National Institute of Economic and Social Research

LINDA: A DYNAMIC MICROSIMULATION MODEL FOR ANALYSING POLICY EFFECTS ON THE EVOLVING POPULATION CROSS-SECTION*

Justin van der Ven[1]

[1]NIESR, London, UK, and MIAESR, The University of Melbourne, Australia

*Development of the LINDA model has been supported by funding from the Joseph Rowntree Foundation, HM Treasury, HM Revenue and Customs, the UK Department for Work and Pensions, the Economic and Social Research Council (Grant number RES-194-23-0005), and the European Commission. The basic architecture for the model was established by James Sefton, and benefitted from supervision by Martin Weale. I thank Guoda Cibaite for useful comments on an earlier draft. The usual disclaimers apply.

NIESR Discussion Paper No. 459

Date: 16th May 2016

National Institute of Economic and Social Research

About the National Institute of Economic and Social Research

The National Institute of Economic and Social Research is Britain's longest established independent research institute, founded in 1938. The vision of our founders was to carry out research to improve understanding of the economic and social forces that affect people’s lives, and the ways in which policy can bring about change. Seventy-five years later, this remains central to NIESR’s ethos. We continue to apply our expertise in both quantitative and qualitative methods and our understanding of economic and social issues to current debates and to influence policy. The Institute is independent of all party political interests.

National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
T: +44 (0)20 7222 7665
E: enquiries@niesr.ac.uk
niesr.ac.uk
Registered charity no. 306083

This paper was first published in May 2016

LINDA: A dynamic microsimulation model for analysing policy effects on the evolving population cross-section*

Justin van de Ven†

Abstract

This paper describes a structural dynamic microsimulation model that generates individual-specific data over a range of demographic and economic characteristics at discrete intervals throughout a simulated time horizon. The model is designed to analyse the distributional implications of policy alternatives over appreciable periods of time. This focus motivates endogenous simulation of savings and labour supply decisions, taking explicit account of uncertainty regarding the evolving decision environment. In contrast to the existing literature of savings in context of uncertainty, the model described here takes an overlapping generations form which is adapted to the needs of policy makers, and which has distinct advantages for empirical investigations.

Key Words: Dynamic Programming, Savings, Labour Supply

JEL Classifications: C51, C61, C63, H31

1 Introduction

Good policy design is a fiendishly difficult business due to the multiplicity, complexity, and inherent uncertainty of the considerations that are involved. One consideration that is often poorly understood is the variable impact that policy can have when considered over alternative time horizons. A welfare benefit may, for example, be interpreted as redistributing income between different members of a population when its incidence is observed at a particular point in time, and be interpreted as redistributing income across the life-course of individuals when considered over longer time horizons. Alternatively, a policy may have very different distributional implications when considered at alternative points in time, especially when endogeneity of behaviour is taken into consideration. Interest in understanding how policy influences individual circumstances over alternative time spans is an important motivation for the development of dynamic microsimulation models. This paper provides a technical description of the Lifetime INcome Distributional Analysis model, or LINDA for short, which is an example of the current state-of-the-art in the field of dynamic microsimulation modelling.

Dynamic microsimulation models suitable for analysing the distributional implications of public policy have been growing in number and sophistication since the ground breaking-work of Orcutt (1957). UK examples from this literature include PenSim2 (Emmerson et al. 2004), SAGE (Zaidi 2007),

*Acknowledgements: Development of the LINDA model has been supported by funding from the Joseph Rowntree Foundation, HM Treasury, HM Revenue and Customs, the UK Department for Work and Pensions, the Economic and Social Research Council (Grant number RES-194-23-0005), and the European Commission. The basic architecture for the model was established by James Sefton, and benefitted from supervision by Martin Weale. I thank Guoda Cibaite for useful comments on an earlier draft. The usual disclaimers apply.

†J. van de Ven: NIESR, London, UK, and MIAESR, The University of Melbourne, Australia: jvandeven@niesr.ac.uk.

SimBritain (Ballas et al. 2005), and CARESIM (Hancock et al. 2007); see Li & O'Donoghue (2013) for a review citing 66 such models for 19 countries. Development of dynamic microsimulation models has benefitted from the increasing availability of detailed microdata, improvements in analytical methods, the advent of generic software packages (e.g. GENESIS, Edwards (2010); LIAM, O'Donoghue et al. (2009)), and a steady rise in computing power. Nevertheless, constructing this type of model remains both technically and computationally challenging, and current implementations consequently all impose non-trivial stylisations of one form or another.

One of the most important stylisations commonly applied in the dynamic microsimulation literature concerns the projection of micro-unit behaviour. The importance of reflecting agent decision making increases with behavioural sensitivity to variation of interest (e.g. policy counterfactuals), and with the bearing that behaviour has on projected characteristics of interest (e.g. government budgets). Such considerations are exaggerated as the projected time-horizon is lengthened, due to feedback effects of behaviour on individual circumstances, and are therefore particularly relevant for dynamic microsimulation models that project circumstances well beyond a short (five year) time horizon. Nevertheless, fewer than one third of the models surveyed by Li & O'Donoghue (2013) are identified as using "behavioural equations" to project decisions through time.[1] Furthermore, even where behavioural variation is projected through time, it is common for these projections to be based on reduced form equations that are ill-suited to respond to evolving incentives; this is the case, for example, for all three of the dynamic microsimulation models for the UK cited by Li & O'Donoghue (2013) that include behavioural projections (PenSim2, SAGE, and a model produced at the IFS described in Brewer et al. (2007)).[2]

A structural model is designed specifically to permit analysis of behaviour in alternative decision contexts. This is achieved by making explicit assumptions concerning how decisions are made, and what aspects of the decision process can be taken to be invariant to the prevailing decision environment. In economics, attention has focussed on understanding behaviour as a product of the incentives that individuals face. The most common method of formalising the relationship between incentives and decisions for individual consumers is through the mathematical framework of utility optimisation. Despite its widespread use by the economics profession, however, utility theory has been the subject of considerable controversy, renewed in recent years following the short-comings of economic theory made clear by the 2007 global financial crisis, and the great recession that followed.

[1] This omission of an explicit allowance for behaviour response is also a stylisation that is commonly employed in the wider empirical literature; see for example Kuang et al. (2011).

[2] All three of these models simulate employment transitions based on probabilities that vary by a range of characteristics, including demographics (e.g. age, sex, relationship status, dependent children), educational attainment, health status, and past work experience. SAGE and the IFS model summarise these probabilities in the form of logit regression equations, which can be derived from a trans-log utility function, and are sometimes therefore described as 'structural'. Nevertheless, these models are denoted 'reduced form' here, because none of them is designed to project labour responses to changes in transfer policy (the explanatory variables being exogenously defined).

Much of the recent controversy concerning utility theory has focussed on the generally perceived lack of realism of the numerical optimisations upon which the theory is based (see, e.g. Simon (1955)). Friedman's (1953) riposte to this criticism is that the usefulness of utility theory lies in its ability to reflect reality, not whether the theory accurately describes the decision process.[3] In this regard it is of note that the most general form of the (utility maximising) life-cycle framework "encompasses many different types of behavior and has almost no testable implications" (Attanasio & Weber (2010), p. 695). Seen from this perspective, the mathematical apparatus associated with utility is best understood as a tool for translating incentives into behaviour; or conversely, for inferring incentives from modelled behaviour. The structural framework that is assumed for the current model reflects our view that, as long as there is a systematic relationship between the incentives that people face and the decisions that they make, utility theory provides a useful tool for considering the behavioural effects of changes to the decision environment.

Some variants of the utility framework permit behaviour to be described by 'closed-form' equations, which are qualitatively similar to the reduced form behavioural descriptions used by the three UK models that are cited above. The analytical convenience of such models, in conjunction with the appeal of a structural framework, has motivated their inclusion in a range of models; see, for example, Pylkkänen (2002) for SESIM. Unfortunately closed-form solutions do not exist for most favoured preference relations in realistic dynamic contexts where the future is understood to be uncertain. In the current context, uncertainty is associated with any variable that includes at least one random component in its functional description through time. Unpredictable labour market opportunities, investment returns, and health care costs are all common examples where uncertainty is important, complicating savings and investment decisions, which in turn can be expected to influence decisions concerning employment. Furthermore, the practical limitations implied by this observation are emphasised by a series of papers which show that even apparently minor forms of uncertainty can have a substantial impact on the behaviour implied by the life-cycle framework – see especially Kimball (1990), Deaton (1991), and Carroll (1992); Browning & Lusardi (1996) provide a simple worked example. These observations detract from models based on utility functions with closed-form solutions, and have spurred academic interest in models that must be solved numerically.[4]

Current best practice in the economic analysis of savings behaviour uses dynamic programming (DP) methods to solve for utility maximising decisions, taking uncertainty explicitly into account (especially

[3] Friedman's riposte can be interpreted as a response to John Searle's (1984) "Chinese room" thought experiment, which is designed to refute the proposition that a computer programme has the potential to replicate a mind or consciousness. Friedman's point is that the objective of utility theory is to replicate human decision making. This objective is weaker than the objective of research into 'strong AI', which seeks to replicate the human thought process.

[4] See Emmerson et al. (2004), Section 4.1, for extended discussion of this criticism in relation to PenSim2.

concerning earnings, investment returns, and survival).[5] However, DP models of intertemporal decision making over continuous domains are complex, time consuming, and costly to implement. It is little wonder then that, despite featuring prominently in the economic literature during the last two decades, these models have not yet gained much traction within policy making institutions. The LINDA model is designed to address this gap, by making current best-practice micro-economic methods of behavioural analysis accessible to, and in a format that is adapted for (UK) policy makers.

Starting from survey micro-data reported for a reference population cross-section, LINDA is designed to (i) project the evolving population cross-section forward through time, and (ii) permit analysis of characteristics measured over alternative time horizons, ranging from a single year, up to the entire life-course. A range of demographic and financial characteristics are projected at discrete intervals through time for each adult represented in the simulated population, treating labour supply and savings as endogenous. The model departs from the DP literature on savings and labour supply by the extent to which household specific heterogeneity is accommodated, including age, year of birth, education status, health status, relationship status, the birth and aging of dependent children, labour market opportunities, sector of employment, and a basket of alternative assets. Much of this heterogeneity is commonly suppressed in focussed academic studies, but is important to policy makers. Furthermore, the overlapping-generations structure adopted for the model – which contrasts with most of the existing literature concerned with savings and labour supply behaviour in context of uncertainty[6] – presents considerable advantages for conducting associated empirical analyses. We are unaware of any DP model that accommodates a comparable degree of heterogeneity to LINDA, or which projects individual circumstances both *forward and backward* through time, where backward projections are made necessary by the objective to describe the life-course of older individuals in a reference population cross-section.[7]

Consistent with the objective of widening access of current best-practice microsimulation modelling beyond specialists, the paper has been written assuming a Bachelor's appreciation of economics. Nevertheless, some jargon is unavoidable, and a glossary of selected terms is provided in Appendix A. An overview of the model is provided in Section 2. Sections 3 to 13 describe each model characteristic in turn; tables defining the variables used are provided in Appendix B. Details regarding the routines that project behaviour through time are described in Section 14, and a summary and directions for future research are provided in the conclusion. In keeping with the objective of developing a model fit for use by non-specialists, this text avoids use of technical terminology associated with dynamic programming

[5] Nagatani (1972) and Zeldes (1989) are early examples of the contemporary literature concerned with savings decisions in context of earnings uncertainty. Most of the related literature that is concerned with savings and employment focusses on the US context: see, for example, Gustman & Steinmeier (1986), Hubbard et al. (1995), Rust & Phelan (1997), and French (2005). For recent analysis of pension policy in the UK, see Sefton et al. (2008) and Sefton & van de Ven (2009).

[6] Livshits et al. (2007) is one exception from the existing DP literature that does take an OLG specification.

[7] The lack of comparable modelling structures has motivated two projects funded by the European Commission to develop LINDA for use in the Republic of Ireland and Italy respectively.

wherever possible.

2 Overview of LINDA

The model is designed to start with detailed micro-data that describe the circumstances of a cross-sectional sample of reference adults, and to project their evolving circumstances at discrete intervals forward and backward through time, eventually building up a complete life-history for each reference person. The reference population is augmented in projections forward through time to accommodate the maturation of dependent children, and the inflow of international immigrants. The decision unit is the benefit unit, defined as a single adult or partner couple and their dependent children. The model can be used to consider endogenous decisions regarding consumption, labour supply of adult benefit unit members, and the portfolio allocation across a range of assets that include safe and risky liquid investments, Individual Savings Accounts[8], and personal pensions. The model assumes by default that decisions maximise expected lifetime utility, given a benefit unit's prevailing circumstances, its preference relation, and beliefs regarding the future. The model can also be directed to use reduced-form equations in place of utility maximisation for a selected set of decisions. The following circumstances of a reference adult may all be projected through time:

- year of birth (constant)
- age
- relationship status (single/couple)
- number and age of dependent children
- student status
- education
- health status
- carer status
- migration status
- self-employed / public-sector / private-sector employee if in work
- labour income earned if reference adult works
- labour income earned if spouse works

[8]Individual Savings Accounts (ISAs) are an investment product available in the UK since 1999 in which interest, dividends and capital gains are tax-free.

- savings held in Individual Savings Account
- own business wealth
- eligibility for alternative private pension schemes
- private pension wealth
- timing of access to private pension wealth
- a contributory state pension, modelled on the UK Basic State Pension
- a contributory state pension, modelled on the UK State Second Pension
- benefit unit wealth not otherwise defined
- time of death

Of the 21 characteristics listed above, eight are modelled exogenously (year of birth, age, relationship status, number and age of dependent children, student status, education, health status, carer status, employment sector, time of death) and the remainder are endogenous. Furthermore, four of the characteristics are assumed by the model to evolve non-stochastically (year of birth, age, contributory state pensions, and the timing of access to private pension wealth) and all others may be uncertain. Tables describing the variables simulated by the model are provided in Appendix B. The preference relation assumed by the model takes an additively separable nested Constant Elasticity of Substitution form that allows for quasi-hyperbolic discounting (on which see Section 3). Expectations are calculated assuming beliefs are 'substantively rational', in a way that is explained below.

2.1 Basic mechanics of the model

Like most Dynamic Programming (DP) models of behaviour, LINDA uses a two-stage process to project the circumstances of a population through time. In the first stage, numerical methods are used to solve for utility maximising decisions (where these are requested) given any feasible combination of individual specific circumstances. The second stage uses the behavioural solutions identified in the first stage, or reduced-form descriptions of behaviour, in conjunction with assumed relationships governing the intertemporal variation of individual-specific circumstances, to generate panel data for the simulated population. These panel data typically form the basis for conducting secondary analyses that are the principal focus of concern. Such analyses include the distributional implications of policy measured over alternative time horizons, behavioural responses to the policy environment, and empirical evaluation of behavioural assumptions.

DP models of behaviour are now sufficiently well known that only a broad outline of the steps involved is given here; see, e.g., Rust (2008) for technical detail.

The first stage: Solving for utility maximising decisions

An analytical solution to the utility maximisation problem assumed for LINDA does not exist, and numerical solution routines are consequently employed. These solution routines are structured around a 'grid' that over-lays all feasible combinations of the individual-specific characteristics (known as the state space, and detailed in the above list). One of these dimensions is age, and another is birth year, each of which is divided into discrete intervals. Time is implicit, as it is a linear combination of age and birth year, an issue that is returned to in Section 2.2. LINDA assumes that there is a maximum potential age to which any individual may survive, denoted by A. The assumption of an upper bound on age and the organisation of time into discrete intervals implies the existence of a 'final period', after which death is certain.[9] In this final period of life, the decision problem is non-dynamic, and is therefore straight-forward to solve. LINDA begins by solving for utility maximising decisions at all intersections of the grid in this final period of life, and stores both the maximising decisions and optimised measures of utility (referred to as the value function).

Having obtained utility maximising solutions for the final period of life, the model then addresses decisions at intersections corresponding to the penultimate period. Here, utility depends upon decisions (e.g. consumption) in the penultimate period, and the impact that decisions in the penultimate period have on circumstances (e.g. wealth) – and utility – in the final period (A). The second of these two factors is evaluated with reference to the optimised measures of utility stored in the solution for the final period of life. Where a given decision alternative in the penultimate period implies a combination of characteristics that corresponds precisely to a grid intersection in the final period, then the associated measure of next-period optimised utility is trivial to identify. In the more general case, where a given decision alternative in the penultimate period implies a combination of characteristics that does not correspond to a grid co-ordinate in the final period, then interpolation methods are used to approximate the value function in the final period by drawing on solutions at 'near-by' grid points.

A further complicating issue arises when, given any feasible decision alternative in the penultimate period, the combination of characteristics in the final period is uncertain. In context of a discrete set of potential alternative state combinations, the assumption of von Neuman Morgenstern preferences permits measures of expected utility to be evaluated as weighted sums. LINDA also allows for margins of uncertainty that are (log) normally distributed. In this case, expectations are evaluated with reference to a discrete set of abscissae, weighted using the Gauss-Hermite quadrature.

[9] Assumed to be age 131 in current model applications.

The above routines allow expected utility in the final period of life to be evaluated for any given decision alternative in the penultimate period. Numerical search routines are used to identify decision combinations that maximise expected lifetime utility at all intersections of the grid that correspond to the penultimate period of life. These maximising decisions and the associated measures of utility are stored by LINDA, and the solution to the lifetime decision problem then proceeds recursively to all earlier periods of life. A brief example of this approach is provided in Appendix C.

The second stage: Simulating a population through time

Having solved for utility maximising behavioural responses at grid nodes as described above, the life-courses of individual benefit units are simulated by 'running them through the grids'.

Model projections begin with detailed micro-data that describe the characteristics (state variables) of a cross-sectional sample of adults and their benefit units observed at a given point in time (the 'reference period'). All adults are represented, including those in co-habitating relationships, so that some circumstances for couples are represented twice (once for each spouse) in the simulated data. The model starts by identifying a single 'reference adult' for each benefit unit. In the case of singles, the reference adult is simply the adult benefit unit member. In the case of couples, if one spouse is identified as previously accessing their pension wealth but the other has not, then the reference adult is set to the spouse who has previously accessed their pension. Otherwise, the reference adult is defined as the spouse with the highest wage potential. Model projections for each benefit unit then proceed with reference to the characteristics of the respective reference adult.

Having loaded in data for the reference cross-section, the model projects characteristics that are exogenous of all decisions that refer to solutions to the utility maximisation problem. These 'exogenous' characteristics include: interest rates, age, time of death, education status, health status, relationship status, age and number of dependent children, disability status of dependent children, carer status, and employment sector. Simulation of exogenous characteristics is relatively fast, and the model consequently simulates the full set, saving only data for characteristics that have been requested for analysis. Where an exogenous characteristic is omitted from the requested simulation, the model continues to save 'working variables' (e.g. simulation random draws), to facilitate replication of results in contexts where alternative sets of characteristics are included for analysis.

Characteristics that are not exogenous are simulated jointly by the model. The micro-data that are fed into the model allow the benefit unit of each reference adult to be located within the grid structure that is described in the preceding subsection. Any decisions that are requested to be generated on the basis of solutions to the utility maximisation problem are read off from the grids, where interpolation methods are employed to identify decisions when circumstances do not lie precisely on a considered grid point. Given each adult's characteristics and the behaviour of their respective benefit unit, their characteristics can be aged one period forward, based on the processes that govern each characteristic's intertemporal variation. Where these processes depend upon stochastic terms, random draws are taken from their defined distributions in a process that is common in the microsimulation literature (sometimes referred to as Monte Carlo simulation).

Similar methods are used to project benefit unit circumstances backward through time, subject to two additional complications. First, it is necessary to ensure that projected decisions satisfy incentive compatibility conditions. Suppose, for example, that the model starts with a given set of benefit unit characteristics at time t. Given an initial guess concerning the benefit unit's decisions at time $t-1$, similar methods to those used to project circumstances forward through time can be used to identify the benefit unit's characteristics to the start of period $t-1$. It is then necessary to check that the characteristics that have been projected to time $t-1$ for the benefit unit are consistent with the guess made about the benefit unit's decisions at time $t-1$; this is the 'incentive compatibility' element of the problem. If the initial guess concerning the benefit unit's decisions at time $t-1$ are found to be incompatible with the solution obtained to the utility maximisation decision, then a search is performed to ensure incentive compatibility.

Secondly, some characteristics display intertemporal persistence that cannot be captured using a simple rule concerning transitions between adjacent years. Child birth, for example, is generally more prevalent among couples than singles, the birth of a child can be expected to influence subsequent transition rates for relationship formation/dissolution, and a child generally remains a dependent of (at least) one parent for many years. These inter-relationships all complicate joint projections of marriage and dependent children backward through time. Similarly decisions concerning the timing of access to pension wealth are made once and are persistent thereafter, which complicates the search routine required to ensure incentive compatibility that is referred to above. In cases where intertemporal persistence complicates backward projections, the model uses search routines to identify projections that are consistent with the set of characteristics observed in the reference cross-section.

2.2 Projecting the population cross-section through time

Variation between the conditions faced by different individuals is commonly decomposed into time, cohort, and age effects. DP models of savings and labour supply typically focus upon the evolving circumstances of individual birth cohorts, in which case the cohort effect is the same for all treated individuals, and time and age can be reduced to a single state (characteristic) due to the linear dependence that exists between the three (time = age + birth year). Extending a traditional (birth) cohort specific DP model to project a population cross-section through time consequently requires one additional dimension to be included in the state space of the decision problem. This dimension is referred to here as the *birth year*.

Differences by birth year have been integrated into almost every aspect of the model structure, reflecting important differences between birth cohorts that are reported in survey data. Survival rates have improved substantially for older people during the last four decades, which has important implications for measures of fiscal sustainability and savings adequacy. At the same time, there has been a distinct weakening of domestic partner relationships (including the rise of cohabitation) and falling fertility rates, influencing (equivalised) earnings potential and consumption needs. Set against these sustained demographic trends are the broad range of labour and capital market characteristics that vary over the economic cycle. Chief among these are variations in labour market conditions (including rates of pay and unemployment) and credit market conditions (including returns to capital and the cost of debt). A birth cohort's relative advantage often depends upon the timing of economic up-swings and down-swings during its life course. Similarly, transfer policy has exhibited substantial variation with time, reflecting changes in public attitudes toward the welfare state. These differences can all be explicitly represented in LINDA.

Having defined variation between birth cohorts, it is also necessary to define how the variation is incorporated into the lifetime decision problem. The approach adopted here is designed to reflect the underlying nature of intertemporal evolution of the policy environment. The smooth temporal transitions that have been observed for each of the demographic factors upon which the model depends – survival rates, marriage rates, rates of marital dissolution, and fertility rates – motivates the assumption that individuals exercise perfect foresight over the respective rates to which they will be subject. This is not to say that an individual is assumed to exercise perfect foresight regarding the out-turn of their own circumstances. Although an individual is assumed to be uncertain about the precise timing of their death, for example, they are assumed to forecast with precision the death rate of their respective birth cohort. Such assumptions are standard in the associated DP literature.

It would be inappropriate to apply the same assumption of perfect foresight to the 'economic factors' that are projected by the model – interest accruing to assets / debt, wage rates, transfer policies, and unemployment rates – due to the temporal volatility exhibited by these factors. This proposition is motivated by more than the theoretical inconsistencies that would otherwise arise in relation to agent expectations. From a technical perspective, the interpolation methods that are used to evaluate the position of individuals for whom an explicit solution to the lifetime decision problem is not obtained assume that reference may reasonably be made to 'near-by' individuals (for whom a solution is obtained). This assumption becomes increasingly prescriptive as the volatility between 'near-by' individuals widens. We therefore employ individually tailored methods to simulate each 'economic factor'.

It is assumed that future returns to risky assets and wage rates are fundamentally uncertain, and this uncertainty is explicitly accounted for when evaluating agent expectations (consistent with the associated literature). Accounting for uncertainty in this way is, however, computationally demanding, and we do not therefore extend the approach to interest charges on safe assets and debt, unemployment rates, or transfer policies. Rather, we assume that individuals from all birth cohorts expect that they will be subject to the same interest rates on safe assets and debt, and the same (age/education/health-dependent) rates of unemployment. That is, for example, that all birth cohorts are assumed to expect that they will be subject to the same unemployment rate if they are a graduate of good health aged 43, which could reasonably be set equal to the average unemployment rate for healthy 43 year old graduates reported by survey data. Nevertheless, LINDA projects the population on observed time-varying rates of return and unemployment, in contradiction of the assumption that individuals expect these to remain constant through time. Conceptually, we assume that people may be aware of the temporal variation of unemployment rates, for example, but choose not to take this variation into account when planning for the future. We refer to this alternative approach as 'substantively rational'.

The final economic factor that is allowed to vary between birth cohorts is transfer policy. The influence of transfer policy on benefit unit budgets is comprised of two key components in the model. The first is a highly flexible 'tax and transfer function' that is capable of capturing much of the detail of transfer policy as it is applied in practice. The second is random variation, implemented through a *tax residual*. The tax residual is designed to correct for differences between simulated and sample moments of disposable income, representing measurement error, and differences between the model tax and welfare structure and policy as it was applied. Any policy variation between birth cohorts that works through the tax function is assumed to be fully anticipated when evaluating agent expectations. It might be sensible, for example, to assume that agent expectations take into account planned increases in the State Pension Age, or trend growth in income tax thresholds. In contrast, individuals are assumed to take no account of any effects that the tax residual may have on their circumstances when evaluating expected lifetime utility. Hence, taxes are simulated in a way that falls somewhere between the approaches adopted for demographic factors on the one hand (where variation between birth cohorts is fully anticipated), and unemployment rates on the other (where variation between birth cohorts – if it is accommodated – is unanticipated).

3 The Preference Relation

Expected lifetime utility of reference adult i, with birth year b, at age a is described by the time separable function:

$$
\begin{aligned}
U_{i,a} &= \frac{1}{1-\gamma}\Bigg\{u\left(\frac{c_{i,a}}{\theta_{i,a}}, l_{i,a}\right)^{1-\gamma} + \Delta_{i,a} + \\
&+E_{a,b}\Bigg[\beta_1\beta_0\left(\phi^b_{1,a}u\left(\frac{c_{i,a+1}}{\theta_{i,a+1}}, l_{i,a+1}\right)^{1-\gamma} + (1-\phi^b_{1,a})\zeta_1\left(\zeta_0 + w^+_{i,a+1}\right)^{1-\gamma}\right) + \\
&+\beta_1\beta_2\sum_{j=a+2}^{A}\beta_0^{j-a}\left(\phi^b_{j-a,a}u\left(\frac{c_{i,j}}{\theta_{i,j}}, l_{i,j}\right)^{1-\gamma} + (1-\phi^b_{j-a,a})\zeta_1\left(\zeta_0 + w^+_{i,j}\right)^{1-\gamma}\right)\Bigg]\Bigg\} \qquad (1)
\end{aligned}
$$

where $\gamma > 0$ is the (constant) coefficient of relative risk aversion; $E_{a,b}$ is the expectations operator for any individual aged a from birth year b; A is the maximum potential age; β_0, β_1, and are discount factors; $\phi^b_{j-a,a}$ is the probability of someone from birth year b living to age j, given survival to age a; $c_{i,a} \in R^+$ is discretionary composite non-durable consumption of the benefit unit of individual i at age a; $l_{i,a} \in [0,1]$ is the proportion of benefit unit time spent in leisure; $\theta_{i,a} \in R^+$ is the benefit unit's adult equivalent size based on the "revised" or "modified" OECD scale; $\Delta_{i,a}$ represents the influence of decision costs on utility; the parameters ζ_0 and ζ_1 reflect the "warm-glow" model of bequests; and $w^+_{i,a} \in R^+$ is liquid net wealth when this is positive and zero otherwise.

The labour supply decision (if it is included in the model) is assumed to be made between discrete alternatives. No upper limit is imposed on the number of discrete alternatives, so that the labour decision can made to approach a continuous margin.[10] Where adults are explicit, then a separate labour supply decision is allowed for each adult. Otherwise, modelled benefit units choose labour supply in a similar fashion to single adults. Where health is explicit, then labour supply options can be constrained to reflect work-limiting conditions and carer responsibilities. We return to discuss labour supply in Section 4.

The modified OECD scale assigns a value of 1.0 to the benefit unit reference person, 0.5 to each additional benefit unit member over age 13, and 0.3 to each child aged 13 and under. This scale is currently the standard for adjusting incomes in European Union countries. It is included in the preference relation to reflect the important influence that benefit unit size has been found to have on the timing of consumption (e.g. Attanasio & Weber (1995) and Blundell et al. (1994)).[11] Similarly, decision costs are included in the preference relation to allow the model to reflect behavioural rigidities that have been cited as important for understanding retirement savings decisions (e.g. Choi et al. (2003)

[10] The search routine used to identify utility maximising labour supply decisions searches over all feasible employment alternatives, implying that increasing employment options can substantively increase computation times.

[11] An empirical study by Fernandez-Villaverde & Krueger (2006) of US data from the Consumer Expenditure Survey suggests that roughly half of the variation observed for lifetime household consumption can be explained by changes in household size, as described by equivalence scales. See Balcer & Sadka (1986) and Muellbauer & van de Ven (2004) on the use of this form of adjustment for household size in the utility function.

and Carroll et al. (2009) for the US and McKay (2006) for the UK). These costs are accommodated by reducing the value of Δ whenever behaviour deviates from pre-assigned default options in relation to private pensions and Individual Savings Accounts.

The model incorporates an allowance for behavioural myopia, through its assumption of quasi-hyperbolic preferences following Laibson (1997). Such preferences are interesting because they are time inconsistent, giving rise to the potential for "conflict between the preferences of different intertemporal selves" (Diamond & Köszegi (2003), p. 1840). Furthermore, the model assumes that all discount parameters are the same for all individuals, and are time invariant. It also assumes that benefit units are aware of any time inconsistency that their preferences display, a condition sometimes referred to as 'sophisticated myopia'. These limitations rule out a number of interesting behavioural phenomena, including the capacity of the model to reflect systematic population heterogeneity with respect to temporal biases (e.g. Gustman & Steinmeier (2005)), and procrastination (e.g. O'Donoghue & Rabin (1999)). Such effects could be accommodated without a qualitative increase in computational burden. Nevertheless, they are omitted here because the limited empirical analysis that we have conducted has failed to reveal important behavioural margins that such effects would help to explain. This is one principal research thread that we hope to pursue during the next few years.

The warm-glow model of bequests simplifies the utility maximisation problem, relative to alternatives that have been considered in the literature.[12] Including a bequest motive in the model raises the natural counter-party question of who receives the legacies that are left. We return to this issue in Section 6.

A Constant Elasticity of Substitution function was selected for within period utility,

$$u\left(\frac{c_{i,a}}{\theta_{i,a}}, l_{i,a}\right) = \left(\left(\frac{c_{i,a}}{\theta_{i,a}}\right)^{(1-1/\varepsilon)} + \alpha^{1/\varepsilon} l_{i,a}^{(1-1/\varepsilon)}\right)^{\frac{1}{1-1/\varepsilon}} \tag{2}$$

where $\varepsilon > 0$ is the elasticity of substitution between equivalised consumption $(c_{i,a}/\theta_{i,a})$ and leisure $(l_{i,a})$ within each year. The constant $\alpha > 0$ is referred to as the utility price of leisure. The specification of intertemporal preferences described by equations (1) and (2) is standard in the literature, despite the contention that is associated with the assumption of time separability (see Deaton & Muellbauer (1980), pp. 124-125, or Hicks (1939), p. 261). This specification of preferences implicitly assumes that characteristics which affect utility, but are not explicitly stated, enter the utility function in an additive way.

[12] See, for example, Andreoni (1989) for details regarding the warm-glow model.

4 Labour Income Dynamics

Earnings are modelled at the benefit unit level, and are described by:

$$\begin{aligned} g_{i,a} &= \max\left(h_{i,a}, h_{a,t}^{\min}\right)\lambda_{i,a} \\ \lambda_{i,a} &= \lambda_{i,a}^{o}\lambda_{i,a}^{emp}\lambda_{i,a}^{ret}\lambda_{i,a}^{ng}\lambda_{i,a}^{hlth} \end{aligned} \tag{3}$$

where $h_{i,a}$ defines benefit unit i's latent wage at age a, $h_{a,t}^{\min}$ is the (statutory) minimum wage, λ^{o} is an adjustment factor to allow for uncertain wage offers, λ^{emp} adjusts for (endogenous) labour supply decisions, λ^{ret} is the impact on earnings of taking up private pension income, $\lambda_{i,a}^{ng}$ is a factor to account for wage premia accruing to alternative (sub-tertiary) education levels, and $\lambda_{i,a}^{hlth}$ is a factor to account for health status. Each of these alternative factors is described below.

Latent wages, h

In most periods, latent wages are assumed to follow the stochastic process described by the equation:

$$\log\left(\frac{h_{i,a}}{m_{i,a}}\right) = \psi_{i,a-1}\log\left(\frac{h_{i,a-1}}{m_{i,a-1}}\right) + \kappa_{i,a-1}\frac{(1-l_{i,a-1})}{(1-l_{ft})} + \omega_{i,a-1} \tag{4a}$$

$$m_{i,a} = m\left(n_{i,a}, ed_{i,a}, sctr_{i,a}, a, b\right) \tag{4b}$$

$$\psi_{i,a} = \psi\left(n_{i,a}\right) \tag{4c}$$

$$\kappa_{i,a} = \kappa\left(n_{i,a}, a\right) \tag{4d}$$

$$\omega_{i,a} \sim N\left(0, \sigma_{\omega}^{2}\left(n_{i,a}, ed_{i,a}, sctr_{i,a}\right)\right) \tag{4e}$$

where the parameters $m(.)$ account for wage growth, which in turn depend on relationship status $n_{i,a}$, education $ed_{i,a}$, self-employment / public / private sector state $sctr_{i,a}$, age a, and birth year b. $\psi(.)$ accounts for time persistence in earnings, $\kappa(.)$ is the return to another period of experience, $(1-l_{ft})$ is the leisure cost of full-time employment by all adult benefit unit members, and $\omega_{i,a}$ is an identically and independently distributed benefit unit specific disturbance term. The variance σ_{ω}^{2} is defined as a function of relationship, education, and employment sector. The only exceptions to equation (4a) are when a reference adult changes their education status (see Section 10), or changes employment sector. In these alternate cases, a new random draw is taken from a log-normal distribution, the mean and variance of which are specific to the benefit unit's age, birth year, relationship, education, and the sector into which they have recently entered.

The form of equation (4a) has a number of desirable properties that motivated its selection. It is a parsimonious wage specification that has been explored at length in the literature (e.g. Sefton & van de Ven (2004)). It requires the addition of just two state variables to the decision problem (h, ω), only one of which is uncertain (ω). The appearance of the $m(.)$ terms on both sides of equation (4a) helps to simplify parameterisation of the model. Increasing $m_{i,a}$ by 3 percentage points, for example, will *ceteris paribus* increase $h_{i,a}$ by 3 percentage points without also feeding through to increase $h_{i,a+1}$ (a property that is lost if the $m(.)$ terms are replaced by a single factor on the right-hand-side of equation 4a). Finally, we have found that the addition of an experience effect to the wage equation can help to match the model to the age profile of labour supply (e.g. Sefton & van de Ven (2004)). Increasing the experience effect acts to increase the cost of leisure early in the working lifetime, which off-sets the low instantaneous wages that are often observed to accrue to young workers.

Minimum wage, $h^{\min}$

The minimum wage allows for a floor to be imposed, with reference to the hourly wage rate. This floor is specified so that it can differ relative to four age thresholds. Each age-specific minimum wage rate can be defined to growth through time at different rates defined by the user.

Wage offers, λ^o

Wage offers are included in the model to allow for the possibility of (involuntary) unemployment among employees (self-employed are unaffected), which we have found to be important in matching the model to rates of employment during peak working years. Separate wage offers can be allowed for both the reference adult and their spouse (if one exists). Receipt of a wage offer is modelled as uncertain between one period and the next, subject to age, education, health, and relationship specific probabilities $p^o\left(n_{i,a}, ed_{i,a}, health^j_{i,a}, a\right)$. If a wage offer is received by an individual, $\lambda^o_{i,a} = 1$, then benefit unit income responds to their labour supply decision. If a wage offer is not received by an individual, $\lambda^o_{i,a} = 0$, then any labour that the respective individual supplies returns no labour income to the benefit unit, implying non-employment where working incurs a leisure penalty.[13]

As discussed in Section 2.2, the solution to the lifetime decision problem assumes that benefit units expect that the probability of a low wage offer is age, relationship, health, and education specific, but is time invariant (as p^o is defined above). When a population is simulated through time, however, allowance is made for historical variation in unemployment rates to reflect observed fluctuations through the economic cycle.

Employment, λ^{emp}

Each discrete labour alternative $l_{i,a}$, which can vary by carer status, $carer_{i,a}$, is associated with its own factor, $\lambda^{emp}(l_{i,a}, carer_{i,a})$. It is usual to define λ^{emp} to be an increasing function of labour supply, and the factor is scaled so that full-time employment of all adult members implies $\lambda^{emp} = 1$. It is assumed

[13] It is assumed that the disutility from a year of employment is more than sufficient to off-set the experience effect on latent wages.

that the benefit unit reference person has the highest wage potential of any adult in the benefit unit. The relationship described by $\lambda^{emp}(l_{i,a})$ permits each adult's share of benefit unit labour income, $g^j_{i,a}$, to be evaluated from total benefit unit labour income: $g^j_{i,a} = \lambda^{emp}\left(l^j_{i,a}\right)/\lambda^{emp}(l_{i,a}).g_{i,a}$.

Pension take-up, λ^{ret}

It is possible to impose wage penalties on benefit units that have started to draw upon their private pension wealth. This is allowed for in the model through the addition of a fixed factor adjustment applied to the benefit unit's latent wage, $\lambda^{ret}_{i,a} < 1$ if the benefit unit has accessed their pension wealth.

Non-tertiary education, λ^{ng}

The model can account for up to five alternative education states for each simulated adult, referring to the highest qualification held (see Section 10). The influence on wages of educational differences between those with and without tertiary education are managed through the evolution of latent wages and wage offers, as described by the entry of $ed_{i,a}$ variable in equation (4a) and p^o. In contrast, all individuals with sub-tertiary education are assumed to be subject to the same relationship governing intertemporal evolution of latent wages. The influence of different sub-tertiary education states on wages is accommodated both through variation in the probability of low wage offers – p^o – and through the term $\lambda^{ng}_{i,a}$, which is designed to reflect the wage premia accruing to alternative levels of sub-tertiary education. These wage premia are permitted to vary by age and through time.

Health status, λ^{hlth}

The model can account for up to ten alternative health conditions for each simulated adult. The influence on wages of health is accommodated both through variation in the probability of low wage offers – p^o – and through the term $\lambda^{hlth}_{i,a}$, which is designed to reflect the wage penalties associated with limiting health conditions. These wage penalties are permitted to vary by age and through time.

4.1 Simulating latent wages backward through time

Simulating latent wages backward through time following the process described by equation (4a) is complicated by the implied non-zero covariance between $\omega_{i,a-1}$ and $h_{i,a}$. It is necessary to take this correlation into account to maintain increasing variances with age, which are commonly evident in survey data. To do this, we assume a linear regression specification between ω_{a-1} and $\log(h_a)$, so that

$h_{i,a-1}$ is given by:

$$\log(h_{a-1}) = \log(m_{a-1}) + \frac{\log\left(\frac{h_a}{m_a}\right) - \left(\kappa_{a-1}\frac{(1-l_{a-1})}{(1-l_W)} + \hat{\omega}_{a-1}\right)}{\psi_{a-1}} \tag{5a}$$

$$\hat{\omega}_{a-1} = \frac{\sigma_\omega^2}{\sigma_a^2}(\log(h_a) - \mu_a) + \eta_{a-1} \tag{5b}$$

$$\eta_{a-1} \sim N\left(0, \frac{\sigma_\omega^2\left(\sigma_a^2 - \sigma_\omega^2\right)}{\sigma_a^2}\right) \tag{5c}$$

where μ_a and σ_a^2 are the mean and variance of log latent wages at age a, $\sigma_a^2 \geq \sigma_\omega^2$, and individual specific indicators have been dropped for simplicity. This specification makes use of the observation that the covariance between ω_{a-1} and $\log(h_a)$, $cov(\omega_{a-1}, h_a) = \sigma_\omega^2$. If experience effects are omitted from the analysis ($\kappa = 0$), then equation (4a) describes a standard regression-toward-the-mean model of earnings (e.g. Creedy (1985)), and we can set $\mu_a = \log(m_a)$. Substituting this restriction into equation (5) and re-arranging, we obtain:

$$\log(h_{a-1}) - \mu_{a-1} = \frac{\left(1 - \sigma_\omega^2/\sigma_a^2\right)(\log(h_a) - \mu_a) - \eta_{a-1}}{\psi_{a-1}}$$

which clarifies the dampening influence that allowing for $cov(\omega_{a-1}, h_a) = \sigma_\omega^2$ has on the dispersion of h_{a-1}.

5 Employment Sector

Three employment sectors can be distinguished in the model: self-employment, public sector employees, and private sector employees. Employment sector is modelled at the benefit unit level, in common with the approach taken to simulate labour incomes. In each simulated period, reference adults of working age and their spouses are jointly assigned an identifier, $sctr_{i,a} = \{0, 1, 2\}$, equal to 0 for self-employed, 1 for private sector employee, and 2 for public sector employee. Any employment that adult benefit unit members engage in during a period is assumed to be in the relevant sector of employment. Transitions between employment sectors are uncertain from one period to the next, with the probability of change stored either in the form of transition matrices or a multinomial logit equation. These probabilities are designed to respond to the reference adult's age, prevailing employment sector, and (in the case of multinomial logit specification) the year of observation.

Individuals in different employment sectors are treated identically, with four exceptions. First, as discussed in Section 4, each of the three employment sectors has its own set of parameters governing wage growth, m, and volatility, ω. Secondly, whereas employees in the private and public sectors may be subject to the risk of a low wage offer, self-employed are not. Thirdly, if private pensions are included for analysis, then individuals in different employment sectors can be subject to different terms regarding private and employer pension contribution rates. These pensions are described in Section 9. Finally, a separate "own business" asset can be included for the self-employed in the model.

5.1 Own business wealth

If an own-business asset is included for analysis, then a fixed proportion of benefit unit earnings during each period of self-employment, π^{ob}, is assumed to be invested in the asset, paid out of pre-tax income. Own-business assets attract a rate of return r_t^{ob} that can be uncertain. When returns are uncertain, then they are assumed to be perfectly correlated with returns to risky liquid assets r^r. Own business wealth w^{ob} consequently evolves following:

$$\begin{aligned} w_{i,a}^{ob} &= r_{t-1}^{ob} w_{i,a-1}^{ob} + \pi^{ob} g_{i,a-1}^{ob} + k_{i,a}^{ob} \\ \ln\left(r_t^{ob}\right) &\sim N\left(\mu_{ob} - \frac{\sigma_{ob}^2}{2}, \sigma_{ob}^2\right), corr(r_t^{ob}, r_t^r) = 1 \end{aligned} \tag{6}$$

where $g_{i,a}^{ob} = g_{i,a}$ if $sctr_{i,a} = 0$ and zero otherwise, $corr\,(.)$ denotes the correlation coefficient, and $k_{i,a}^{ob}$ represents capital-transfers into the own-business asset (negative when there are out-flows). The only departure from equation (6) is when a reference adult is identified as suffering a divorce, in which case a fixed factor adjustment is used to divide the own-business wealth between the reference adult and their spouse.

The assumption that own business wealth is illiquid during the period of self-employment requires $k_{i,a}^{ob} = 0$ for all periods other than those in which the benefit unit transitions out of self-employment. Denote $k_{i,a}^{end} = k_{i,a}^{ob}$ when a benefit unit transitions out of self-employment ($sctr_{i,a} = 0; sctr_{i,a+1} > 0$), and $k_{i,a}^{begin} = k_{i,a}^{ob}$ when it transitions into self-employment ($sctr_{i,a} > 0; sctr_{i,a+1} = 0$). It is useful to assume that all newly self-employed benefit units hold zero net equity in their own businesses, $w_{i,a}^{ob} = k_{i,a}^{begin} = 0$, as this ensures that any wealth held in the own business asset at the end of a self-employment episode is assessable for capital gains tax. At the end of an episode of self-employment, all own-business assets are converted into liquid net wealth; $k_{i,a}^{end} = -\left(r_{t-1}^{ob} w_{i,a-1}^{ob} + \pi^{ob} g_{i,a-1}^{ob}\right)$. An important distinction between the forward and backward projections for own-business assets is that the value of $k_{i,a}^{end}$ is derived as a product of simulating benefit unit circumstances between the reference period and age a in the forward projections, but must be imputed for the backward projections.

The backward projections impute $k_{i,a}^{end}$ in a way that is designed so that $k_{i,a-T^{ob}}^{begin} \simeq 0$, where T^{ob} is the duration of the relevant episode of self-employment. To do this, the model takes advantage of the fact that the self-employment state of each benefit unit is simulated exogenously (see Section 2.1), which permits T^{ob} to be evaluated when a benefit unit first enters a self-employment episode in the backward projections. The model then approximates the aggregate value of own business wealth that would be held at the end of the episode if own business wealth were zero at the beginning of the episode, $k^{end}\left(k^{begin}=0\right)$, by multiplying the time spent in the episode by the benefit unit's prevailing latent wage and the (exogenous) accrual rate; $\hat{k}_{i,a}^{end}=T^{ob}\pi^{ob}g_{i,a-1}^{ob}$. In this case, $k_{i,a}^{end}$ will tend to over-predict $k^{end}\left(k^{begin}=0\right)$ to the extent that latent wages grow during the period of self-employment, and where the earnings potential is limited by the receipt of low wage offers or less than full take-up of employment. In contrast, $\hat{k}_{i,a}^{end}$ will tend to under-predict $k^{end}\left(k^{begin}=0\right)$ to the extent that it omits associated investment returns and fails to reflect any decline in the latent wage during the period of self-employment. Any mis-match between $\hat{k}_{i,a}^{end}$ and $k^{end}\left(k^{begin}=0\right)$ is ignored for the purposes of simulating capital gains tax in the backward projections.

6 The Budget Constraint

Equation (1) is maximised, subject to an age specific credit constraint imposed on liquid net wealth, $w_{i,a}\geq D_a$ for benefit unit i at age a.[14] In context of income uncertainty, and a preference relation where marginal utility approaches infinity as consumption tends toward zero, rational individuals will never choose to take on debt equal to or greater than the discounted present value of the minimum potential future income stream that they face (however unlikely that stream might be). This rule is used to define D_a, subject to the additional constraint that all debts be repaid by age $a_D\leq A$.[15] Intertemporal variation of $w_{i,a}$ is, in most periods, described by the simple accounting identity:

$$w_{i,a}=w_{i,a-1}+\tau_{i,a-1}+ur_{i,a-1}^{h}-c_{i,a-1}-ndc_{i,a}^{x}+k_{i,a-1}+B_{i,a-1} \tag{7}$$

where τ denotes disposable income, ur^h is un-realised returns to owner-occupied housing, c is discretionary non-durable composite consumption, ndc^x is non-discretionary expenditure, k represents net investment flows with other asset classes (i.e. own business assets and Individual Savings Accounts), and $B_{i,a-1}$ is the value of bequests received.

Non-discretionary costs are included to reflect the view that there exists a basket of goods and services that represent the basic necessities of life, and which individuals therefore exercise little discretion concerning purchase (sometimes referred to as "committed expenditure"). Non-discretionary costs are disaggregated into child care, housing (rent and mortgage interest), health, and 'other' categories of expenditure to facilitate simulation of welfare benefits that make explicit reference to any one of these expenditure categories. Simulated child care costs, ndc^c, are described as a function of the number and age of dependent children, and of the employment status of the least employed adult benefit unit member. Non-discretionary health costs, ndc^h, can either be set equal to the value of dedicated welfare

[14] Note that $w_{i,a}^{+}$ referred to above is related to $w_{i,a}$, with $w_{i,a}^{+}=0$ if $w_{i,a}<0$, and $w_{i,a}^{+}=w_{i,a}$ otherwise.

[15] The lower bound D_a is assumed to be the same for all households, to simplify the interpolation routines that evaluate over variable birth years. Interpolation methods are discussed in Section 14.1.

benefits generated in respect of recognised health conditions, or to the value of exogenously supplied model parameters. Non-discretionary housing expenditure is comprised of rent and mortgage payments, $ndc^{hg} = rent + mort$, and is described in Section 6.2. 'Other' non-discretionary expenditure, ndc^{o}, is loosely designed to reflect the minimum expenditure required to participate in society, consistent with standard definitions of poverty. Consumption on other basic necessities is defined in terms of equivalised (non-housing / non-child care / non-health) consumption, and can vary by age and year.

The only potential departures from equation (7) occur when a benefit unit is identified as accessing pension wealth, or when a reference adult is identified as getting married or incurring a marital dissolution. Wealth effects at the time a benefit unit accesses its pension wealth are discussed in Section 9. In relation to marital transitions, backward projections assume that spouses are identical clones (see Section 11.1 for discussion), so that wealth is halved in context of a dissolution and doubled in context of a formation. In forward projections, spouses are identified from within the simulated sample. A marriage between two simulated singles consequently results in the liquid net wealth of each being combined in the common benefit unit. A divorce is assumed to see liquid net wealth split evenly between each divorcee, whereas widowhood sees all liquid net wealth bequeathed to the surviving spouse.

The methods used to simulate transfer policy and to project returns to liquid net wealth are now described, before describing two incentive compatibility problems that are introduced when wealth is projected backward through time and how each of these problems is accommodated by the model.[16]

6.1 Simulated transfer policy

As the model has been designed to undertake public policy analysis, particular care was taken concerning formulation of the module that simulates the effects of taxes and benefits. The model allows the measures of income accruing to each adult benefit unit member to be accounted for separately, so that it can reflect the taxation of individual incomes that is applied in the UK. The tax function assumed for the model is represented by:

$$\tau_{i,a} = \tau \left(\begin{array}{c} b, a, n_{i,a}, n^{c}_{i,a}, health^{j}_{i,a}, health^{c}_{i,a}, care^{j}_{i,a}, sctr_{i,a}, l^{j}_{i,a}, g^{j}_{i,a}, \\ hh_{i,a}, mh_{i,a}, w^{h}_{i,a}, rent_{i,a},, mort_{i,a}, rr^{h}_{i,a}, w^{nh,j}_{i,a}, r^{nh}_{i,a} w^{nh,j}_{i,a}, \\ pc^{c,j}_{i,a}, pc^{nc,j}_{i,a}, py^{j}_{i,a}, k^{ISA}_{i,a}, w^{ISA}_{i,a}, r^{ISA}_{t} w^{ISA}_{i,a}, ndc^{c,h}_{i,a}, \eta_{b,a} \end{array} \right) \tag{8}$$

which depends on the birth year of the reference adult b; age of the reference adult, a; number of adults (relationship status), $n_{i,a}$; number and age of all dependent children, represented by the vector $n^{c}_{i,a}$; health status of each adult j in the benefit unit, $health^{j}_{i,a}$; health status of each child, $health^{c}_{i,a}$; carer status of each adult, $care^{j}_{i,a}$; employment sector of the benefit unit during the period, $sctr_{i,a}$; labour supply of each adult, $l^{j}_{i,a}$; the labour income of each adult, $g^{j}_{i,a}$; indicator variables for home-

[16]The methods used to project wealth forward through time have been thoroughly discussed in the dynamic microsimulation literature, and are not restated here.

owners, $hh_{i,a}$, and mortgage holders, $mh_{i,a}$; the net owner-occupied housing wealth held by the benefit unit, $w^h_{i,a}$; the rent paid by non-home-owners, $rent_{i,a}$; the mortgage interest paid by mortgage holders, $mort_{i,a}$; the realised returns to (gross) housing wealth, rr^h; the non-housing net liquid wealth held by each adult in the benefit unit, $w^{nh,j}_{i,a}$; the investment return on liquid net wealth of each adult in the benefit unit, $r^{nh}_{i,a}w^{nh,j}_{i,a}$ (which may be negative); the concessional and non-concessional pension contributions made by each adult in the benefit unit, $pc^{(n)c,j}_{i,a}$; the (retirement) pension income received by each adult in the benefit unit, $py^j_{i,a}$; net contributions to Individual Savings Accounts made during the prevailing year, $k^{ISA}_{i,a}$ (which may be negative); the wealth held in Individual Savings Accounts by the benefit unit, $w^{ISA}_{i,a}$; the income earned on savings in ISAs, $r^{ISA}_t w^{ISA}_{i,a}$; non-discretionary child care costs, $ndc^c_{i,a}$; non-discretionary health costs $ndc^h_{i,a}$; and the tax residual, $\eta_{b,a}$. All of the inputs to the tax function are described in other subsections of this paper.[17]

The form of the tax function described by equation (8) was selected to minimise the computational burden of the utility maximisation problem. Calculating taxes with respect to wealth held at the beginning of a period (as it is here) implies that disposable income is made independent of consumption. This is advantageous when consumption is a choice variable, as it implies that the numerical routines that search for utility maximising values of consumption do not need to evaluate disposable income for each consumption alternative that is tested.

Projecting wealth backward through time introduces two principal complications are introduced when a model is structured to project circumstances backward through time. The first is the incentive compatibility problem that is referred to in the immediately preceding subsection. The second is the difficulty associated with match up random innovations that are assumed to apply through an extensive period of the life course, with individual specific circumstances observed at some advanced age. These are two themes that are returned to in relation to backward projections throughout the paper.

6.2 Dis-aggregating liquid net wealth

Liquid net wealth includes all assets other than those that are otherwise explicitly represented in the model. This composite asset is divided into three sub-categories by the model: net wealth held in owner occupied housing $w^h_{i,a} \in [0, \infty)$; non-housing risky assets $w^r_{i,a} \in [0, \infty)$; and non-housing safe assets $w^s_{i,a} \in [D, \infty)$; $w_{i,a} = w^h_{i,a} + w^s_{i,a} + w^r_{i,a}$.

Given a measure of liquid net wealth, the model begins by distinguishing housing from non-housing wealth $\left(w^{nh} = w^s + w^r\right)$. Although formal modelling of housing investment decisions is analytically feasible (see, e.g. Attanasio et al. (2012)), it is also computationally burdensome. Computational feasibility of the model is maintained by adopting an exogenous procedure for identifying the following

[17] See Appendix C for details.

housing-related features: home owners (hh), mortgage holders (mh), net housing wealth $\left(w^h\right)$, mortgage debt $\left(md^h\right)$, gross housing wealth $\left(w^h + md^h\right)$, realised return on gross housing equity $\left(rr^h\right)$, unrealised return on gross housing equity $\left(ur^h\right)$, mortgage interest costs ($mort$), and rent ($rent$). The procedure adopted to identify these characteristics is comprised of 12 steps.

1. Identify home-owners, hh: All benefit units with non-pension/non-business wealth $\left(w_{i,a} + w_{i,a}^{ISA}\right)$ under a threshold are assumed to not be home owners; all benefit units with non-pension/non-business wealth above a threshold are assumed to be home-owners, and those with wealth between these two thresholds are identified using a logit equation that refers to their non-pension/non-business wealth, marital status, and the age of the reference person. The assumed thresholds grow at a fixed rate through time.

2. If home owner, then go to step 3. Otherwise go to step 11.

3. Identify mortgage holders, mh: mortgage holders identified using a logit equation that refers to each benefit unit's non-pension/non-business wealth, marital status, and the age of the reference person.

4. Identify net housing equity, w^h: An age specific proportion of liquid net wealth held in housing is assumed for all home owners.

5. If mortgage holder, then go to step 6. Otherwise go to step 8.

6. Identify mortgage debt, md: Each mortgage holder's mortgage value, defined as a multiple of non-pension/non-business wealth, is identified using a linear function of log non-pension/non-business wealth. The slope and intercept of this function are allowed to vary between singles and couples, and the multiple is restricted to values between 0 and 20.

7. Identify annual interest charge on mortgage debt, $mort$: The annual interest change on mortgage debt is evaluated by multiplying the (gross) mortgage value identified in step 6 by an (exogenously assumed) fixed rate of mortgage interest.

8. Identify returns to gross housing wealth: Gross housing wealth is assumed to attract an exogenous rate of return. A fixed rate of return is applied for solving the lifetime decision problem, and a year-specific rate is permitted when projecting the population through time (see discussion of 'substantial rationality' in Section 2.2).

9. Identify realised and unrealised returns to gross housing wealth, rr and ur: Total return to gross housing wealth identified in step 8 is disaggregated into realised and unrealised components using an age specific ratio, based on the age of the benefit unit reference person.

10. End
11. Identify rent paid, *rent*: One bedroom is assumed for the reference adult and their spouse (if married). One bedroom is assumed for each dependent child aged 13 or over, and for every two children aged under 13 years. The model parameters include assumed rental charges distinguished by the number of bedrooms, which vary through time following an assumed growth rate.
12. End

The portfolio allocation decision is represented in the model as a choice over the proportion of non-housing liquid net wealth that is invested in risky assets during each year, $\rho_{i,a} \in [0,1]$, and is only possible if non-housing liquid net wealth is positive. The division of non-housing liquid net wealth into safe and risky assets affects only the effective rate of return, $r^{nh}_{i,a}$, and is only evaluated if the associated asset allocation decision is included for analysis. The return to non-housing net liquid wealth is given by $r^{nh}_{i,a} = \rho_{i,a} r^r_t + \left(1 - \rho_{i,a}\right) r^s_{i,a}$, where r^r_t is the return to risky assets, and r^s the return to safe assets/debt. The rate of return to risky assets $\ln\left(r^r_t\right) \sim N\left(\mu_r - \frac{\sigma^2_r}{2}, \sigma^2_r\right)$ is assumed to be the same for all benefit units at any point in time, t. The interest rate on safe liquid assets is assumed to depend upon whether $w^s_{i,a} = \left(1 - \rho_{i,a}\right) w^{nh}_{i,a}$ indicates net investment assets, or net debts. Where $w^s_{i,a}$ is (weakly) positive, then r^s takes the value r^I. When $w^s_{i,a}$ is (strictly) negative then, r^s is designed to vary from r^D_l at low measures of debt to r^D_u when debt exceeds the value of working full time for one period $\left(g^{ft}\right)$:

$$r^s = \begin{cases} r^I & \text{if } w^s \geq 0 \\ r^D_l + \left(r^D_u - r^D_l\right) \min\left\{\frac{-w^s}{g^{ft}}, 1\right\}, r^D_l < r^D_u & \text{if } w^s < 0 \end{cases} \tag{9}$$

Specifying $r^D_l < r^D_u$ reflects a so-called 'soft' credit constraint in which interest charges increase with loan size. As discussed in Section 2.2, the model parameters r^I, r^D_l, and r^D_u take fixed values when solving for utility maximising decisions, and are allowed to vary when simulating the intertemporal evolution of a population.

6.3 Near-term incentive compatability

Near-term incentive compatibility refers to the issue of identifying a decision vector in period a, d_a, that is consistent with any given characteristic vector in the immediately succeeding period, v_{a+1}, and with the solution to the lifetime decision problem. In the current context, this problem centres on the relationship between discretionary non-durable consumption and liquid net wealth. Preliminary analysis revealed that this problem is generally well-behaved, and the model consequently uses a simple recursive routine to solve it. Consider the problem as it is defined above. An initial guess is made, that equates the benefit unit's decisions at age a, represented by the vector d^0_a, to their decisions at age $a+1$: $d^0_a = d_{a+1}$.

Given v_{a+1} and d_a^0, standard simulation methods are employed to identify the benefit unit's implied characteristics at age a, v_a^0. The solution to the dynamic programming problem is then referenced to identify the benefit unit decisions that are consistent with circumstances v_a^0; d_a^1. An adjustment rule is used to identify a revised guess for the benefit unit decision vector $d_a^2 = \hat{\phi} d_a^0 + \left(1 - \hat{\phi}\right) d_a^1$, $0 < \hat{\phi} < 1$, and the process repeated until the absolute difference between d_a^j and d_a^{j-1} is sufficiently small (defined as £0.01 of weekly non-housing discretionary non-durable consumption).

Two key factors ensure that, for the most part, the search problem defined above is well-behaved. First, consumption and wealth lie at the heart of this problem, and the marginal propensity to consume out of wealth is usually not equal to 1. Secondly, most of the other decisions that are endogenous to the model are between discrete alternatives (e.g. employment, the timing of pension take-up, participation in personal pensions). The first of these factors helps to ensure that the search routine described above will converge. The second helps to limit the domain of viable alternatives over which the search must be conducted. Matters become slightly more complicated when either of these two factors break down.

The first factor identified above breaks down when benefit units are liquidity constrained, in which case the marginal propensity to consume out of liquid net wealth is 1. Liquidity constraints are most likely to be encountered at the extremes of the life course. For the young, this poses little problem for the simulations, whereas the opposite holds true for the old. This is because the intertemporal connection between utility maximising decisions is severed whenever behaviour is constrained. Hence, the behaviour generated back in time for older individuals by the model will tend to provide a more approximate guide to their actual circumstances than for individuals earlier in life.

The second factor identified above breaks down when account is taken of decisions over multiple continuous domains (e.g. pension contribution rates, and/or investment in liquid assets). The resulting complexity can increase computation times considerably.

6.4 Long-term incentive compatibility

As an individual ages, their assets portfolio will generally evolve in response to their accumulated life-history, responding to positive and negative shocks to a wide range of factors including labour market success, investment returns, health, relationship status, and so on. An individual who is in the top wealth decile at a given age is consequently likely to have experienced more favourable forms of variation during their lifetime than an otherwise similar individual in the bottom decile. Failure to accommodate this feature can result in unrealistic projections for wealth backward through time. For example, in cases where an individual is associated with insufficiently favourable variation in backward projections to reflect the assets they hold in the simulation reference period, the backward projections will indicate unrealistically high wealth holdings early in the adult lifetime. The problem is to model random variation during the life course that is broadly compatible with characteristics described in the reference period from which backward projections are made.

The model uses 'inheritances' as a tool in the backward projections for meeting the 'long-term incentive compatibility' problem that is described above. This approach mitigates computational burden by bridging the gap between an individual's assumed history of random variation and their financial circumstances observed in the simulation reference period using a single unrequited transfer.

Receipt of inheritances

Receipt of inheritances is only modelled in backward projections by the model. Each benefit unit represented in the reference population cross-section is assumed to receive at most one inheritance during the course of a reference person's lifetime, and all inheritances are assumed to be unanticipated. The model is designed to search for bequests which ensure that the liquid net wealth of each benefit unit falls below a threshold value at the beginning of the simulated lifetime. The search starts from the assumption that all benefit units in the reference cross-section did not receive a bequest at some time prior to the observed cross-section. Each benefit unit is then projected backwards, and any benefit unit that is simulated to hold liquid net wealth at the beginning of the simulated lifetime that exceeds the assumed threshold is identified as having received an inheritance at some time prior to the reference cross-section. The age of inheritance receipt is then identified for each of these benefit units, based upon age specific probabilities of incidence, and subject to the upper bound described by their respective ages in the reference cross-section. All benefit units that are assumed to receive an inheritance prior to the reference cross-section are then re-projected back through time, taking into account the age of inheritance receipt and an assumed (minimum) inheritance value. Any benefit unit whose liquid net wealth at the beginning of the simulated lifetime continues to exceed the upper threshold following this iteration of the backward simulation then passes to the next iteration of the search routine, in which the inheritance value is increased by a fixed increment, and projections back through time are re-simulated. The process is repeated until the upper threshold condition on liquid net wealth is satisfied for all simulated benefit units.[18]

Leaving of inheritances

Implicit in the specification of preferences described by equation (1) is the assumption that inheritances are left at the time of death of each adult (possibly to their spouse, if one exists). By definition, no reference adult described by the cross-sectional data from which model projections are made can have died prior to the year for which the cross-sectional data were observed. Hence, inheritances can

[18] This process does not guarantee that all benefit units will satisfy the upper bound on liquid net worth at the beginning of the simulated life-course. Nevertheless, practical use of the approach suggests that it achieves the desired objective in the vast majority of cases.

only be left in the model in periods that are projected forward through time. In this case, simulating inheritances imposes no additional computational burden beyond that required to simulate liquid net wealth forward through time.

7 Individual Savings Accounts

Individual Savings Accounts (ISAs) are an asset class that is designed to encourage savings for retirement in the UK. There are three principal elements to ISAs. First, investment income and capital gains within an ISA are tax free, both at the time earned and upon withdrawal. Secondly, annual contributions are subject to upper limits. And thirdly, ISAs impose no limits on when accumulated funds can be withdrawn. The first of these elements encourages contributions into the scheme, the second discourages withdrawals, while the third relaxes the liquidity disincentives associated with traditional pension schemes.

Each benefit unit is assumed to be able to contribute to a single ISA account. Annual contributions to the ISA account are made out of post-tax income, and are subject to a per-period cap that doubles where the benefit unit is comprised of an adult couple. Although a distinction currently exists in the UK between so-called 'cash' and 'stocks-and-shares' ISAs, the model is adapted to consider only one of these types at a time. At the start of each period, all wealth held in an ISA is assumed to accrue the same rate of return, r_t^{ISA}, which can be specified as uncertain. Uncertain returns to ISAs are assumed to be perfectly correlated with the returns to the risky liquid asset (r_t^r in Section 6.2). In most periods, wealth held in an ISA, w^{ISA}, is assumed to vary intertemporally as described by the equation:

$$\begin{aligned} w_{i,a}^{ISA} &= r_{t-1}^{ISA} w_{i,a-1}^{ISA} + k_{i,a}^{ISA} \\ \ln\left(r_t^{ISA}\right) &\sim N\left(\mu_{ISA} - \frac{\sigma_{ISA}^2}{2}, \sigma_{ISA}^2\right), corr(r_t^{ISA}, r_t^r) = 1 \end{aligned} \tag{10}$$

where $k_{i,a}^{ISA}$ denotes net contributions into the scheme (negative when there are net out-flows), and $corr\,(.)$ denotes the correlation coefficient. The only departure from equation (10) is when the relationship status of a reference adult is identified as changing, in which case associated fluctuations in ISAs are modelled in the same fashion as described for liquid net wealth (see Section 6).

As noted in Section 3, the preference relation assumed for analysis also allows for the possibility that contributions to ISAs are influenced by decision costs, $\Delta_{i,a}^{ISA}$. In this case, utility is assumed to decline discontinuously when the first contribution to a benefit unit's ISA is made.

8 Contributory State Pensions

The model is designed to permit up to two contributory state pensions to be run in parallel. The terms of these pensions are based upon the basic State Pension and the State Second Pension, as these were applied in the UK in 2011.

8.1 The basic State Pension (BSP)

The BSP is a flat-rate contributory state pension, rights to which are accrued through accreditation in respect of National Insurance contributions during the working lifetime. The model tracks the number of years, $yr_{i,a}^{BSP}$, for which each benefit unit, i, at age a, has been accredited with National Insurance contributions, up to the maximum defined by the number of years required for a full BSP for each adult benefit unit member. Accreditation for National Insurance contributions is derived if the earnings of an adult exceed a minimum threshold (the primary threshold), and can also be allowed for in respect of child care (non-employment during peak child-rearing ages), or involuntary unemployment (periods in which a low-wage offer is received – see Section 4). In most years prior to state pension age, the number of years of accreditation for National Insurance contributions is defined by:

$$yr_{i,a}^{BSP} = yr_{i,a-1}^{BSP} + k_{i,a-1}^{BSP} \tag{11}$$

where $k_{i,a-1}^{BSP}$ are the additional contributions accredited to benefit unit i at age $a-1$. The only exception to equation (11) is when the relationship status of a reference adult is identified as changing, in which case associated fluctuations in BSP rights are modelled in a similar fashion as described for liquid net wealth (see Section 6).

Each benefit unit is assumed to draw its basic State Pension from state pension age, a^{SPA}, which is permitted to vary between birth cohorts, and this public transfer is added to pension income for tax purposes. The value of the BSP payable to each benefit unit depends upon the contributions history of the benefit unit, the value of the full BSP assumed for the reference year, a growth rate applied until the time when the reference adult of the benefit unit attains state pension age, and another growth rate applied from state pension age. Two values of the full BSP are taken into consideration; one for single adults, and another for adult couples. The model assumes that each benefit unit is paid the greater of the single allowance, paid in respect of the number of complete contribution histories accrued by all adult benefit unit members, and the couple allowance, paid in respect of a single adult's contribution history for couples. The model does not track each adult's contribution history separately, but instead assumes that all contribution years accrue to the reference adult up to the number of years required for a full BSP, and to the spouse (if one exists) thereafter.

8.2 The State Second Pension (S2P)

The S2P is an earnings-related contributory state pension. Like the BSP, rights to the S2P are modelled at the benefit unit level, are accumulated prior to state pension age, and are associated with a (taxable) income stream from state pension age, a^{SPA}. Unlike the BSP, the model tracks rights to the S2P in the form of the associated annuity income. The annuity to which a benefit unit is eligible from state pension age is assumed to grow at the rate r_w^{S2P} until state pension age, and at the rate r_r^{S2P} from state pension age. During accumulation, rights to the S2P are calculated with respect to three earnings thresholds. Any benefit unit with earnings in excess of the Lower Earnings Threshold, LET_t, is assumed to gain a flat-rate increase in their S2P rights. This flat-rate contribution can also be allowed to accrue in respect of child care and (involuntary) unemployment, as described above for the BSP. Benefit unit earnings between the Lower Earnings Limit, $LEL_t > LET_t$, and the Upper Accrual Point, $UAP_t > LEL_t$, are assumed to increase S2P rights by a fixed accrual rate. Benefit unit earnings in excess of the UAP_t have no bearing upon S2P rights.

In most periods, rights to the S2P follow:

$$py_{i,a}^{S2P} = r_{r/w}^{S2P} py_{i,a-1}^{S2P} + k_{i,a-1}^{S2P} \tag{12}$$

where $k_{i,a-1}^{S2P}$ denotes the additional rights to the S2P derived by the benefit unit's earnings at age $a-1$. The only exception to this equation is when the relationship status of a reference adult is identified as changing, in which case associated fluctuations in S2P rights are modelled in a similar fashion as described for liquid net wealth (see Section 6).

9 Private Pensions

A flexible structure that is capable of reflecting the broad features of private pensions in the UK is included in the model. The UK private pension system is comprised of Occupational Pensions (OP) run by companies on behalf of their employees, and Personal Pensions (PP) that individuals provide for themselves. Furthermore, a conspicuous feature of private pensions in the UK is the diversity of schemes that exist. This diversity is accommodated in the model, by allowing OPs and up to five PP schemes to be run in parallel. These schemes can differ from one another concerning the terms of pension contributions, and returns to pension wealth.

All private pensions are modelled at the benefit unit level, and are Defined Contribution in the sense that every benefit unit is assigned an account into which their respective pension contributions are (notionally) deposited. Where OP and PPs are run in parallel, then any benefit unit with a labour income in excess of a lower bound is assumed to participate in the OP, while any benefit unit in which the highest adult earner has a labour income within an income band can be given the option to participate in a PP. The income thresholds used to manage eligibility to private pensions can (but do not have to) overlap. Where multiple PPs are accommodated in the model, then each benefit unit is considered to be eligible for a single PP in each simulated period, where the evolution of pension eligibility is determined by a stochastic process.

The stochastic process used to identify the PP for which each benefit unit is eligible, PP, proceeds as follows. Suppose, for example, that the model allows for three alternative PP schemes, $PP \in \{A, B, C\}$. Further, suppose that a benefit unit chooses to work and to contribute to their eligible pension $PP = B$ at age 35. For any benefit unit that chooses to contribute to their eligible pension, the model first assumes that there is a probability that they will remain eligible for the same pension in the immediately succeeding period. This step is skipped for benefit units that choose not to contribute to their eligible pension. Assume here that this probability is set to 85%, where the remaining 15% can be interpreted as the probability of job change from one period to the next. The model then identifies into which of three income bands the wage potential of the benefit unit falls. We refer to wage potentials here because the model considers for this purpose a measure of earnings potential, and not actual earned income (see Section 4). Each income band has a separate set of probabilities describing the likelihood of eligibility for the alternative pension schemes. Suppose that our benefit unit falls into the middle income band, where the probabilities are: $\Pr(A) = 0.1$, $\Pr(B) = 0.6$, and $\Pr(C) = 0.3$. Given this information the model calculates the probability that the benefit unit will be eligible for each of the alternative pension schemes in the following period; in the above example, the probability is skewed toward B, with a probability of 0.94. These probabilities are used to evaluate agent expectations when solving the lifetime decision problem, and are used to project the circumstances of each benefit unit through time using Monte Carlo methods.

9.1 Private pension contributions

Contributions to private pensions are defined as rates of employment income (implying that they are limited to benefit units that work), and are distinguished by whether they are made by the employer or the employee. Employer contributions are assumed to be exempt from taxation, and labour income is reported net of these contributions. Employee contributions can be subject to taxation, and labour income is reported gross of these contributions.

Contributions to an OP are simulated exogenously in the model. The model assumes that any benefit unit that has not previously accessed its pension wealth, and which has labour income in the prevailing period in excess of a lower threshold, will contribute to the OP. The employee contribution rate to the OP, $\pi_{ee}^{OP,k}$, and employer contributions rate, $\pi_{er}^{OP,k}$, are defined as fixed proportions of benefit unit labour income, which may vary by employment sector k (self-employed / private sector / public sector employees; see Section 5). Hence, for any benefit unit that contributes to an OP in a given period, the value of the contribution is defined as: $pc_{i,a}^{OP} = \left(\pi_{ee}^{OP,k} + \pi_{er}^{OP,k}\right) g_{i,a}$.

Each PP scheme, h, has a single employer contribution rate, $\pi_{er}^{PP,h}$, and a (minimum) employee contribution rate, $\pi_{ee,0}^{PP,h}$, which apply to labour income within an income band defined by upper and lower bounds. Contributions to a PP can be simulated either endogenously or exogenously. Where multiple PPs are considered, then endogenous decisions are limited to the extensive margin. Where only one PP is considered for analysis, then the benefit unit can choose whether to make fresh contributions to their eligible scheme (the extensive margin) and / or how much to contribute (the intensive margin) in excess of $\pi^{PP,h}_{ee,i,a} \geq \pi^{PP,h}_{ee,0}$. Hence, for any benefit unit that contributes to an PP in a given period, the value of the contribution is defined as: $pc^{PP}_{i,a} = \left(\pi^{PP,h}_{ee,i,a} + \pi^{PP,h}_{er}\right) \max\left[\min\left(g_{i,a}, g^{\max,PP}_t\right), g^{\min,PP}_t\right]$, where $g^{\max/\min,PP}$ describe the upper and lower bounds on income assessable for contributions.

If endogenous PP contributions are considered, then it is also possible to allow for the influence of assumed 'default options', which may vary between alternative pension schemes. This is achieved by allowing for 'decision costs' that discontinuously reduce welfare when a benefit unit decision departs from the pre-assigned default. Defaults may be considered over both contribution rates and / or participation alternatives. Defaults concerning participation, δ^{PP}, are defined when a benefit unit first becomes eligible to a given scheme (i.e. auto-enrolment or active opt-in), and thereafter track a benefit unit's pension decision in the preceding period. As indicated by equation (1), decision costs (represented by Δ) have been included in the assumed preference relation in an additively separable form, so that they influence the level of welfare, but not first order conditions.

Pension contribution caps

Aggregate private pension contributions (from both OPs and PPs) can be subject to three separate caps. Two of these caps are year specific, can be altered over three age intervals, and can be administered at either the individual or benefit unit level. The first of these period-specific caps defines the upper limit on pension contributions that are eligible for tax relief. Employer contributions are accommodated first within this cap, after which private contributions are included. Private pension contributions that can be accommodated within this first cap are referred to in the model as *concessional contributions*, pc^c. Any contribution in excess of the first cap can also be subject to a second cap defining the maximum contributions that are permissable each simulated period. Any employer contributions not accommodated within the concessional contributions cap are preferentially accommodated within the second cap. If employer contributions are in excess of the sum of these two period-specific caps, then the excess of employer contributions is considered to be returned to the employer and no private contributions are permitted. Any private contribution accommodated within the second cap is recognised as a *non-concessional contribution*, pc^{nc}, and is assessable for tax purposes. The third cap on pensions is an upper bound on the maximum size of the private pension pot; any contribution made that would result in this third cap being breached is assumed to be taxed at the rate of 100%. The first two of the caps defined above can be allowed to vary through time, at year specific rates within a closed and bounded period, and at fixed rates beyond this period. The upper bound on the size of the pension pot is assumed to remain fixed through time.

9.2 Evolution of private pension rights during accrual

Where multiple PP schemes are considered for analysis, balances of PP wealth are assumed to be perfectly portable between schemes, and each benefit unit is assumed to hold all of their personal pension wealth in the scheme for which they are eligible at the given point in time (which evolves stochastically as described above). The same basic rules are used to project through time wealth held in occupational and personal pensions, $p = \{OP, PP\}$. Until the year in which a benefit unit accesses its pension wealth, returns to private pension savings attract a rate of return that can be defined as uncertain. When returns are assumed to be uncertain, then they are perfectly correlated with the returns to the risky liquid asset (r_t^r in Section 6.2). Hence, accrued pension rights do not hedge against uncertainty in the liquid asset portfolio. Intertemporal accrual of private pension wealth, w^p, is described by equation (13):

$$\begin{aligned} w_{i,a}^p &= \max\left\{0, \min\left[w^{p,\max}, r_{t-1}^p w_{i,a-1}^p + pc_{i,a}^p\right]\right\} \\ \ln\left(r_t^p\right) &\sim N\left(\mu_p - \frac{\sigma_p^2}{2}, \sigma_p^2\right), corr(r_t^p, r_t^r) = 1 \end{aligned} \tag{13}$$

where p distinguishes OP, and each of the PP schemes considered for analysis, $w^{p,\max}$ defines the maximum size of a (PP or OP) pension pot, and $corr\,(.)$ is the correlation coefficient. The returns accruing to alternative PP schemes can be allowed to vary to reflect differences in assumed management charges.

Note that equation (13) indicates that any pension wealth accrued in excess of the pension pot cap is assumed to be lost, possibly through an excess contributions tax. The equation also indicates that the only difference between concessional and non-concessional contributions is their tax treatment at the time the contributions are made – they are treated equivalently after entering the pension notional account. Equation (13) holds in all periods prior to pension receipt except following relationship transitions, in which case associated fluctuations in pension rights are modelled in a similar fashion as described for liquid net wealth.

9.3 Accessing pension wealth and retirement

The age at which pension dispersals are accessed, a^P, can be determined either endogenously or exogenously in the model. If both OP and PP schemes are modelled, then rights to both pensions are assumed to be accessed at the same time. At the time that pension wealth is accessed, a fixed fraction of accrued pension wealth (that may differ between occupational and personal pensions) is received as a lump-sum cash payment, and the remainder converted into an annuity. The model can be used to simulate either fixed term or life annuities. Annuity rates for life annuities are calculated to reflect birth cohort-specific survival probabilities in the model, subject to assumed rates of investment returns, real growth, and transaction costs levied at time of purchase. The tax treatment of both the lump-sum and pension annuity can also be specified.

When the timing of pension dispersals is exogenously imposed, then all benefit units are assumed to access their pension wealth at their respective state pension ages, a^{SPA} (see Section 8). When the timing of pension dispersals is endogenously determined, then this decision can be made subject to minimum thresholds on age and annuity income. Furthermore, limits can be imposed on a benefit unit's pension contributions and employment opportunities following pension take-up. Employment opportunities can be subject to both hard limits (on the ability to find employment following pension take-up), and soft limits (in the form of wage penalties imposed on pensioner benefit units, discussed in Section 4).

9.4 Simulating pensions through time

Up to five state variables, three of which may be continuous, are required to reflect private pensions as they are described above: the value of occupational pension wealth (w^{OP}, continuous); the value of personal pension wealth (w^{PP}, continuous); the type of personal pension to which an individual is eligible (PP); the default decision regarding private pension contributions (δ^{PP}, may be continuous); and whether private pension wealth has previously been accessed $\left(a^{P}\right)$. If, however, occupational and personal pensions are assumed to attract the same rates of return, and are subject to the same conditions when accessed (ie same proportions taken as a lump-sum, same annuities purchased), then there is no need to distinguish occupational from personal pension rights and the model is designed to take advantage of this possibility. Nevertheless, accommodating all of the potential variation that is outlined in the above description is likely to be prohibitively demanding on existing computing technology, and the model allows selected features to be suppressed where they are not of immediate concern.

Simulating pensions backward through time shares many similarities with the simulation of liquid net wealth described in Section 6. The principal innovations associated with the simulation of pensions concern the treatment of the timing of pension access when this is endogenous. Consider, for example, the case where the model is projecting backward through time the circumstances of a benefit unit that is identified as having started to draw down their pension wealth in some period prior to the reference cross-section. The model must then identify when the benefit unit first started to access its pension wealth. Ensuring incentive compatibility of backward simulations is complicated in this context by the existence of multiple ages at which pension access will be optimal, given the assumption that pension income was not taken up at some preceding age, so that it is necessary to test over a large number of alternatives.

LINDA addresses this problem by projecting back in time two separate life histories for each benefit unit that is identified as having accessed their pension wealth in the reference population cross-section. Denote these two life histories A and B. Life history A assumes that each benefit unit first accesses their pension wealth at the earliest possible opportunity (e.g. currently at age 55 in the UK). Life history B, in contrast, assumes that each benefit unit accesses its pension wealth at the latest possible opportunity that has not been shown to be incentive incompatible. Where an inconsistency is shown to exist between the assumptions of life history B and decisions of the dynamic program, then the results stored in life history B are discarded, and replaced by those stored in A.

Consider, for example, a benefit unit that is identified as having previously accessed its pension wealth, with a reference adult who is aged 72 in the reference period. Life history A proceeds on the assumption that pension wealth was first accessed by this benefit unit at age 55. Life history B would proceed on the assumption that pension wealth was first accessed at age 71. No inconsistency is possible until the model comes to solving for benefit unit characteristics at age 70. In this case, life history B assumes that the benefit unit will not want to access its pension wealth at age 70, because it is identified as first accessing this wealth at age 71. If, however, decisions consistent with the solution to the dynamic program suggest that the benefit unit, having not previously accessed its pension wealth at age 70, would choose to access its pension wealth at age 70, then results stored for age 71 in life history B are discarded, and replaced by those stored in A, with the additional assumption for life history B that pension wealth is first accessed at age 70. This procedure is repeated until the earliest age at which pension wealth can be accessed, at which time life history A is discarded.

10 Education

Each adult is allocated an education state at entry into the model, $ed_{i,a}$, referring to the highest qualification held. It is possible to distinguish between up to five education states, one of which is reserved to reflect tertiary education, and the other four to reflect alternative levels of pre-tertiary education. The pre-tertiary education states can differ from one another in relation to the assumed probabilities of receiving a low wage offer and assumed wage premia; λ^o and λ^{ng} described in Section 4. In addition, individuals with tertiary education can be distinguished from non-tertiary educated in relation to the age specific evolution of latent wages (h in Section 4), and transition probabilities governing marriage and divorce.

Individuals who do not enter the simulated population with tertiary education may be identified as tertiary students, $stud_{i,a}$. Any individual who first appears as a tertiary student is assumed to leave tertiary education at an exogenously defined age (assuming that they survive), at which time they may transition to tertiary educated, depending on a stochastic process that represents whether they pass their final exams. At the time an individual leaves tertiary education, they receive a new random draw for their wage potential from a log-normal distribution, where the terms of the distribution differ for graduates and non-graduates. This approach for students is inverted in the projections backward through time for any individual who is identified as a graduate when they first appear in the model population. All processes that govern transitions between alternative education states when projecting a population through time are assumed to be fully consistent with the associated expectations adopted to solve the lifetime decision problem.

11 Relationship Status and Dependent Children

11.1 Relationship status

A 'relationship' is loosely defined in the model as a cohabitating partnership, and may be parameterised to reflect alternative arrangements, including formal marriages and civil partnerships. In each period of a reference adult's life, their relationship status in the immediately succeeding period is uncertain, reflecting the uncertain likelihood of marriages, divorces, and widowhood. The transition probabilities that govern relationship transitions depend upon a reference adult's existing relationship status, age and birth year, the age and birth year of their spouse (if they have one), and may also vary with respect to their education status, health status, or the presence of dependent children. An important feature of the approach taken to simulated relationship status is that it is exogenous of the decision process, and can therefore be simulated before endogenous model behaviour (including consumption, employment, and savings decisions) is projected. Probabilities of marriage and divorce are stored in a series of 'transition matrices', each cell of which refers to a discrete relationship/age/birth year combination; separate matrices are also stored that distinguish reference adults by education status, health status, and whether there are dependent children in the benefit unit. Similar transition matrices are used to model mortality (and widowhood), as described in Section 12.

When solving the lifetime decision problem, individuals are assumed to anticipate the probabilities of relationship formation, divorce, and death that govern relationship transitions. The decision problem is simplified by assuming that each individual expects to marry – if they marry – an identical clone of themselves. This assumption omits the uncertainty that would otherwise need to be accommodated concerning the characteristics of each potential spouse.

There are two principal approaches to projecting marital status in dynamic microsimulation models.

The first assumes that the partners of newly married reference adults are drawn from outside the simulated population, and is used by LINDA for projections backward through time. In this case, the characteristics of spouses are simulated on the assumption that relationships form between identical clones. This assumption is consistent with the assumptions used to project agent expectations for evaluating the lifetime utility maximisation problem, and has the added advantage that it requires little additional data to be generated in respect of the spouse. The second approach adopts a 'closed model' specification that identifies married partners from within the simulated sample, and is used by LINDA to simulate the population forward through time. This difference in simulation approach between the forward and backward projections is motivated by the dual observations that a 'closed model' for relationships is both facilitated, and facilitates a model context that captures the evolving population cross-section.

The cross-sectional data that are loaded into the model to initialise the simulated population include the marital status of each represented adult. These data also include a personal reference number, and a benefit unit reference number. The same benefit unit reference number is recorded for each member of a cohabitating relationship. Matching of spouses in forward simulations is performed on a year-by-year basis. At the start of each simulated year, the pool of marrying adults is identified. Spouses are sorted using a 'points system', with reference to each individual's age and education status. First, the model lists the population of newly married adults in no particular order, from 1 to n. Starting with the first listed individual, the model identifies their age and education status. The model then considers the suitability of individual 2 in the list as the spouse for individual 1. It does this by assigning one "point" for each year difference in age and adds five points if the two individuals hold different education levels. The model stores the result, and then passes to individual 3 in the list. The model evaluates the points associated with a union between individuals 1 and 3. If the aggregate is less than the points associated with a union between individuals 1 and 2, individual 3 is stored as the preferred spouse for individual 1. The model then passes to individual 4 in the list, and continues its search until it either identifies a perfect match with individual 1 (zero points in aggregate), or reaches the end of the list (individual n). If it reaches the end of the list, the currently stored best match for individual 1 is adopted as individual 1's spouse, and the two adults are assigned the same benefit unit number, equal to the person identifier for individual 1.

The model proceeds by checking whether individual 2 has been matched to individual 1. If not, then it searches for a match to individual 2 in the same fashion as described for individual 1, with the addition that it checks to ensure that each tested individual has not been previously assigned a spouse. The model works through the entire pool of newly married individuals in this way. If an odd number of newly married individuals is identified by the model, then the last individual considered in the list is assumed to remain single.

The model simulates endogenous behaviour with reference to a range of individual specific characteristics, including age, year of birth, and education. Where marriages are between identical clones, as is assumed in backward projections, then this feature is trivial to apply. Potential ambiguity, however, arises wherever spouses have different characteristics. This ambiguity is resolved in forward projections by simulating benefit unit decisions with reference to characteristics held by the benefit unit's 'reference adult'.

All simulated benefit units are assigned a reference adult in every simulated year. In backward projections, each benefit unit includes one adult in the simulated sample, who is defined as the reference adult. Similarly, un-married benefit units in the simulated sample are comprised of a single adult who is defined as the respective reference adult. It is assumed that the reference adult of each benefit unit can only vary in a period when marital status changes. These conditions fully identify reference adults in the model, except in the reference population cross-section, or when adults are projected to get married in forward projections. In both of these cases, the model begins by checking whether one spouse has accessed their pension wealth, but the other has not (see Section 9.3). If this is the case, then the reference adult is set to the spouse that has accessed their pension wealth. Otherwise, the reference adult is defined as the spouse with the highest wage potential (see Section 4).

At the time of a union in forward projections, jointly held assets are the sum of the assets held individually by each spouse. Widowhood in forward projections of the model is based on the age of death simulated for each adult (described in Section 12.4). In backward projections, widowhood is identified randomly, based upon the mortality rates of the simulated adult (given the assumption of marriages between clones). In the event of widowhood, all assets and children of the benefit unit are assumed to reside with the surviving spouse. Divorce is simulated based on the transition probabilities applicable for the benefit unit reference person, in which case all assets and children are divided evenly between the respective spouses (to the nearest integer in the case of children).

11.2 Birth and aging of children

The model is designed to take explicit account of the number and age of dependent children of reference adults. The birth of dependent children is assumed to be uncertain in the model, and is described by transition probabilities that vary by the age, birth year, relationship status, and previously born children of a reference adult. These transition probabilities are stored in a series of transition matrices, in common with the approach used to model relationship status (described above). Having been born into a benefit unit, children are assumed to remain dependants until an exogenously defined age of maturity. A child may, however, depart the modelled benefit unit prior to attaining maturity, if the reference adult experiences a relationship dissolution (to account for the influence of divorce).

Allowing for dependent children in the way set out in the preceding paragraph can lead to a very significant increase in the computational burden of the lifetime decision problem. If, for example, a benefit unit was considered to be able to have children at any age between 20 and 45, with no more than one birth in any year, and no more than six dependent children at any one time, then this would add an additional 334,622 states to the decision problem (with a proportional increase in the associated computation time).[19] In cases where children are not an issue of concern, the model consequently allows associated uncertainty and heterogeneity to be suppressed. In this case, the number of dependent children in each benefit unit is described as a deterministic function of the age and relationship status of the reference adult. Where the number and age of dependent children are considered to be important, then the model is made computationally feasible by limiting child birth to a fixed number of reference person ages. The model may be directed, for example, to consider child birth only when the reference adult is aged 22, 26, or 33 years.

Capturing realistic benefit unit sizes in context of limited child birth ages will usually require that multiple births be allowed at each birth age. We might, for example, allow up to two children to be born at each of the three child birth ages referred to above, in which case the maximum number of children in a benefit unit at any one time would be limited to 6. In this example, the computation burden of the decision problem would be increased by a factor of 231, which is sufficiently constrained to make the solution to the decision problem feasible on contemporary computing technology.

Restricting the number of ages at which a child can be born in the model raises a thorny problem regarding identification of the transition probabilities that are used to describe fertility risks. The model calculates the required probabilities internally, based upon the assumed birth ages and fertility rates reported at a highly dis-aggregated level. This approach has been adopted both because statistical agencies tend to publish data at the dis-aggregated (annual age band) level, and because it facilitates associated sensitivity analyses to be conducted around the number and precise birth ages assumed.

Consider the above example, in which child birth is limited to three discrete birth ages – 22, 26 and 33 – and where the maximum number of children that can be born at each birth age is limited to 2. Suppose also that age specific fertility rates calculated on survey data extend between ages 16 and 45 inclusive. The fertility transition probabilities associated with each birth age are calculated by dividing life into mutually exclusive age bands, where the thresholds between adjacent bands are set to the mid-points between the respective 'birth ages'. A diagrammatic representation of this division for the example considered here is provided in Appendix D. Monte Carlo methods are used to generate data for the complete life history of a large number of reference adults for each potential birth year

[19] This assumes an age of maturity of 17.

considered for analysis.[20] This simulated population is calculated in a way that is consistent with the wider analytical framework, but with the exception that fertility transitions are permitted at each individual age (subject to the limitation that no more than one child be born at each age). Fertility transition probabilities for each birth age are then calculated by aggregating up the data calculated for the simulated population within the respective age band.

11.3 Simulating relationships and dependent children backward through time

As noted above, the model assumes for backward projections that the spouses of reference adults enter and exit the simulated frame with their marriage. There is, however, the additional problem of ensuring that the projected marital status of a reference adult in the reference cross-section coincides with their observed relationship status. The same problem applies for the projected number and age of dependent children. In the case of children, however, the problem is further complicated by the fact that associated shocks – the birth of a new child or the reduction in the number of children due to marital dissolution – can influence benefit unit circumstances over a large number of simulated periods. The problems here are conceptually similar to those described for the backward simulation of the timing of access to pension wealth, and could be addressed in a similar fashion to the approach described for pension access in Section 9.4. In contrast to the timing of access to pension wealth, however, the assumed model structure permits benefit unit demographics to be simulated separately from all other model characteristics. This feature of the model motivated the adoption of a 'trial-and-error' approach to simulate benefit unit demographics backward through time.

A set of random draws from a uniform distribution are taken for each reference adult, one draw for each possible period of the simulated lifetime. These random draws are used to project relationship status and dependent children for the life course of each reference adult up until the age that they appear in the reference cross-section, assuming that all individuals are single without children at the youngest age represented by the transition probabilities (16 in most applied contexts). The simulated relationship status, number and age of dependent children projected for each reference adult are compared against the data loaded in for the adult in the reference cross-section. A new set of random draws is taken for each adult until a set of draws consistent with their circumstances in the reference cross-section is found. To ensure that this random search routine does not continue indefinitely, the model limits the number of sets of random draws that it tests for each simulated adult, reporting the proportion of the simulated population for whom a precise match was not found.

[20] Sensitivity analysis suggested that a simulated population size of 1,000,000 benefit units obtains reliable fertility transition probabilities in an applied context considered for the UK.

11.4 Child maturity

When dependent children reach an assumed age of maturity they depart their parental benefit unit and enter one of their own. In the backward projections, dependent children of reference adults enter the simulation frame when they are aged 17, and depart again in the year prior to their birth, in a similar fashion to the treatment of spouses of reference adults over marital transitions. In the forward projections, dependent children can be followed into adulthood. This feature is implemented to reflect the evolving population cross-section, and is governed by two key model boundaries. The first is the maximum population size that the model is directed to take into account, and the second is the number of periods into the future that projections are made. The more restrictive of these two boundaries determines the time horizon over which the evolving population cross-section is projected, and the model reports this as part of its standard on-screen output. The time horizon of the projected population cross-section in turn determines whether a dependent child will mature and be simulated as an adult by LINDA; any child who matures beyond this cross-sectional time-horizon is dropped from the simulated sample.

When a child first achieves their maturity, a series of characteristics must be identified to continue their projection into adulthood. Each maturing child is assigned a unique person identifier. Their age and birth year are carried over from their parental benefit unit, and their year of entry into the simulated sample is recorded. Simulation of survival, relationship status, and dependent children proceeds in the same way as described for adults in the reference cross-section, with the exception that no search routine is required to match the characteristics at the time that they enter the model sample frame as an adult. Health status (discussed in Section 12) is randomly assigned, based on age and year specific transition rates. The health states to which maturing children are allocated can be limited in the case where a child is identified as disabled.

All maturing children are identified as non-graduates. Education status of maturing children is otherwise allocated randomly, based on transition probabilities that can vary by child disability status, year and the education status of the reference adult in the parental benefit unit. All maturing children are assumed to not be self-employed, or public sector employees, if these employment sectors are included for analysis. Low wage offers are simulated in the same way as for the wider population, based on age, year, and education specific probabilities. All assets are set to zero for maturing children. This assumption is made because the model does not account for child income, and unrequited transfers between benefit units other than inheritances (see Section 6.4) are not included in the model. Wage potential at age of maturity is based on a random draw from a log-normal distribution, the means and variances of which are age, year, and education specific.

12 Health and Mortality

Each individual is allocated a health status at the time that they enter the simulated population, and a set of age-specific random draws from a uniform $[0, 1]$ distribution, which refer to below as the 'health vector'. The model can distinguish the health status of dependent children, the health status of each simulated adult, associated carer responsibilities, and each simulated adult's time of death. As projections for health push computing technology to the limits of what is currently feasible, the model has been designed only to project health and carer states forward through time.

12.1 Health status of children

The health status of children, $health^{c}_{i,a}$, distinguishes between two discrete alternatives that are designed to identify those with and without a persistent disability. In the case of children entering the population in the data for the reference cross-section, the cross-sectional data include a disability identifier. In the case of children entering due to simulated child birth, child health status is allocated by comparing the first element of the child's health vector with an exogenous incidence rate on the assumption that up to one disabled child may enter each benefit unit at each 'birth age'. The health status of children is assumed to remain unchanged until they mature into adulthood. Furthermore, child disability status is assumed to influence health and education status upon maturity (Section 11.4). It is consequently possible to use the model to track the influence of disability throughout the life course, from birth through to death.

Child disability influences simulated welfare benefits (discussed in Section 6), the carer responsibilities of parents (discussed in Section 12.3), and benefit unit costs, ndc^{h}. Benefit unit costs associated with child disability can either be exogenously defined, or be set equal to the value of associated welfare benefits.

12.2 Health status of adults

The model is designed to distinguish between up to 10 discrete health conditions for each adult in each period projected forward through time. Behavioural solutions are structured around a health state described at the benefit unit level, $health^{j}_{i,a}$. In the case of single adults, this health state defines the health condition of the relevant adult. In the case of couples, the health state defines the 'health combination' of the two spouses. The health conditions recognised by the model are drawn from a discrete set of alternatives, the number of which is a model parameter that can be adjusted as desired. Simulated health conditions evolve through time, based on exogenously defined transition probabilities that vary by each adult's prevailing health condition, education, age, and year.

The health state can influence benefit units in a variety of ways. As noted above, the health condition of each adult can affect their likely health condition in the future. This feature is required to capture the persistence that is associated with many health conditions, which may have an important bearing on the life-course. The health condition of one adult in a couple can affect the carer responsibilities of their spouse (discussed in Section 12.3). The health state can also be defined to limit the discrete set of labour alternatives available to each adult, the probabilities of receiving a low wage offer and wages earned (Section 4), welfare benefits (Section 6), the likely evolution of relationship status in prospective years (Section 11.1), and non-discretionary costs, ndc^h. As for children, benefit unit costs associated with adult health conditions can either be exogenously defined, or be set equal to the value of associated welfare benefits.

Care must be exercised when defining this aspect of the model to ensure feasible computational times in context of prevailing computing technology. Consider, for example, specifying the model to distinguish between 4 discrete health conditions in each year throughout each adult's simulated lifetime. In this case, any single adult would be associated with four potential health conditions, and any couple would be associated with 16 alternative health combinations (4 for each spouse). This would expand the size of the state space by a factor of 10 (= (4 + 16)/2). If transitions between simulated years are permitted between any given health condition and any of the alternative conditions, then this would expand the computations involved when evaluating expectations by a factor of 100 (= (4 + 16)/2 x (4 + 16)/2). Increases in computational burden of this magnitude can be highly undesirable.

The model is designed to limit the computational problem outlined in the preceding paragraph in two ways. First, some health transitions occur with low probabilities. The model is designed to ignore evaluation of expectations that are associated with (near) zero probabilities. Secondly, it may be practical to ignore some health combinations for adult couples. This would be the case, for example, if the incidence of two severely disabled adults in a couple was very low over a specific age band. Ignoring very unusual combinations of this sort can help to improve model performance. The model is designed to take advantage of these types of economies by allowing the set of health conditions considered for couples to be restricted as desired, with restrictions permitted to vary over two alternative age bands.

12.3 Carers

A "carer state" is generated for each benefit unit in each simulated period, $carer_{i,a}$, where carer benefit units include one adult with carer responsibilities. The carer state evolves through time, based on exogenously defined transition probabilities that vary by the individual's prevailing carer state, the disability state of their spouse (see Section 12.1) and/or dependent children (see Section 12.2), age, and year. Carers can be limited to benefit units with at least one adult who is sufficiently healthy, as defined by a pre-defined value of the health state.

Carers can be distinguished from other adults in regards to the benefits that they are eligible for (Section 6), their employment opportunities, and the time that they have available for leisure (Section 4).

12.4 Mortality

The survival of each adult is modelled individually, governed by the adult's age and prevailing simulation year. The timing of death, a^d, can be considered uncertain, and depends on age and year specific survival probabilities that are commonly reported as components of official life-tables by national statistics authorities. The same set of probabilities are used to simulate mortality in the population, as are assumed for individual specific expectations when evaluating the solution to the lifetime decision problem. The model is designed so that the associated input closely reflects the format of statistics reported by the Office for National Statistics in the UK, which it is hoped will facilitate sensitivity analysis of mortality assumptions.

All adults in the reference cross-section are assumed to survive until at least the age at which they appear in the reference cross-section, with death possible in any year subsequent to the reference cross-section. The age at death for each individual is evaluated by comparing the age specific elements of their health vector against the probability of death for an individual of the relevant age and year. If the element of the health vector is less than or equal in value to the respective probability, then the individual is assumed to die at the respective age.[21] Otherwise, they survive into the succeeding year. The age at death is consequently the first age at which the health vector is less than or equal to their respective probability of death.

13 International Migration

It is necessary to account for international migratory flows in any projection of the population cross-section. As the review by O'Donoghue et al. (2010) makes clear, there are a wide range of alternative approaches used to simulate the effects of migration in the microsimulation literature. Key modelling decisions include whether to model net migration or immigration and emigration separately, the variables that describe the likelihood of emigration, the approach taken to generate the characteristics of immigrants, and whether to accommodate re-entry of emigrants. These decisions depend upon the reasons for the respective model's development, and the data that are available for parameterisation.

[21] Note that the health vector used to adjust mortality is the same as is used to evaluate the evolving health state, which serves to economise on the data that need to be recorded to permit replication of a simulation. Each element of the health vector is adjusted after it is referenced for identifying health state to ensure that an independent random uniform number exists for evaluating mortality. For example, in the case where there are only two health states considered for analysis, then the model begins by considering whether the element of the health vector is under the relevant transition probability for distinguishing health state 1 from 2. If it is under the relevant transition probability, then health state is set to 1, and the element of the health vector is divided by the respective transition probability to obtain a new independent random number from a uniform distribution for use when evaluating mortality.

Migration has been accommodated in LINDA to meet two key objectives. First, the model should be capable of reflecting official projections for the age distribution of the population through time. Secondly, the model should reflect the bearing that contemporary trends concerning migration would have on the distribution of income if they were to continue into the projected future. Although the first of these two objectives could be achieved by modelling net migration only, this approach would complicate achieving the second objective. LINDA is consequently designed to accommodate explicitly both immigration and emigration in each simulated period.

There are two principal approaches for generating the characteristics of recent immigrants in a microsimulation context (e.g. Duleep & Dowhan (2008)). The 'regression' based approach involves estimating a system of equations describing all of the characteristics of interest, and uses these equations to generate characteristics for new immigrants. Valid implementation of this approach is, however, exceptionally challenging in any context where more than a few characteristics are involved, as is the case in the current context. LINDA consequently generates the characteristics of new immigrants using the alternative approach, by 'cloning' benefit units from 'donors' drawn from targeted population subgroups. This means that the approaches taken to simulate immigrants and emigrants share close similarities with one-another.

The model parameters include the total numbers of immigrants and emigrants to be assumed for each prospective year. The parameters also include the proportions of immigrants and emigrants to assume within a set of mutually exclusive and exhaustive population subgroups. These subgroups are defined with respect to age, education, marital status, and dependent children. Subgroups are further distinguished by disposable income quintiles for immigrants, and by past migrant status for emigrants. These model parameters permit evaluation of target numbers of immigrants and emigrants who fall into each considered population subgroup in each simulated year. The model divides the domestic population simulated for each year into the same subgroups distinguished for migrants, and randomly selects members from these subgroups as either emigrants, or to be cloned as new immigrants, to match migrant targets. Variables are generated that report the age of immigration, a^{im}, and emigration, a^{em}, for each simulated adult. It is also possible to distinguish between up to two source regions (in addition to the UK) for migrant flows, $region$, so that the model is capable of reflecting, for example differences between migrants by whether their country of origin was the host country, the EU, or some other country.

14 Behavioural Projections

A central feature of the model is the ability to base projections on decisions that solve the lifetime utility maximisation problem. It is also possible, however, to direct the model to project some aspects of behaviour using reduced form statistical equations. Where reduced form equations are used to project all decisions associated with a particular simulation, then the utility maximisation problem is omitted from analysis, and the model effectively takes the form of a classical dynamic microsimulation model. This section begins by outlining the methods used to solve the utility maximisation problem, before describing the reduced form behavioral descriptions that have also been implemented in the model.

14.1 Solution routines for utility maximisation and model validation

The value function considered by LINDA is not guaranteed to be either smooth or concave. Non-smoothness arises due to the focus of the model on decisions between discrete alternatives, the inclusion of decision making costs in the preference relation, the imposition of various constraint conditions, and the allowance for a flexible budget set that may be non-convex (due, for example, to means testing of welfare benefits). Non-concavities of the value function imply that the optimisation problem described in Section 2.1 can have local maxima. This observation, combined with the idiosyncratic nature of the model and its level of complexity, emphasises the importance of checks to determine the validity of model inputs and outputs, and for methods to determine the degree of numerical accuracy obtained.

Basic programming issues

A long list of checks and balances have been implemented to minimise the risk of error in the data generated by the model. Given the complexity of the programming problem, there is little substitute for intensive testing of model output. The single most important factor in minimising errors in the model is the time during which it has been in use: work first started on the model in 2003, and has continued without pause ever since. This relatively long process of development has led to an extensive system of internal checks and balances; the model will, for example, stop and report a critical warning message if it encounters any one of 142 problems that range from poor convergence properties of the numerical search routines, the validity of assumptions that underly the interpolation routines, to basic issues concerning the internal consistency of the model parameters. These checks have been introduced to capture the most common errors that have been identified during the model's development and both guard against erroneous model projections, and facilitate model development (as they help to identify programming errors introduced by new model code).

The long history of model development also helps to validate the introduction of new simulation routines by providing a tested basis for comparison. The typical development cycle involved when introducing new functionality into the model has three key stages. The first involves restructuring the model code to include the new functionality. This stage is facilitated by the adoption of a modular form for the code, in which each new individual characteristic is stacked on top of the last in a fixed order of priority. This means that the introduction of new functionality typically requires little variation of the existing model code. The second stage involves suppressing the new functionality, and checking to ensure that results produced by the altered model code are the same as those produced prior to the model amendments. The third stage is to activate the new functionality, and to check that the simulated output looks sensible. This third stage may itself run over a number of variants. Following introduction of a new pension asset, for example, it would be normal to allow the benefit units of reference adults to invest in the new asset, but specify returns to the asset that were so low that no rational individual should want to participate. Under these conditions, the model should produce results that are identical to those obtained in the absence of the new pension asset. Following such a test, the return on the new pension asset might be increased in increments, and the output checked to ensure that the influence on pension participation rates look sensible.

A toolbox of solution routines

Three key components of the solution to the lifetime decision problem can be subject to variation in the model: the solution detail can be increased, alternative interpolation methods can be used, and alternative numerical search routines can be employed. Source code relating to the last two of these components, upon which any dynamic programming model depends, are available from the authors upon request.

Increasing the solution detail involves increasing the size and number of points of the grid that overlays the feasible state space, and which is used to approximate utility maximising solutions at any conceivable combination of individual specific circumstances. It is also possible to increase the number of abscissae that are used to evaluate expectations over normal distributions of uncertainty via the Gaussian quadrature.[22] Increasing the grid points provides a more detailed solution of the utility maximising problem, though it also implies a rapid increase in computational burden; increasing the grid points in multiple dimensions increases the computational burden geometrically rather than arithmetically, a problem that is commonly referred to as the curse of dimensionality.

Linear or cubic interpolation methods may be employed by the model to evaluate behaviour between discrete grid points.[23] Relative to linear interpolation, cubic interpolation produces a smoother functional form, and ensures continuous differentiability. Cubic interpolation also requires evaluations at 4^n grid points, rather than 2^n points, where n is the number of dimensions over which the interpolation is taken. If cubic interpolation is used, then the model performs an internal check to determine whether the surface over which an interpolation is being taken is reasonably smooth, before selecting the cubic

[22] Evaluation of weights and abscissae of the Gauss-Hermite quadrature are based upon a routine reported in Chapter 4 of Press et al. (1986).

[23] The interpolation routines that are used are based on Keys (1981).

interpolation for analysis; otherwise, it conducts a linear interpolation.[24] It is of note that the cubic and linear interpolation routines have been programmed separately, and so can be used to validate against one another.

Three alternative numerical search routines can be used to identify solutions over continuous control variables to the lifetime decision problem. All three of these routines can search over multiple continuous dimensions, and are based only on value function calls. This approach is taken because first derivatives are not guaranteed in many problems for which the model has been developed. Even in conceptual contexts where well-behaved first derivatives do typically exist, our past experience suggests that the computation time required to evaluate first derivatives often exceeds any attendant gain in efficiency derived from search routines that identify a zero rather than a minimum/maximum.

The first search routine is based upon Powell's method in multiple dimensions, in conjunction with Brent's method in a single dimension, as described in Press et al. (1986). This approach combines parabolic interpolation with a golden section search, and previous analysis suggests that it is both robust in context of non-smoothness of the value function, and efficient where the value function is well behaved. The second search routine is based on the Nelder-Meade simplex search method, as described in Lagarias et al. (1998). The third method is based on the multi-level coordinate search method described in Huyer & Neumaier (1999), as implemented in the Numerical Algorithms Group (NAG) library. The model allows each of these alternative search methods to be applied in isolation, or in serial with the others, which offers a useful tool to determine whether results are driven by local and not global maxima.

14.2 Parallelisations

Solutions to the utility maximisation problem can require a great deal of time to compute, and the model is consequently designed to take advantage of parallel processing. 'Parallel processing' is the ability to divide a computational problem between two or more processes that work at the same time (in parallel). Two forms of parallelisations are implemented in the model.

14.2.1 Shared memory parallelisations

OpenMP is the application programming interface (API) that is used to take advantage of multi-core systems with shared memory, as are now present in most personal computing (including individual workstations). As discussed in Section 2.1, solutions to the utility maximisation problem are evaluated at intersections of a grid that overlays the state space of the decision problem. The states of the decision

[24] This involves distinguishing the "inner" 2^n points in closest proximity to the co-ordinate to be interpolated, from the "outer" $(4^n - 2^n)$ points considered in evaluating the cubic interpolation. If the smallest difference between any of the outer points and any of the inner points is more than 5 times the maximum difference between the inner points, then the model reverts to linear interpolation.

problem are organised in a strict hierarchy, with all states lower in the hierarchy being evaluated before consideration of alternative state values higher in the hierarchy. The approach taken to solve the lifetime decision requires that age be placed at the top of this hierarchy; thereafter the state order for solving the decision problem is arbitrarily defined, subject to the condition that discrete states (e.g. age, education, health condition, numbers of adults and children) are placed at the top, and continuous states (e.g. liquid net wealth, wage potential, private pension rights) at the bottom. This division between discrete and continuous states is computationally useful, because many of the calculations required to evaluate agent expectations vary by the respective discrete state combination. For example, all single adults in good health and without graduate education are subject to the same probabilities of marrying, having children, and transitioning to an alternative health state. By placing the discrete states at the top of the hierarchy, the computations required to evaluate a broad set of expectations can be evaluated once for all associated combinations of continuous state variables, and discarded when a new discrete state combination is considered.

Parallelisations implemented by OpenMP impose some computational burden, especially when a new computational process (or 'thread') is created or closed. At the same time, a substantial volume of data is typically required to describing agent expectations for each discrete state combination. The first of these considerations motivates inclusion of the parallelisation higher in the hierarchy, while the second precludes its extension into the discrete state space. Hence, a single OpenMP parallelised loop is included for the upper-most continuous state in the hierarchy adopted for solving the decision problem.

14.2.2 Distributed memory parallelisations

Most High Performance Computing (HPC) – sometimes referred to as supercomputer – environments are based on a distributed memory architecture, in which each computational 'worker node' is a separate computing entity that is essentially unaware of any progress made by other worker nodes, unless the relevant information is explicitly passed to them. Note that each worker node may (and usually will) include a multi-core processor with shared memory, in which case the OpenMP directives referred to above remain relevant. Something more, however, is required to co-ordinate the computations of individual worker nodes – MPI is the API that is used for this task.

The method used to divide the utility maximisation problem between worker nodes on a distributed memory system involves starting a unique version of the model on each worker node. All worker nodes perform precisely the same tasks, up to the point where utility maximisations start to be computed. These initial tasks include loading in model parameters, and establishment of the various 'containers' used to store solutions to the utility maximisation problem. From this point, computations for each age are divided between the available worker nodes by allocating each node an approximately equal share of the set of discrete state combinations that require solutions to the utility maximisation problem at that age. When all worker nodes complete their assigned share of the utility maximisation problem at a given age, they communicate their results to all other worker nodes, before proceeding to the immediately proceeding age. This process is repeated until utility maximising solutions for the full lifetime are evaluated, at which time a single worker node proceeds to generate data for a simulated population (see Section 2.1).

14.3 Reduced-form descriptions for behaviour

Four decisions can be simulated via reduced-form equations in LINDA: labour / leisure; consumption / saving; pension scheme participation; and the time of pension access. The functional descriptions adopted for all four of these decisions are broadly the same, represented by:

$$\begin{aligned} y_{i,a} \quad = \quad & \theta_1 nk_{i,a}^{0-5} + \theta_2 nk_{i,a}^{6-12} + \theta_3 nk_{i,a}^{13-17} + \theta_4 d_{i,a}^{student} + \theta_5 d_{i,a}^{graduate} + \theta_6 d_{i,a}^{pension} + \\ & \sum_{j=1}^{3} d_{i,a}^{ageband(j)} \sum_{k=1}^{10} \beta_{j,k} d^{decile(k)} + \sum_{l=-9}^{46} \alpha_l d_{i,a}^{age(l)} + \varepsilon_{i,a} \end{aligned} \tag{14}$$

where $y_{i,a}$ denotes the decision of benefit unit i at age a, nk^{x-z} is the number of dependent children in the benefit unit within the age band x to z, d^D denote indicator variables that equal to 1 when condition D is true and zero otherwise. *student* denotes (tertiary) students, *graduate* denotes (tertiary) graduates, *pension* denotes benefit units in receipt of retirement pensions, *ageband* denotes age bands distinguishing between three periods of life, *decile* denotes the respective decile group into which the benefit unit falls, *age* denote year specific ages, and ε denotes a residual term. Separate functions are considered for single adults and for couples.

Age and relationship specific deciles groups are evaluated using non-pension wealth for evaluating employment and the timing of pension access, earnings for evaluating pension scheme participation, and disposable income for evaluating total non-durable consumption (including non-discretionary expenditure, on which see Section 6). Given the structure of the model, this approach requires employment and the timing of pension access to be evaluated first, which permits earnings to be calculated. Having calculated earnings, pension scheme participation can then be identified, which permits disposable income to be calculated. Finally, given disposable income, consumption can be identified.

An ordered logit equation is used to project labour decisions. Logit equations are also used to evaluate pension scheme participation, and the timing of pension access. Finally, the expenditure/savings decision is represented by a reduced form equation, where the dependent variable is defined as log consumption less log disposable income.

Equation (14) was adopted because it has a number desirable properties. First, it is a parsimonious specification that depends only on a set of variables that are simulated by the model, which is a necessary pre-condition for use as a basis for projecting simulated decisions. Secondly, it takes into account important distributional relationships concerning the decisions of interest, wealth, earnings, and disposable income in a way that adapts in a plausible manner to temporal shifts in levels and dispersion.

15 Conclusions

This paper describes a structural dynamic microsimulation model. The model is designed to be a Lifetime INcome Distributional Analysis (LINDA) tool, to improve the evidence base for policy design and reform. The key distinguishing features of the model are that it is based on current best-practice (dynamic programming) methods for analysing savings and employment decisions, incorporates heterogeneity between individuals that is of primary interest to policy makers, and is capable of projecting circumstances forward and backward through time to build up lifetime histories for each individual represented in a reference population cross-section. It is also designed to project the evolving population cross-section forward through time, and to be used by policy analysts without any specialist experience.

Models that permit policy makers to consider the effects of policy measured over alternative time-horizons can help to distinguish the extent to which redistribution takes place between individuals, as distinct from between life-periods of individuals. Such information can be crucially important, both during the process of designing desirable policy reforms, and in explaining reforms to the population at large. It is also detail that facilitates an appreciation for the longer-term implications of policy, which is often absent from the contemporary public policy debate.

The last half century has seen a substantial improvement in the resources available to model builders, resulting in a proliferation of dynamic microsimulation models that are designed to project the implications of public policy for individuals through time. Coincident advances in our theoretical understanding of savings and labour supply decisions have, however, complicated the task of including the evolving theory into such models. This limitation is particularly relevant for dynamic microsimulation models, due to the extended time-periods over which such models are often used to project individual circumstances. A tax on saving, for example, may reduce savings rates in the near term, resulting in lower wealth in longer-term projections, and a consequent delay in the timing of retirement. Population variation of such effects can have important distributional implications that are exceedingly difficult to foresee without some analytical assistance. LINDA is designed to address this gap.

LINDA simulates the evolving circumstances for the benefit units of a population cross-section of reference adults forward and backward through time, and projects the evolving population cross-section forward through time. Dynamic programming methods are used to simulate endogenously a range of consumption/savings and labour/leisure decisions, which are considered to be made to maximise expected lifetime utility in context of an uncertain future. The model is designed to allow for differences between reference adults regarding their year of birth, age, relationship status, number of dependent children, health status, migration status, student status, education status, employment status (including employment sector), labour income, liquid net wealth, savings in Individual Savings Accounts (ISAs), private pension eligibility, private pension wealth, state pension rights, own business assets, timing of access to pension wealth, and time of death. Decisions that are endogenous to the model include (non-durable) consumption, labour supply, investments in risky assets, investments in ISAs and private pensions, and the timing of access to pension wealth. Uncertainty can be taken into account with respect to prospective labour market opportunities, investment returns, education status, relationship status, dependent children, health status, and time of death. Particular care has been taken to allow the model to reflect a detailed description of tax and benefits policy.

It is clear that the brief description of the model that is given in the preceding paragraph provides a blinkered version of reality. Key omissions from the list of characteristics that are explicitly taken into account by the model include distinguishing individuals by gender, and a formal recognition of housing wealth. Furthermore, the model does not include endogenous decisions regarding a range of important factors, including relationship status, fertility, intra-household allocations, housing investments, educational, and occupational choices. The view of the world that is presented by the model is also disconnected in the sense that macro-economic influences on the decision making environment are exogenously assumed. These restrictions are imposed on the model primarily due to the limitations of contemporary computing technology, which currently precludes the possibility of a single 'best' analytical tool for all purposes. LINDA is consequently designed to complement a suite of analytical tools – including dynamic microsimulation models that abstract from behaviour – rather than as a stand-alone all-purpose model framework.

Our on-going research agenda for LINDA extends along three principal dimensions. First, we continue to build upon the model's structure to mitigate the limitations that are outlined in the preceding paragraph. This developmental work is facilitated by advances in computing technology, and has accounted for the lion's share of our work with the model to date. Secondly, we are actively using the model to conduct empirical analyses designed to improve our understanding of the decisions that people make. Can relative risk aversion be accurately identified, how do decision making rigidities influence behaviour, how far is habit formation important in understanding the choices that people make? Improving our understanding of such questions are key areas of research interest. Finally, we are interested in using the existing model framework to explore the implications of policy alternatives. In this second respect, the most recent model incarnation is designed to tailor the model for use as part of the Joseph Rowntree Foundation's Programme to develop sustainable anti-poverty strategies for the UK. Such an objective is highly ambitious, and we believe a perfect context for consideration with the type of model described here.

References

Andreoni, J. (1989), 'Giving with impure altruism: applications to charity and Ricardian equivalence', *Journal of Political Economy* **97**, 1447–1458.

Attanasio, O., Bottazzi, R., Low, H., Nesheim, L. & Wakefield, M. (2012), 'Modelling the demand for housing over the life cycle', *Review of Economic Dynamics* **15**, 1–18.

Attanasio, O. P. & Weber, G. (1995), 'Is consumption growth consistent with intertemporal optimization?', *Journal of Political Economy* **103**, 1121–1157.

Attanasio, O. P. & Weber, G. (2010), 'Consumption and saving: Models of intertemporal allocation and their implications for public policy', *Journal of Economic Literature* **48**, 693–751.

Balcer, Y. & Sadka, E. (1986), 'Equivalence scales, horizontal equity and optimal taxation under utilitarianism', *Journal of Public Economics* **29**, 79–97.

Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B. & Rossiter, D. (2005), 'SimBritain: A spatial microsimulatino approach to population dynamics', *Population, Space and Place* **11**, 13–34.

Blundell, R., Browning, M. & Meghir, C. (1994), 'Consumer demand and the life cycle allocation of household expenditure', *Review of Economic Studies* **61**, 57–80.

Brewer, M., Browne, J., Emmerson, C., Goodman, A., Muriel, A. & Tetlow, G. (2007), Pensioner poverty over the next decade: what role for tax and benefit reform? IFS Commentary No. 103.

Browning, M. & Lusardi, A. (1996), 'Household saving: micro theories and macro facts', *Journal of Economic Literature* **34**, 1797–1855.

Carroll, C. (1992), 'The buffer-stock theory of saving: Some macroeconomic evidence', *Brookings Papers on Economic Activity* **23**, 61–156.

Carroll, G. D., Choi, J. J., Laibson, D., Madrian, B. C. & Metrick, A. (2009), 'Optimal defaults and active decisions', *Quarterly Journal of Economics* **124**, 1639–1674.

Choi, J. J., Laibson, D., Madrian, B. C. & Metrick, A. (2003), 'Optimal defaults', *American Economic Review* **93**, 180–185.

Creedy, J. (1985), *Dynamics of Income Distribution*, Basil Blackwell, Oxford.

Deaton, A. (1991), 'Saving and liquidity constraints', *Econometrica* **59**, 1221–1248.

Deaton, A. & Muellbauer, J. (1980), *Economics and Consumer Behavior*, Cambridge University Press, Cambridge.

Diamond, P. & Köszegi, B. (2003), 'Quasi-hyperbolic discounting and retirement', *Journal of Public Economics* **87**, 1839–1872.

Duleep, H. & Dowhan, D. (2008), 'Adding immigrants to microsimulation models', *Social Security Bulletin* **68**.

Edwards, S. (2010), 'Techniques for managing changes to existing simulation models', *International Journal of Microsimulation* **3**, 80–89.

Emmerson, C., Reed, H. & Shephard, A. (2004), An assessment of PenSim2. IFS WP04/12.

Fernandez-Villaverde, J. & Krueger, D. (2006), Consumption over the life cycle: facts from Consumer Expenditure Survey Data. Working Paper, University of Pennsylvania.

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

Friedman, M. (1953), The methodology of positive economics, *in* M. Friedman, ed., 'Essays in Positive Economics', University of Chicago Press, Chicago.

Gustman, A. L. & Steinmeier, T. L. (1986), 'A structural retirement model', *Econometrica* **54**, 555–584.

Gustman, A. L. & Steinmeier, T. L. (2005), 'The social security early entitlement age in a structural model of retirement and wealth', *Journal of Public Economics* **89**, 441–464.

Hancock, R., Wittenberg, R., Pickard, L., Comas-Herrera, A., Juarez-Garcia, A., King, D. & Malley, J. (2007), Paying for long-term care of older people in the UK: Modelling the costs and distributional effects of a range of options. PSSRU Discussion Paper 2336/2.

Hicks, J. R. (1939), *Value and Capital*, Clarendon Press, Oxford.

Hubbard, R. G., Skinner, J. & Zeldes, S. P. (1995), 'Precautionary saving and social insurance', *Journal of Political Economy* **103**, 360–399.

Huyer, W. & Neumaier, A. (1999), 'Global optimization by multi-level coordinate search', *Journal of Global Optimization* **14**, 331Ű355.

Keys, R. G. (1981), 'Cubic convolution interpolation for digital image processing', *IEEE Transactions on Acoustics, Speech, and Signal Processing* **29**, 1153–1160.

Kimball, M. S. (1990), ‘Precautionary saving in the small and in the large’, *Econometrica* **58**, 53–73.

Kuang, Y., Englebrecht, T. & Gilley, O. W. (2011), ‘A distributional analysis of the FairTax Plan: Annual and lifetime income considerations’, *Southern Economic Journal* **78**, 358–381.

Lagarias, J., Reeds, J. A., Wright, M. H. & Wright, P. E. (1998), ‘Convergence properties of the Nelder-Mead simplex method in low dimensions’, *SIAM Journal of Optimization* **9**, 112–147.

Laibson, D. (1997), ‘Golden eggs and hyperbolic discounting’, *Quarterly Journal of Economics* **112**, 443–477.

Li, J. & O’Donoghue, C. (2013), ‘A survey of dynamic microsimulation models: uses, model structure and methodology’, *International Journal of Microsimulation* **6**, 3–55.

Livshits, I., MacGee, J. & Tertilt, M. (2007), ‘Consumer bankruptcy: A fresh start’, *American Economic Review* **97**, 402–418.

McKay, S. (2006), Employers’ pension provision survey 2005. Department for Work and Pensions Research Report 329.

Muellbauer, J. & van de Ven, J. (2004), Estimating Equivalence Scales for Tax and Benefits Systems. NIESR Working Paper 229.

Nagatani, K. (1972), ‘Life-cycle saving: Theory and fact’, *American Economic Review* **62**, 344–353.

O’Donoghue, C., Lennon, J. & Hynes, S. (2009), ‘The Life-cycle Income Analysis Model (LIAM)’, *International Journal of Microsimulation* **2**, 16–31.

O’Donoghue, C., Redway, H. & Lennon, J. (2010), ‘Simulating migration in the PENSIM2 dynamic microsimulation model’, *International Journal of Microsimulation* **3**, 65–79.

O’Donoghue, T. & Rabin, M. (1999), Procrastination in preparing for retirement, *in* H. J. Aaron, ed., ‘Behavioral Dimensions of Retirement Economics’, Brookings Institute, Washington, pp. 125–156.

Orcutt, G. (1957), ‘A new type of socio-economic system’, *Review of Economics and Statistics* **58**, 773–797.

Press, W. H., Flannery, B. P., Teukolsky, S. A. & Vetterling, W. T. (1986), *Numerical Recipes: the art of scientific computing*, Cambridge University Press, Cambridge.

Pylkkänen, E. (2002), Modelling wages and hours of work, *in* E. Pylkkänen, ed., ‘Studies on household labor supply and home production’, Goteborgs Universitet, Gothenburg.

Rust, J. (2008), Dynamic Programming, *in* S. N. Durlauf & L. E. Blume, eds, 'New Palgrave Dictionary of Economics', Palgrave Macmillan, New York.

Rust, J. & Phelan, C. (1997), 'How social security and medicare affect retirement behavior in a world of incomplete markets', *Econometrica* **65**, 781–831.

Searle, J. (1984), *Minds, Brains and Science: The 1984 Reith Lectures*, Harvard University Press, Harvard.

Sefton, J. & van de Ven, J. (2004), 'Simulating household savings and labour supply: an application of dynamic programming', *National Institute Economic Review* **188**, 56–72.

Sefton, J. & van de Ven, J. (2009), 'Optimal design of means tested retirement benefits', *Economic Journal* **119**, F461–F481.

Sefton, J., van de Ven, J. & Weale, M. (2008), 'Means testing retirement benefits: fostering equity or discouraging savings?', *Economic Journal* **118**, 556–590.

Simon, H. A. (1955), 'A behavioural model of rational choice', *Quarterly Journal of Economics* **69**, 99–118.

Zaidi, A. (2007), The SAGE model: A dynamic microsimulation population model for Britain, *in* W. Barnett, ed., 'Modelling Our Future: Population Ageing, Health and Aged Care', Emerald Group Publishing Limited, London.

Zeldes, S. P. (1989), 'Optimal consumption with stochastic income: deviations from certainty equivalence', *Quarterly Journal of Economics* **104**, 275–298.

A Glossary

- *Additively separable*: A function that can be written as the sum of individual components. In context of intertemporal utility, 'additively separable' can be used interchangeably with 'time separable' if the utility function can be written as the sum of utility enjoyed in each of a number of discrete periods of the lifetime. Utility functions of this type imply that utility enjoyed in any given period depends on consumption in that period, and is otherwise independent of utility enjoyed in other periods. This rules out some interesting forms of behaviour, such as habit formation.

- *Application Programming Interface (API)*: A set of routines and tools to aid development of a computer program.

- *Closed form expression*: This concept is used here to denote any expression that can be fully evaluated using a finite number of operations limited to the endogenous and exogenous variables and parameters of a utility model.

- *Constant Elasticity of Substitution (CES)*: A property of some utility (and production) functions, *** FROM MY NOTES ***. A 'nested' CES function is one in which a CES function is taken of a CES function.

- *Continuous space*: A state that is closed, in the sense that all points between any two points within the space are also within the space.

- *Counterfactual context*: A context that has not existed in practice.

- *Discrete state*: A state that is comprised of points that are isolated from one another.

- *Distributed memory architecture*: a computing structure in which system memory is divided between computation nodes.

- *Dynamic microsimulation model*: A model that projects the circumstances of individual micro-units (microsimulation) through time (dynamic).

- *Exogenous*: Something defined outside of the system of interest.

- *Endogenous*: Something defined within the system of interest.

- *Global maximum*: The highest peak defined by a function.

- *High Performance Computing (HPC)*: A HPC, sometimes referred to as a supercomputer, is a computer that has a high computational capacity, relative to a personal computer.

- *Incentive compatibility*: A condition requiring that the assumed decisions of all individuals be consistent with their best potential outcomes, given their respective circumstances, expectations, and preferences.
- *Intertemporal*: Term meaning 'through time'. Intertemporal utility is a utility function that accounts for individual specific circumstances through time.
- *Local maxima*: The set of local peaks defined by a function.
- *Message Passing Interface (MPI)*: An API for facilitating co-ordination of analysis between multiple compute-nodes on a distributed memory computing system.
- Multi-core processor: A processor with multiple processing cores, capable of conducting multiple computations at the same time (in parallel).
- *Node*: A processing location, with a common set of shared resources (e.g. personal computer, laptop, or server-based resources).
- *Numerical search routine*: A trial and error search algorithm based on numerical comparisons. The objective of the search algorithm can vary widely, but often focuses on identifying a minimum, maximum, or zero of an assumed objective function.
- *Open Multi-Processing (OpenMP)*: An API designed to facilitate shared-memory parallel processing for a wide range of programming languages and operating systems.
- *Overlapping generations (OLG) model*: A model that represents multiple birth cohorts at any point in time, which is necessary feature to reflect a population cross-section.
- *Preference relation*: See Utility
- *Parallel processing*: Processing involving multiple computations undertaken simultaneously.
- *Quasi-hyperbolic discounting*: A form of intertemporal discounting that applies a higher discount rate in the short-run than the longer-run. This form of discounting is interesting because it is capable of capturing the potential that (some) people are myopic, in the sense that they exhibit a propensity to spend more at any given point in time than they would choose to do if they had the option to commit to a consumption plan at some prior point in time.
- *Reduced form*: A reduced form is the solution of a system of endogenous variables as functions of exogenous variables and model parameters. In econometrics, a 'reduced form' regression model usually explores potential relationships between variables, without explicit reference to an assumed theoretical framework.

- *Shared memory architecture*: a computing structure in which multiple processing cores share a common pool of memory.
- *State space*: The set defining all feasible combinations of state variables.
- *State variable*: A variable that is required to describe fully the prevailing condition of a dynamic process. In the current context, a state variable is an individual specific characteristic that is not perfectly described by some function of other individual characteristics.
- *Stochastic variable*: A variable influenced by at least one random term.
- *Structural model*: A structural model projects behaviour based on explicit assumptions concerning how decisions are made, and what aspects of the decision process can be taken to be invariant to the prevailing decision environment. In economics, attention has focussed on understanding behaviour as a product of the incentives that individuals face, commonly operationalised through the framework of utility maximisation. The central advantage of such models is they permit analysis of behaviour in alternative (counterfactual) decision contexts. In econometrics a 'structural form' regression model explores the empirical support for functional relationships between variables derived from an underlying theoretical framework.
- *Substantive rationality*: A form of expectations that simplifies the intertemporal utility maximisation problem by replacing some stochastic variables by their respective means, thereby reducing the multiplicity of potential future outcomes that it is necessary to take into account.
- *Thread*: A sequence of programmed instructions that can be managed independently.
- *Utility*: In economics, utility is a theoretical function describing the 'benefit' that an individual enjoys through consumption of an assumed set of variables. Conceptually, the idea behind utility is to translate incentives into behaviour (and vice-versa).
- *Value function*: The best possible value of the objective (utility) function, written as a function of the state variables
- *Worker node*: See node.

B Definition of Model Variables

Symbol	Definition	Control	State	Endogenous	Stochastic	Section
α	utility price of leisure					3
A	maximum age of life					2.1
a	age		X			3
a^d	age of death		X		X	12.4
a^{em}	age of emigration		X	X	X	13
a^{im}	age of immigration		X	X	X	13
a^P	age accessed private pensions	X	X	X		9.3
a^{SPA}	state pension age					8
β_0	exponential discount factor					3
$\beta_{1/2}$	short-run excess discount factors					3
B	value of bequests received					6
b	birth year		X			2.2
c	discretionary composite non-durable consumption	X				3
$carer^j$	carer status of each adult j		X		X	12
δ^{PP}	default decision for Personal Pension contributions					9.1
Δ	decision cost					3
D	age specific maximum debt					6
ε	elasticity of substitution, consumption and leisure					3
η	tax residual					2.2
$E_{a,b}$	expectations operator, person aged a of birth year b					3
ed	education status of benefit unit		X		X	10
$\phi^b_{j-a,a}$	survival probability to age j, given survival to age$a$ and birth year b					3
γ	coefficient of relative risk aversion					3
g	benefit unit employment income					4
g^j	employment income of adult j					4
h	benefit unit latent wage		X	X	X	4
$h^{\min}$	minimum wage					4
hh	home-owner identifier					6.2
$health^j$	health status of adult j		X		X	12
$health^c$	vector defining health status per dependent child		X		X	12
i	individual identifier					3
κ	experience effect on wages					4
k^{BSP}	contribution year to basic State Pension					8.1
k^{ISA}	transfers into and out of Individual Savings Accounts	X				7

Symbol	Definition	Control	State	Endogenous	Stochastic	Section
k^{ob}	transfers into and out of own-business assets					5
k^{S2P}	accrual of rights to State Second Pension					8.2
λ^{o}	job offer		X		X	4
λ^{emp}	employment effect on benefit unit earnings					4
λ^{ret}	retirement effect on benefit unit earnings					4
λ^{ng}	non-graduate earnings premium					4
λ^{hlth}	effect of health status on benefit unit earnings					4
l	proportion of benefit unit time spent in leisure	X				3
l^{j}	proportion of time spent in leisure by adult j					4
l^{ft}	leisure time when all adults full-time employed					4
m	parameters for earnings growth					4
md^{h}	mortgage debt					6.2
mh	mortgage holder identifier					6.2
$mort$	mortgage charges					6.2
n	number of adults in benefit unit		X		X	11.1
n^{c}	vector of dependent children by age		X		X	11.2
ndc^{c}	non-discretionary child care costs					6
ndc^{h}	non-discretionary health costs					6
ndc^{hg}	non-discretionary housing costs					6
ndc^{o}	other non-discretionary expenditure					6
ω	wage potential random innovation					4
pc^{c}	concessional pension contributions					9.1
pc^{nc}	non-concessional pension contributions					9.1
pc^{OP}	Occupational Pension contributions					9.1
pc^{PP}	Personal Pension contributions	X				9.1
PP	eligible Personal Pension type		X	X	X	9
py^{j}	retirement pension income received by adult j		X	X	X	6
py^{S2P}	rights to State Second Pension		X	X		8.2
ψ	parameter governing earnings persistence					4

Symbol	Definition	Control	State	Endogenous	Stochastic	Section
ρ	proportion of w^{nh} invested in risky assets	X				6.2
$region$	source region for migration		X	X	X	13
r^{ISA}	return to Individual Savings Account assets					7
r^{nh}	return to non-housing liquid net wealth					6.2
r^{ob}	return to own-business assets					5
r^{r}	return to risky liquid assets					6.2
r^{s}	return to safe liquid assets					6.2
$r^{S2P}_{r/w}$	return to State Second Pension rights					8.2
$rent$	rental charges					6.2
rr^{h}	realised returns to housing wealth					6.2
σ^2_{ω}	variance of ω					4
$sctr$	employment sector		X		X	5
$stud$	student status		X		X	10
τ	disposable income					6
θ	adult equivalence scale of benefit unit					3
ur^{h}	unrealised returns to housing wealth					6.2
w	liquid net wealth		X	X	X	6
w^{ob}	own-business assets		X	X	X	5
w^{OP}	wealth held in Occupational Pensions		X	X	X	9.2
w^{ISA}	Individual Savings Account assets		X	X	X	7
w^{h}	liquid net wealth held in housing					6.2
w^{nh}	non-housing liquid net wealth					6.2
w^{PP}	wealth held in Personal Pensions		X	X	X	9.2
w^{s}	non-housing liquid net wealth held in safe assets					6.2
w^{r}	non-housing liquid net wealth held in risky assets					6.2
w^{+}	liquid net worth that is non-negative					3
yr^{BSP}	years of contributions to basic State Pension		X	X		8.1
$\zeta_{0/1}$	parameters of warm-glow bequest model					3

C Modelling Methodology: a brief example

Suppose that LINDA is used to evaluate consumption and investment decisions between safe and risky assets. In the final period of life, given liquid net wealth w_A, the implied decision problem involves balancing motives for consumption, c_A, against those for bequests, $w_A - c_A$, to maximise expected lifetime utility. In this case, expected lifetime utility, $E(U_A)$, will correspond to simple utility $U_A = u(c_A) + b(w_A - c_A)$, as there is no possibility of a subsequent period of life. LINDA evaluates and stores consumption decisions that maximise U_A at a series of measures of liquid net wealth, $\hat{c}_A(\hat{w}_A)$. In addition to these measures of consumption, LINDA also stores the value function; $V_A(\hat{w}_A) = u(\hat{c}_A(\hat{w}_A)) + b(\hat{w}_A - \hat{c}_A(\hat{w}_A))$.

In the penultimate period of life, the decision problem involves selecting within-period consumption and investments in risky assets to maximise expected lifetime utility. In this case, expected lifetime utility will equal the sum of within-period utility from consumption and expected utility derived in the final period of life. Denote wealth held in the penultimate period by w_{A-1}, the proportion of wealth invested in risky assets ρ_{A-1}, and the probability of surviving to period A, given survival to period $A-1$, ϕ_{A-1}. Ignoring temporal discounting, and assuming that optimal decisions are taken in period A, we have: $E(U_{A-1}) = u(c_{A-1}) + \phi E\left\{V_A\left(w_A\left(c_{A-1}, \rho_{A-1}, w_{A-1}\right)\right)\right\} + (1-\phi) E\{b(w_{A-1} - c_{A-1})\}$. LINDA uses numerical search methods to identify the feasible measures of consumption and risky asset investments $\left(\hat{c}_{A-1}(\hat{w}_{A-1}), \hat{\rho}_{A-1}(\hat{w}_{A-1})\right)$ that maximise $E(U_{A-1})$, at a series of measures of liquid net wealth $\hat{w}_{A-1}$. When $\left(\rho_{A-1}, w_{A-1}\right) \neq 0$, then $w_A\left(c_{A-1}, \rho_{A-1}, w_{A-1}\right)$ is uncertain, and is assumed to be normally distributed. Expectations are then evaluated with reference to a discrete set of points in w_A using the Gauss-Hermite quadrature and interpolation over the solutions obtained for the final period of life, $V_A(\hat{w}_A)$. In addition to storing the optimised measures of consumption and risky assets investment, $\left(\hat{c}_{A-1}(\hat{w}_{A-1}), \hat{\rho}_{A-1}(\hat{w}_{A-1})\right)$, LINDA also stores the value function $V_{A-1}(\hat{w}_{A-1}) = u(\hat{c}_{A-1}) + \phi E\left\{V_A\left(w_A\left(\hat{c}_{A-1}, \hat{\rho}_{A-1}, \hat{w}_{A-1}\right)\right)\right\} + (1-\phi) E\{b(\hat{c}_{A-1} - \hat{w}_{A-1})\}$. The model then proceeds recursively to solve the utility maximising decisions for the remainder of the considered life course.

D Simulating Child Birth

Table 1: Schematic describing the division of life assumed for the purpose of evaluating fertility transition probabilities

	child birth period		
age	1	2	3
16			
17			
18			
19			
20			
21			
22			
23			
24			
25			
26			
27			
28			
29			
30			
31			
32			
33			
34			
35			
36			
37			
38			
39			
40			
	: child birth age		
	: ages covered by period		