly saved: none, Monthly saved: (£0-£50], Monthly saved: (£50-£150], Monthly saved: (£150-£500], Monthly saved: over £500 (derived, based on **saved**)
  - categories of current subjective financial wellbeing (based on **finnow**) -- Living comfortably, Doing alright, Just about getting by, Finding it quite difficult, Finding it very difficult
  - categories of future financial wellbeing (based on **finfut**) -- Better off, Worse off than now, About the same
  - categories of housing situation (based on **hsownd**) -- Owned outright, Owned/being bought on mortgage, Shared ownership, Rented, Rent free, Other
  - categories of being up to date with bills (based on **xphsdba**) -- Up to date with all bills, Behind with some bills, Behind with all bills
  - debt levels (derived, based on debtc1-5) -- No debt or under £100, Debt levels: [£100-£500), Debt levels: [£500-£1500), Debt levels: [£1500-£5000), Debt levels: [£5000-£10000), Debt levels: £10000 or more
  - categories of pensions contributions (derived, based on **ppen, penmex, ppreg, penmcn**) -- Doesn't contribute/no pension scheme, Contributes towards pensions, Non-contributory, Stopped contributing

We define the following list of control variables:

- Age (**dvage**)
- Sex (**sex**)
- Ethnicity (**ethn_dv**)
- Marital Status (based on **mastat_dv**)
- Number of children aged 15 or under (**nkids_dv**)
- Number of people in household (**hhsize**)
- Housing situation (**hsownd**)
- Region (**gor_dv**)

- Job industry (**jbiindb_dv**)
- Current job type (**jbnssec3_dv**)
- Highest educational qualification achieved (**hiqual_dv**)
- Long standing illness or disability (**health**)
- Inheritance (**rtfnd6**)
- Job change (derived, based on **jbbgm, jbbgd, jbbgy, jbbgdatm, jbbgdatd, jbbgdaty**)
- Job type: permanent, contract, seasonal (derived, based on **jbterm**)
- Pension contributions as a percent of salary (based on **penmpy** and **penscheme**)
- On housing benefits (derived, based on **rent_dv** and **rentgrs_dv**)
- Being responsible for childcare (based on **ccare**)
- Caring for sick/disabled/elderly (**aidhh**),
- On benefits (derived, based on **fimnsben_dv**)
- Sample Type (main sample, self-employed, below AE threshold)
- Absolute Income Volatility and Annual Income Volatility.

The two measures of income volatility are defined following for instance Moffitt & Zhang (2022), Carr et al. (2022):

- annual volatility: AnVol=$\left|\frac{Y_t-Y_{t-1}}{\left[\frac{Y_t-Y_{t-1}}{2}\right]}\right|$, where Y is the total household net real income
- absolute volatility (variance): volatility=$\frac{\sum_1^N(AnVol_t-\overline{AnVol})^2}{N}$, where $\overline{AnVol}$ is the average volatility for individual.

Persona definitions

With personas we aim to bring to life the diversity of individual and household circumstances by introducing numerous persona individuals who are living in different households. We have drawn up a set of these personas, based on our data analysis, to illustrate and compare different aspects of the financial and household situations that people are living in today.

We define personas based on realistic, yet fictional individuals. We use the USoc sample to define a subsample of individual characteristics a persona might have based on their initially assigned income quintile, age and gender. From that subsample we then select persona ethnicities, then given the ethnicity we select the region. From that point the rest of the variable value inputs are based on this sample, so that no individuals are indefinable. We use some of the probabilistic outcome projections described above to aid persona definitions, specifically when defining the housing situation.

Following the approach above we define 30 personas:

- 24 personas are based on the AE main sample, out of which 2 personas (male and female) represent statistically average individuals
- 2 personas are from the sample below AE threshold
- 4 personas represent the self-employed sample

Once personas are defined and modelled, we then define persona yearly life events up until retirement age. The life events are based on manipulation of control variables described above and are based on expert assumptions of realistic life events and combination of events at certain points of personas' working life. Personas' income variables are based on the combination of these characteristics and are results of model predictions, and hence we do not manually manipulate income or other output variables.

Thus, personas and their life events are not defined based on statistical mode, rather they are based on realistic representation of personas in an inclusive manner within the pool of personas in the gross income quintiles 1-4.

Modelling

We use data from USoc and the samples defined above for the modelling. The modelling consists of several approaches and stages.

1. In Step I of the modelling we model income variables:

$$\log(income)_i = \alpha + \beta\, log(individual\ gross\ labour\ income)_i + \boldsymbol{\gamma}\, \boldsymbol{X_i} + \varepsilon_i$$

   where $\log(income)_i$ is the logarithmic transformation of the income variables listed above (except for individual gross labour income) $\boldsymbol{X_i}$ includes a vector of control variables as defined above for an individual i . We use pooled cross-section linear regression to estimate this model. We estimate this set of models for the three samples described above (AE, below AE threshold, Self-employed).

2. In Step II we model the binary and categorical variables.
   a. The binary variables: $\Pr(\mathrm{Y}_i = 1) = \alpha + \beta\, log(individual\ gross\ labour\ income)_i + \boldsymbol{\gamma}\, \boldsymbol{X_i}$, we use probit and logit approaches to estimate this model, both exhibiting similar results.
   b. The categorical variables: $\Pr(\mathrm{Y}_i = j) = \alpha + \beta\, log(individual\ gross\ labour\ income)_i + \boldsymbol{\gamma}\, \boldsymbol{X_i}$, we use ordered probit to estimate this model.

   We estimate this model for all three samples combined using a binary variable for each of the three samples (AE, below AE threshold, Self-employed).

   The estimates of the effect of pensions contributions based on a smaller sample due to smaller number of respondents to the question. The outcome variables of likelihood of being a member of a pension scheme and categories of pensions contributions have limited availability in USoc sample due to the question being asked in the USoc questionnaire every other wave, and/or due to it being asked to individuals of the age of 30 or over. Therefore, the estimates based on the models for these variables are likely to be biased due to sample selection issue. To correct for the bias we apply the two-step sample selection bias correction approach (Heckman, 1976, 1979), using (1) a binary variable for individuals of the age of 30 or over as an instrument for being a member of a pension scheme, and (2) the binary variable for individuals of the age of 30 or over and the binary variable for being a member of a pension scheme as instruments for the variable whether contributes towards pension.

3. In Step III we apply dynamic models to model income variable dynamics over time:
   a. For individual gross labour income:

$$\log(individual\ gross\ labour\ income)_{it} = \alpha_i + \beta\, log(individual\ gross\ labour\ income)_{i,t-1} + \boldsymbol{\gamma}\, \boldsymbol{X_{it}} + \tau_t + \varepsilon_{it}$$

   b. For other income outcome variables:

$$\log(income)_{it} = \alpha_i + \beta\, log(income)_{i,t-1} + \delta\, log(individual\ gross\ labour\ income)_{i,t} + \boldsymbol{\gamma}\, \boldsymbol{X_{it}} + \tau_t + \varepsilon_{it}$$

   We use Arellano and Bond approach (Arellano & Bond, 1991) to estimate these dynamic panel models with fixed effects. A separate Hausman test confirms that fixed effects approach is appropriate for the model.

4. In Step IV we model the binary variables, and the categorical variables presented as binary using panel data with fixed effects. This step is aimed at estimating the effect of pensions contributions on the likelihood outcomes.

5. We use probit and ordered probit for modelling life events. Estimating panel probit using Correlated Random Effects (Mundlak, 1978; Perales & Schunck, 2021) that would account for time-invariant fixed component, is not appropriate for predicting life events. When predicting probabilities for personas, CRE would require additional assumptions on the mean of covariates, thus reducing the added value of CRE and control for unobserved effect, therefore the predictions of probabilities and categorical variables are conducted using probit and ordered probit cross-sectional models using income predictions from Arellano and Bond estimators.

The predictive power and goodness of fit of the income variables is high, and it is moderate to high for probabilistic models.

The regression outcomes are included in the main spreadsheet, and persona outcomes are predicted based on their individual and household characteristics, as well as life events, and regression results. The cross-sectional results are based on the latest year in the USoc sample, specifically, 2022-2023, while longitudinal predictions are not inflation adjusted, and therefore in the final results they are adjusted for inflation, based on the assumption of inflation being 2% post-2023, using Bank of England inflation target.

Adequacy optimisation

We used gross individual labour income predictions as inputs for Nest's internal pensions projections model to make projections about each persona's, and their partners', pension incomes. This uses investment market data to project investment returns using the Nest scheme's default investment strategy, which is based on a target date fund approach tailored to the individual's retirement age. It is based on the qualifying earnings calculation used to calculate automatic enrolment contributions. Retirement age is set to be 68n for all personas. The personas are assumed to convert all their DC savings into a retirement income rather than taking a lump sum at retirement, in order to maximise their potential incomes. State pension amount is based on the 2024 rate, which we assume to remain constant in real terms[1].

We apply pension contribution histories for each persona based on their labour incomes, using Office for National Statistics data on distribution of frequency of different numbers of years of historic defined benefit and defined contribution savings. The observed personas' contribution patterns then follow two scenarios:

› **Baseline scenario**, when all personas are assumed to contribute 5% (including tax relief) from the year they enter into our sample up until the age of 68 when they retire.
› **Optimised scenario**, when pension contributions are computed as a result of an optimisation task as described below.

Adequacy optimisation is aimed at smoothing working life income and income at retirement. We specifically look at per adult disposable income:

› During working life, it is calculated as per adult [household net income from all sources (–) any housing costs (mortgage, rent)]
› At retirement, it is calculated as per adult of [total pension projected income (sum of persona's and partner's income, if a partner is present at retirement, and it includes projections of state pension income, assumed maximum projected income for all working personas) (+) any defined benefits pension projections (+) any non-labour income persona and their household has at retirement that is expected to continue into retirement (–) any housing costs that is expected to continue into retirement (mortgage, rent)]

Thus the optimisation tasks is defined as follows:

› **Optimisation task** (lifetime income smoothing) – minimise the difference between average disposable income over working live and disposable income at retirement
› **Constraints**:
  – contributions can change once every decade
  – contributions can range [0,12]
  – disposable income every year should be above the Minimum Income Standard[2]
  – the difference between average disposable income over working live and disposable income at retirement should be non-negative

It should be noted that the constraints and the optimisation task are not always achievable, in which case the model aims to achieve next best solution.

1 In reality the full State Pension amount has risen faster than inflation for a number of years, because of the 'triple lock' mechanism currently in place. We have opted not to assume the triple lock mechanism continues indefinitely, given that many of our personas have decades to go before retirement, and to model future outcomes that included the triple lock would significantly inflate their outcomes. As such, we believe that flat real terms growth in the State Pension provides a better baseline for possible outcomes.

2 *A Minimum Income Standard for the United Kingdom in 2023*, Loughborough University. Minimum Income Standard (hereinafter, MIS) is calculated based on monthly equivalent of weekly MIS budgets of the report. It includes "Total excluding rent and childcare" for single working age adult, for lone parent, and couples. The report specifies the budget for parents with two children aged 2-4, however we include the figure for couples as a measure for all couples, divided by two, while for lone parents the respective figure is included for a parent with children aged 15 or younger. This is certainly an approximation, and ideally we would have a measure of MIS for each type of couple, however, the assumption we make is that as childcare and rent are not included in our measure of MIS, so the approximation brings it close to single working age adult for each individual.

References

1. Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. The review of economic studies, 58(2), 277-297.
2. Carr, M. D., Moffitt, R. A., & Wiemers, E. E. (2022). Reconciling Trends in Male Earnings Volatility: Evidence from the SIPP Survey and Administrative Data. Journal of Business & Economic Statistics, 41(1), 26-32.
3. Heckman, J. J. (1976). The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models. In Annals of economic and social measurement, volume 5, number 4 (pp. 475-492). NBER.
4. Heckman, J. J. (1979). Sample selection bias as a specification error. Econometrica: Journal of the econometric society, 153-161.
5. Moffitt, R., & Zhang, S. (2022). Estimating trends in male earnings volatility with the panel study of income dynamics. Journal of Business & Economic Statistics, 41(1), 20-25.
6. Mundlak, Yair. "On the pooling of time series and cross section data." Econometrica: journal of the Econometric Society (1978): 69-85.
7. Perales, F. P., & Schunck, R. (2021). XTHYBRID: Stata module to estimate hybrid and correlated random effect (Mundlak) models within the framework of generalized linear mixed models (GLMM).

Contact us:
insight@nestcorporation.org.uk

To find out more, visit our website:
nestinsight.org.uk