THE UNIVERSITY OF MELBOURNE

FACULTY OF BUSINESS & ECONOMICS

Melbourne Institute Working Paper Series

Working Paper No. 20/13

Modelling the Dynamic Effects of Transfer Policy: The LINDA Policy Analysis Tool

Paolo Lucchino and Justin van de Ven

MELBOURNE INSTITUTE®
of Applied Economic and Social Research

Modelling the Dynamic Effects of Transfer Policy: The LINDA Policy Analysis Tool*

Paolo Lucchino[†] and Justin van de Ven[‡]

[†] National Institute of Economic and Social Research, London

[‡] National Institute of Economic and Social Research, London; Melbourne Institute of Applied Economic and Social Research, The University of Melbourne

Melbourne Institute Working Paper No. 20/13

ISSN 1328-4991 (Print)

ISSN 1447-5863 (Online)

ISBN 978-0-7340-4310-8

June 2013

* This study was made possible by funding from the Economic and Social Research Council, UK Treasury, and UK Revenue and Customs. The usual disclaimer applies. Corresponding author: Justin van de Ven, email <justin.van@unimelb.edu.au>.

Melbourne Institute of Applied Economic and Social Research
The University of Melbourne
Victoria 3010 Australia
***Telephone* (03) 8344 2100**
***Fax* (03) 8344 2111**
***Email* melb-inst@unimelb.edu.au**
***WWW Address* http://www.melbourneinstitute.com**

Abstract

This paper describes a structural dynamic microsimulation model that generates individual-specific data over a range of demographic and economic characteristics at annual intervals over the life-course. The model is specifically designed to analyse the distributional implications of policy alternatives in terms of their bearing on income and consumption measured over alternative time periods, from one year up to the entire life-course. This focus on economic characteristics measured over appreciable periods of life motivates endogenous simulation of savings and labour supply decisions, taking explicit account of uncertainty regarding the evolving decision environment. Reflecting the demands of policy makers, and in contrast to the majority of the associated literature, the model described here is designed to project from data observed for a population cross-section.

JEL classification: C51, C61, C63, H31

Keywords: Dynamic programming, savings, labor supply

1 Introduction

Good policy design is a fiendishly difficult business, due to the multiplicity and complexity 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. Interest in understanding how policy influences individual circumstances over alternative time spans is an important motivation for the development of dynamic microsimulation models. The model described in this paper represents the current state-of-the-art in this stream of research.

Most dynamic microsimulation models allow for behavioural responses to the policy environment via stylised analytical specifications (e.g. Pensim2 for the UK, or APPSIM for Australia). These stylised specifications are often based upon structural descriptions of decision making that have been selected for their analytical convenience (at least) as much as their plausibility. The reason for this is not that behavioural responses to the policy environment are considered unimportant; it is widely recognised that behavioural variation becomes increasingly important as the time horizon of analysis is extended. Rather, behaviour rarely lies at the heart of dynamic microsimulation models due to the computational burden that is involved. The rapid advance of computing technology during the last half century has consequently seen increasingly sophisticated attempts to integrate behavioural responses into dynamic models of individual circumstances.

Current best practice in the economic analysis of dynamic decision making is based upon dynamic programming (DP) methods. 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 decade, these models have not yet gained much traction within policy making institutions. This stands in stark contrast to the pervasive use made of current best-practice macro-economic models by monetary authorities. The LINDA (Lifetime INcome Distributional Analysis) model is designed specifically to address this gap, by making current best-practice micro-economic methods of behavioural analysis accessible to (UK) policy makers.

LINDA projects a range of demographic and financial characteristics for a sample of 'reference individuals' and their evolving families, treating labour supply and savings as endogenous. The model departs from most of the related DP literature by projecting from a population cross-section rather than a single birth cohort.[1] Projecting from a population cross-section substantively extends the capacity

[1] 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

of the model to consider issues of practical policy concern. Furthermore, projecting the circumstances of a population cross-section forward and backward through time presents considerable advantages for conducting associated empirical analyses, which we believe have not been fully exploited by the existing literature.

The empirical advantages obtained by projecting from a population cross-section are subtle and non-trivial. Empirical identification of savings preferences requires for identification behaviour observed over a considerable period of the life-course. If behaviour for a single birth cohort are considered for analysis, then we require data observed over a substantial period of time. This complicates the analysis, due to coincident variation of the underlying policy environment. A model that is explicitly designed to project from a population cross-section is capable of describing behaviour observed throughout the life-course at a single point in time, albeit for individuals drawn from different birth cohorts. As such, it permits preferences to be identified on cross-sectional survey data, which considerably simplifies the task of describing the incentives underlying observed behaviour. Although overlapping-generations models of savings in context of uncertainty do exist (e.g. Livshits et al. (2007)), we are unaware of any study that takes advantage of the possiblity to parameterise such a structure on cross-sectional data. This subject is taken up by our companion study Lucchino & van de Ven (2013), which discusses how the model that is described here has been parameterised to survey data for the UK.

This paper provides a technical description of the LINDA model, including discussion of how the issues that are specific to the use of DP methods have been addressed. As noted above, the principal motivation for creating LINDA is to make current best practice economic methods accessible to policy makers. This is an objective that we are keen to help facilitate more widely, and so we also discus the special needs of policy makers that have been drawn to our attention, and how we have gone about meeting those needs.

An over-view of the model is provided in Section 2. Sections 3 to 10 describe each model characteristic in turn. Details regarding the routines that solve the lifetime utility maximisation problem are described in Section 11, and a summary and directions for future research are provided in the conclusion.

2 Overview of LINDA

The model projects the evolving circumstances of a sample of reference adults at discrete intervals through time. The decision unit is the nuclear family, defined as a single adult or partner couple and their dependent children, which we refer to throughout as the 'family'. Intra-family allocations are not addressed. The model can be used to consider endogenous decisions regarding consumption, labour

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).

supply of adult family members, and the portfolio allocation across a range of assets that include safe and risky liquid investments, Individual Savings Accounts[2], and personal pensions. The model assumes that decisions maximise expected lifetime utility, given a family's prevailing circumstances, its preference relation, and beliefs regarding the future. 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
- self-employed / employee if in work
- wage earned if reference adult works
- wage earned if spouse works
- liquid wealth
- savings held in Individual Savings Account
- 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
- time of death

Of the 17 individual specific characteristics listed above, three are assumed by the model to evolve non-stochastically (year of birth, age, and the timing of access to private pension wealth, which can be made an endogenous decision) and all others may be uncertain. The preference relation assumed

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

by the model takes an additively separable nested Constant Elasticity of Substitution form that allows for quasi-hyperbolic discounting. Expectations are calculated assuming beliefs that are ‘substantively rational’, in a way that we explain 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 individuals described by an initial reference population through time. In the first stage, numerical methods are used to solve for utility maximising decisions, given any feasible combination of individual specific circumstances. The second stage uses the behavioural solutions identified in the first stage, in conjunction with assumed relationships governing the intertemporal variation of individual-specific circumstances, to generate panel data for the reference population. These panel data typically form the basis for conducting secondary analyses that are the principal focus of concern, including empirical evaluation of alternative behavioural assumptions, and exploration of implied behavioural responses to policy counterfactuals.

DP models of behaviour are now sufficiently well known that we provide only a broad outline of the actual steps involved 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 considered in 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). LINDA assumes that there is a maximum potential age to which any individual may survive, denoted by A. Given that time is divided into discrete intervals, the assumption of an upper bound on age implies that there exists a ‘final period’, after which death is certain. In this final period of life, the decision problem is non-stochastic, and is therefore straightforward 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.

Having obtained utility maximising solutions for the final period of life, the model then considers decisions at intersections corresponding to the penultimate period. Here, utility depends upon consumption in the penultimate period, and the impact that decisions in the penultimate period have on circumstances – 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 preceding stage of analysis. 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 obtain. 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 measure of optimised utility in the final period by drawing on '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.

The above routines allow expected utility in the final period to be evaluated for any given decision alternative in the penultimate period. Given this, 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.

Suppose, for example that LINDA is used to evaluate (non-durable) consumption and investment decisions between safe and risky assets. In the final period of life, given liquid 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 remaining uncertainty. LINDA evaluates and stores consumption decisions that maximise U_A at a series of measures of liquid wealth, $\hat{c}(\hat{w}_A)$. In addition to these measures of consumption, LINDA also stores the optimised measures of utility, $V_A(\hat{w}_A) = u(\hat{c}(\hat{w}_A)) + b(\hat{w}_A - \hat{c}(\hat{w}_A))$ (commonly referred to as the value function).

In the penultimate period of life, the decision problem considered here 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 and expected utility derived in the final period of life. Ignoring temporal discounting, we have: $E(U_{A-1}) = u(c_{A-1}) + E(U_A)$. Decision options that involve investing only in safe assets will imply that survival, ϕ, is the only uncertain factor. In this case, assuming that optimal decisions are taken in period A, and denoting wealth carried over to period A by w_A, we have $E(U_A) = \phi V_A(w_A) + (1-\phi) b(w_A)$. LINDA uses interpolation methods over the discrete set of wealth alternatives considered for period A, $\hat{w}_A$, to evaluate $V_A(w_A)$. Numerical search methods analyse feasible alternatives of c_{A-1} to maximise $E(U_{A-1})$. Expanding the decision set to allow for investments in risky assets introduces uncertainty over w_A, in addition to survival. In this case, LINDA assumes that the returns to risky investments are (log) normally distributed, so that expectations regarding future wealth, w_A, can be approximated as the weighted sum of evaluations taken at a discrete set of abscissae, N, via the Gauss-Hermite quadrature: $E\left(U_{A-1}\right) = \sum_{k=1}^{N} \alpha_k \left[\phi V_A\left(w_k\right) + \left(1-\phi\right) b\left(w_k\right)\right]$, where w_k denote the quadrature abscissae for w_A and α_k denote the weights.

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 families are simulated by 'running them through the grids'. Starting from a given population cross-section, the decisions of each simulated family are interpolated from the solutions calculated about their respective grid co-ordinates. Given a family's characteristics (state variables) and behaviour, its characteristics are then aged one period, based on the processes that are considered to 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 family circumstances backward through time, but with the addition of search routines that ensure that projected decisions and family characteristics satisfy incentive compatability constrains. That is, given a family's characteristics at time t, and a set of assumed decisions at time $t-1$, similar methods to those used to project circumstances forward through time can be used to identify the family's characteristics back in time to the start of period $t-1$. But having identified a family's characteristics to the start of time $t-1$, it is necessary to check that the decisions implied by the solution to the utility maximisation problem for the given set of characteristics at time $t-1$ are consistent with the assumed behaviour used to identify the associated characteristics.[3] These methods can be applied recursively to generate complete panel data for the life-course of each reference adult in the model.

It may be helpful to point out that the second stage of the simulation procedure that is referred to here is very similar to classical dynamic microsimulation. The distinguishing difference between the two approaches to simulation is that classical microsimulation replaces the first stage of the simulation described above with a series of (reduced form or structural) regression equations. Notably, it is in this first stage that the computational burden cited as the principal draw-back of the dynamic programming approach arises. Hence, the choice between the two approaches hinges upon the relative importance

[3]For example, we might assume that a family with £1000 at age 35 consumes £500 at age 34. Ignoring income eared at age 34, this combination of wealth and consumption would imply that £1500 of wealth was held by the family at age 34. It would be necessary, in this context, to ensure that the behavioural solution of the model implies that a family with £1500 at age 34 would want to consume £500 in order for the population simulation to be internally coherent.

assigned to dynamic decision making (the relative advantage of the DP framework) and population heterogeneity (the relative advantage of the classical microsimulation framework).

2.2 Using dynamic programming (DP) methods to simulate from a population cross-section

Although many of the issues involved in projecting a population cross-section through time have been discussed at length in the dynamic microsimulation literature, projecting from a population cross-section is a novelty for the associated DP literature. This section discusses the issues that are specific to DP models of savings and labour supply, and describes how projections for a population cross-section have been implemented in the current context.

Variation between the conditions faced by different individuals is commonly decomposed into time, cohort, and age effects. Until now, DP models of savings and labour supply have focused 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. We refer to this dimension here as the *birth year*. This raises two key design questions; what differences between birth cohorts will the revised model take into account; and how will those differences be integrated into the lifetime decision problem. We describe our responses to each of these questions in turn.

Following careful consideration of the list of characteristics that are explicitly treated by the model (see above), heterogeneity between birth cohorts was limited to the following nine factors:

- survival rates	- marriage rates	- marital dissolution rates
- fertility rates	- rates of return	- interest charges on debt
- wage rates	- transfer policies	- unemployment rates

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, we have observed a distinct weakening of domestic partner relationships (including the rise of cohabitation) and falling fertility rates, influencing 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 to debt). A birth cohort's relative advantage will often depend 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 regarding the welfare state.

Having defined the range of factors that are considered to vary between birth cohorts, the next problem is to define how the variation is incorporated into the lifetime decision problem. The approach that we have adopted is designed to ensure the model's computational feasibility. The smooth temporal transitions that have been observed for each of the four demographic factors – 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, for example, about the precise timing of their death, 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 to the remaining 'economic factors' – interest accruing to assets / debt, wage rates, transfer policies, and unemployment rates – due to the relatively volatile temporal variation that these factors exhibit. This conclusion is attributable to more than the associated conceptual inconsistencies that 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 of the 'economic factors' listed above

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-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 when they are each aged 43, which could reasonably be set equal to the average unemployment rate for 43 year olds reported by survey data.

Given the above, the model allows a population cross-section to be simulated through time in one of two alternative ways. The first alternative imposes the same assumptions for projecting a population through time as adopted for agent expectations. In this case, for example, we would assume that the return to safe liquid assets remains invariant through time. This is consistent with the classical assumption of perfect rationality. The second alternative 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 age-specific 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 family budgets is comprised of two key components in the model. The first is a highly flexible 'tax and welfare function' that is capable of reflecting a detailed description of policy. The second is random variation, implemented through a 'tax residual'. We envisage the tax residual will correct for differences between simulated and sample moments of disposable income, representing the difference between the model tax and welfare structure and that which applied in practice. 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).

2.3 Meeting the needs of policy makers

Three key criterion are at the heart of our effort to meet the modelling needs of policy makers; (1) to allow the analysis of the impact of policy measured in terms of lifetime income; (2) to provide a close reflection of population sub-groups of interest; and (3) to provide sufficient flexibility to allow government analysts to explore margins of interest without the need to seek third-party assistance. These three criteria have influenced every stage of the development of LINDA, and we discuss here a number of key effects that they have had on the model structure.

Analysis in terms of lifetime income

While much of the public debate on tax and welfare interventions often focuses on 'next-day' effects for 'winners and losers' of reform, many policy reforms are designed with a view to longer term effects. This is particularly notably for pension policies, but has increasingly been the case in relation to the work-incentive effects of taxes and welfare benefits in the UK. This motivates our interest in a tool that is sufficiently flexible to consider redistributive effects over alternative time horizons that include the entire life course. As noted in the introduction, increasing the time horizon of interest puts emphasis on behavioural responses to the underlying policy environment, which exaggerates the conceptual problems associated with reduced forms of analysis. This motivates our adoption of a structural framework for projecting the population through time, in which parameters that are assumed to remain invariant to policy change – so-called 'deep parameters' – are explicitly defined. By simulating individuals through time, and by considering optimal decisions at each point in time, the objective is to build up a sensible picture of how policy counterfactuals may bear out, now and well into the future. By focussing upon individuals strung out over the age and income/wealth distributions, the objective is to provide a framework that is suitable for exploring lifetime distributional implications for key population subgroups of interest.

Simulating from a population cross-section

To remain relevant to the needs of policy makers, the above stated focus on intertemporal effects of policies is implemented in a way that is sufficiently flexible to consider the implications for the contemporary population cross-section. Before presenting the advantages this may deliver to policy-makers, it is worth pointing out that projecting from a population cross-section is not the same as projecting the population cross-section through time. For example, the latter would include migratory flows or the birth of new population cohorts which this model makes no allowance for. Nevertheless, the cross-sectional design of LINDA will be intuitively attractive to policy-makers. It will allow them to map the simulated lifetime impacts back to actual groups in the current population, which is of clear political relevance. Doing so will also allow policy analysts to characterise the size and direction of the effects on (lifetime) income for each of these population groups, and to estimate the current numbers and characteristics of (life-cycle) winners and losers from policy. Finally, LINDA's cross-sectional approach is also more aligned to other models commonly used in Treasuries around the world. While beyond the current capabilities of LINDA, it is not inconceivable to think that its cross-sectional set-up may pave the way for successive extensions that might allow linking to existing macro models, or to deliver estimates of the tax yield or loss from simulated policy changes.

Calibration and validation

The unobserved preference parameters of a model that projects from a population cross-section can be identified on behaviour observed for a single population cross-section. This is notable, given that preference parameters are often a central focus of interest in the related literature. It is also extremely useful because it simplifies specification of the policy context underlying the behaviour considered for identification, and omits feed-back effects that can otherwise complicate adjustment of parameters. This is in contrast to a birth-cohort specific model of household sector savings of the type more commonly considered in the associated literature, which requires for identification behaviour representative for a given birth-cohort over a substantive period of the life course, with the implication that agent specific characteristics are endogenous and variable. Suppose, for example, that consumption implied by the model early in the life course was low, relative to survey data, indicating that preferences should reflect greater impatience. Adjusting preferences in this way would imply lower wealth later in life, which might then influence the timing of retirement. Such feed-back effects can be ignored in an empirical analysis of household sector savings when the analysis is based on a structural model of the population cross-section, which helps to simplify the identification problem considerably. Related issues are discussed at greater length in Lucchino & van de Ven (2013).

Nevertheless, the model that is described here involves complex interactions between a wide array of family specific circumstances. Validating such a model against survey data is a difficult and time-consuming problem. In the context of the UK, this problem is mitigated somewhat by the rich survey data that are available for validation, particularly since the advent of the Wealth and Assets Survey. The task is also made more transparent, from the perspective of the policy maker, by the accessible format of the micro-panel data that are generated by the model, which facilitates third-party evaluation (discussed below).

Inputs, outputs, and model design

The model described here, like most DP models of individual behaviour, is computationally intensive. This limits the scope for use of integrated analytical packages in which many of the necessary routines are pre-programmed, as most such packages involve a considerable loss of computational efficiency. The model described here has been programmed in Fortran, which is a computationally efficient language that has had extensive optimised libraries developed for it. Such an approach tends to complicate independent use of the model by policy makers. LINDA has been designed to off-set this drawback in three important ways. First, all inputs for the model are delivered through an Microsoft Excel spread sheet. A series of 'front-end' dialog boxes (user forms) walk a policy maker through all of the parameters that they are most likely to want to alter, each of which is assigned a plain-English description. Similarly, the model produces a standard set of output for each simulation to a series of Excel spread sheets, designed to facilitate preliminary analyses.

Secondly, the micro-data that the model generates for each simulated population are stored in a standard format that can be read by most statistical packages (comma separated variable, csv, format). These data are crucial, both to facilitate an on-going process of model validation, and because it is impossible, as a modeller, to anticipate all of the statistics that will be useful for any given analysis.

Finally, the tax and benefit routines that are a central focus of policy interest in the model are provided in a self-contained module (a dynamic link library) that policy makers can alter in whatever way they like, without the need for external consultation. Like the main program, this module is programmed in Fortran, which has the added advantage of being an accessible language for non-specialists.

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:

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

where $\gamma > 0$ is the (constant) coefficient of relative risk aversion; E_a is the expectations operator; A is the maximum potential age; β_1, β_2, and δ 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 consumption; $l_{i,a} \in [0,1]$ is the proportion of family time spent in leisure; $\theta_{i,a} \in R^+$ is 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 net liquid wealth when this is positive and zero otherwise.

The labour supply decision (if it is included in the model) is considered to be made between discrete alternatives. No upper limit is imposed upon the number of discrete alternatives, so that the labour supply decision can be made arbitrarily close to a decision over a continuous domain. Where adults are explicitly considered for analysis then a separate labour supply decision is considered for each adult. Otherwise, the family is considered to choose its labour supply in a similar fashion to a single adult.

The modified OECD scale assigns a value of 1.0 to the family reference person, 0.5 to each additional adult member and 0.3 to each child, and is currently the standard scale for adjusting before housing costs incomes in European Union countries. Its inclusion in the preference relation reflects the fact that family size has been found to have an important influence on the timing of consumption (e.g. Attanasio & Weber (1995) and Blundell et al. (1994)).[4] Similarly, decision costs are included in the preference relation to allow the model to reflect behavioural rigidities that have been cited as important in understanding retirement savings decisions (e.g. Choi et al. (2003) and Carroll et al. (2009) for the

[4] 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.

US and McKay (2006) for the UK). These costs are accounted for 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).

The model assumes that all discount parameters are the same for all individuals, and time invariant. It also assumes that families 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 easily 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 associated analytical problem, relative to alternatives that have been considered in the literature.[5] Including a bequest motive in the model raises the natural counter-party question of who receives the legacies that are left. The model assumes that all bequests accrue to the state, to avoid adding uncertainty bequest receipts to the decision problem.

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.

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

4 The Wealth Constraint

Equation (1) is maximised, subject to an age specific credit constraint imposed on liquid net worth, $w_{i,a} \geq D_a$ for family i at age a.[6] 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 (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 $a_D \leq A'$.[7] Liquid net worth is defined as the sum of safe liquid assets, $w^s_{i,a} \in [D_t, \infty)$, and risky liquid assets, $w^r_{i,a} \in [0, \infty)$. Intertemporal variation of $w_{i,a}$ is, in most periods, described by the simple accounting identity:

$$w_{i,a} = w_{i,a-1} - c_{i,a-1} + k_{i,a-1} + \tau_{i,a-1} \tag{3}$$

where k represents net investment flows with other asset classes in the model, and τ denotes disposable income net of non-discretionary expenditure. The only potential departures from equation (3) occur when a reference adult is identified as getting married or incurring a marital dissolution, in which case wealth is adjusted by a fixed factor to reflect the proportion of family assets accruing to the spouse.

As the model has been designed to undertake public policy analysis, particular care was taken in formulating the module that simulates the effects of taxes and benefits on family disposable incomes. The tax function assumed for the model can be represented by:

$$\tau_{i,a} = \tau(l_{i,a}, x_{i,a}, n_{i,a}, n^c_{i,a}, r^s_{i,a} w^s_{i,a}, r^r_t w^r_{i,a}, pc_{i,a}, b, a, \varepsilon_{b,a}) \tag{4}$$

which depends on age, a; labour supply, $l_{i,a}$; private non-capital income, $x_{i,a}$; the number of adults, $n_{i,a}$; the number and age of children (represented by the vector), $n^c_{i,a}$; the return to safe liquid assets, $r^s_{i,a} w^s_{i,a}$ (which is negative when $w^s_{i,a} < 0$); the return realised on risky liquid assets, $r^r_t w^r_{i,a}$ (possibly negative); private pension contributions, $pc_{i,a}$; birth year, b; and the tax residual, $\varepsilon_{b,a}$.

Non-capital income $x_{i,a}$ considered for tax purposes is equal to labour income $g_{i,a}$ plus the proportion of pension annuity income that is considered taxable. The model allows the measures of income accruing to each adult family member to be accounted for separately, so that it can reflect the taxation of individual incomes that is applied in the UK. Tax residuals are model parameters that can vary by birth year and age. The influence that these residuals have on disposable income is determined within the tax function, and so is otherwise unconstrained.

The rate of return to risky assets $\ln(r^r_t) \sim N\left(\mu_r - \frac{\sigma^2_r}{2}, \sigma^2_r\right)$ is assumed to be the same for all families at any point in time, t. The interest rate on safe liquid assets, $r^s_{i,a}$, is assumed to depend upon whether

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

[7] The lower bound D 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 11.

$w^s_{i,a}$ 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 $\left(g^{ft}\right)$:

$$r^s = \begin{cases} r^I & \text{if } w^s \geq 0 \\ r^D_l + \left(r^D_u - r^D_l\right)\min\left\{\frac{-w^s}{g^{ft}}, 1\right\}, r^D_l < r^D_u & \text{if } w^s < 0 \end{cases} \quad (5)$$

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 with time when projecting a population through time.

The form of the tax function described by equation (4) was selected to minimise the computational burden of the utility maximisation problem. For the purposes of taxation, and in a discrete time model such as this, investment income depends upon the timing of pension receipt. 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 require repeated evaluations of disposable income for each consumption alternative that is tested.

4.1 Projecting wealth backward through time

The issues involved in projecting family circumstances forward through time have been thoroughly discussed in the dynamic microsimulation literature. As noted in Section 2.1, dynamic programming methods do not change the fundamental nature of this problem, as such methods simply replace the stylised relationships that are otherwise used to describe family decisions. When projecting family wealth forward through time, for example, traditional microsimulation methods impute the consumption measure that is required to evaluate the associated accounting identity (see equation 3) using estimated regression equations. The dynamic programming environment applies similar methods, but replaces the estimated regression equations for consumption with an endogenous numerical specification that has been calculated to maximise an assumed preference relation.

Projecting a population backward through time is far less common in the associated literature, and introduces the additional problem of ensuring consistency between simulated decisions and evolving family circumstances. The relationship between consumption and wealth lies at the heart of this problem in the LINDA model: given liquid wealth at age $a+1$, consumption at age a, c^0_a, implies a measure of wealth at age a: $w^0_a = w_{a+1} + c^0_a - \tau_a$ (ignoring investment flows between asset classes k_a); at the same time, the solution to the dynamic program implies that a family with wealth at age a, w^0_a, will choose to consume c^1_a. The problem then is to search for a value of consumption, such that $c^0_a = c^1_a$.

We have found that the search problem described above is well-behaved, and we have therefore implemented a simple recursive routine to solve it. Consider the problem of projecting the circumstances of a family aged a, with characteristic vector v_a, back one period in time. An initial guess is made, that equates the family's decisions at age $a-1$, represented by the vector x_{a-1}^0, to their decisions at age a: $x_{a-1}^0 = x_a$. Given v_a and x_{a-1}^0, standard simulation methods are employed to identify the family's implied characteristics at age $a-1$, v_{a-1}^0. The solution to the dynamic programming problem is then referenced to identify the family decisions that are consistent with circumstances v_{a-1}^0; x_{a-1}^1. An adjustment rule is then used to identify a revised guess for the family decision vector $x_{a-1}^2 = \hat{\phi}x_{a-1}^0 + \left(1-\hat{\phi}\right)x_{a-1}^1$, $0 < \hat{\phi} < 1$, and the process repeated until the absolute difference between x_{a-1}^j and x_{a-1}^{j-1} is sufficiently small (defined as £0.01 of weekly consumption).

Two key factors ensure that, for the most part, the search problem defined above is well-behaved. First, as noted above consumption and wealth lie at the heart of the problem, and the marginal propensity to consume out of wealth is usually less than 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 suppress the multiplicity of alternative decision vectors that are consistent with a given set of family circumstances. Matters become slightly more complicated when either of these two factors break down.

The first factor identified above breaks down when families are liquidity constrained, in which case the marginal propensity to consume out of liquid 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 less useful 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.

5 Labour Incomes Dynamics

We have chosen to model latent wages at the family level, rather than for each individual, to simplify the model framework. The labour income of family i at age a, $g_{i,a}$, is given by:

$$
\begin{aligned}
g_{i,a} &= \lambda_{i,a} h_{i,a} \\
\lambda_{i,a} &= \lambda_{i,a}^{o} \lambda_{i,a}^{ed} \lambda_{i,a}^{emp} \lambda_{i,a}^{ret}
\end{aligned}
$$

where $h_{i,a}$ defines the family's latent wage, λ^{o} is an adjustment factor to allow for uncertain wage offers, λ^{ed} represents the shock to earnings of the reference adult attaining tertiary education, λ^{emp} adjusts for (endogenous) labour supply decisions, and λ^{ret} is the impact on earnings of taking up private pension income. Each of these alternative factors is described below.

5.1 Wage offers

Wage offers are included in the model to allow for the possibility of involuntary unemployment, which we have found to be important in matching the model to rates of employment during peak working years. Separate wage offers can be considered 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, and relationship specific probabilities $p^{o}\left(n_{i,a}, ed_{i,a}, a\right)$. If a wage offer is received by an individual, $o_{i,a} = 1$, then family income is an increasing function of their labour supply. If a wage offer is not received by an individual, $o_{i,a} = 0$, then any labour that the respective individual supplies returns no labour income to the family, implying non-employment where working incurs a leisure penalty.[8]

The solution to the lifetime decision problem assumes that families expect that the probability of a low wage offer is age, relationship, and education specific, but time invariant (as p^{o} is defined above). A population cross-section can be simulated through time, either assuming the same time invariant probabilities of a low wage offer that are assumed to solve the lifetime decision problem (consistent with perfect rationality), or allowing for historical variation in unemployment rates (to allow for fluctuations through the economic cycle).

5.2 Tertiary qualifications

The qualifications of each reference adult remain time-invariant for the greater part of the simulated lifetime, and can be defined to influence labour market opportunities (through p^{o} discussed above) and the evolution of relationship status (discussed in Section 10). The sole exception occurs toward the beginning of the simulated lifetime, when the model can be defined to consider the circumstances of

[8] We assume here that the disutility from a year of employment is more than sufficient to off-set the assumed experience effect on latent wages that is discussed below.

students in tertiary education. Two classes of tertiary student are permitted in the model, distinguished by their age at graduation, their probability of successful completion, and the influence that leaving education has on latent wage rates. The probabilities of gaining a qualification are exogenously defined for each student type. Similarly, the influence that leaving tertiary education has on latent wages is uncertain, and depends upon whether a reference adult is identified as achieving graduate status, $ed_{i,a} = 1$. The shock to latent wages of leaving tertiary education is represented as a random draw from a log normal distribution, $\lambda_{i,a}^{ed} \sim \ln N\left(\mu\left(ed_{i,a}\right), \sigma_{ed_{i,a}}^2\right)$ if reference adult i leaves tertiary education at age a, and is equal to 1 otherwise.

5.3 Employment and pension take-up

Family earnings respond to two endogenous decisions in the model. Each discrete labour alternative $l_{i,a}$ is associated with its own factor, $\lambda^{emp}\left(l_{i,a}\right)$. 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 $\lambda^{emp} = 1.0$. Furthermore, it is possible to impose wage penalties on families 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 family's latent wage, $\lambda_{i,a}^{ret} < 1$ if the family has accessed their pension wealth and one otherwise.

5.4 Latent wages

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

$$\log\left(\frac{h_{i,a}}{m\left(n_{i,a}, b, a\right)}\right) = \psi\left(n_{i,a-1}\right)\log\left(\frac{h_{i,a-1}}{m\left(n_{i,a-1}, b, a-1\right)}\right) + \kappa\left(n_{i,a-1}, a-1\right)\frac{\left(1-l_{i,a-1}\right)}{\left(1-l_W\right)} + \omega_{i,a-1} \quad (6)$$

where the parameters $m\left(.\right)$ account for wage growth (and depend on age, a, birth year, b, and the number of adults in the family, $n_{i,a}$), $\psi\left(.\right)$ accounts for time persistence in earnings, $\kappa\left(.\right)$ is the return to another period of experience, and $\omega_{i,a} \sim N\left(0, \sigma_{\omega,n_{i,a-1}}^2\right)$ is an identically and independently distributed family specific disturbance term. The only exceptions to this specification are when a reference adult enters or departs a cohabitating relationship, in which case the latent wage defined by equation (6) is adjusted by a fixed factor to reflect the impact of spouse earnings.

The form of equation (6) has a number of desirable properties that motivated its selection. First, 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). Secondly, the appearance of the $m\left(.\right)$ terms on both sides of equation (6) helps to simplify parameterisation of the model. Increasing $m\left(a\right)$ by 3 percentage points, for example, will *ceteris paribus* increase $h_{i,a}$ by 3 percentage points without also feeding through to increase $h_{i,a+1}$. This property is lost if the $m(.)$ terms are replaced by a single factor on the right-hand-side of equation (6). And thirdly, we have found that the addition of an experience effect to the wage equation helps 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.

5.5 Simulating latent wages backward through time

Simulating latent wages backward through time following the process described by equation (6) 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:

$$\log(h_{a-1}) = \log(m_{a-1}) + \frac{\log\left(\frac{h_a}{m_a}\right) - \left(\kappa_{a-1}\frac{(1-l_{a-1})}{(1-l_W)} + \hat{\omega}_{a-1}\right)}{\psi_{a-1}} \tag{7a}$$

$$\hat{\omega}_{a-1} = \frac{\sigma_\omega^2}{\sigma_a^2}\left(\log(h_a) - \mu_a\right) + \eta_{a-1} \tag{7b}$$

$$\eta_{a-1} \sim N\left(0, \frac{\sigma_\omega^2\left(\sigma_a^2 - \sigma_\omega^2\right)}{\sigma_a^2}\right) \tag{7c}$$

where μ_a and σ_a^2 are the mean and variance of log latent wages at age a, and relationship specific indicators have been dropped for simplicity. This specification makes use of the observation that the covariance between $\omega_{i,a-1}$ and $\log(h_{i,a})$, $cov(\omega_{i,a-1}, h_{i,a}) = \sigma_\omega^2$. If experience effects are omitted from the analysis ($\kappa = 0$), then equation (6) describes a standard regression-toward-the-mean model of earnings (e.g. Creedy (1985)), and we can set $\mu_a = \log(m_a)$. Introducing this restriction into equation (7) clarifies the dampening influence that allowing for $cov(\omega_{i,a-1}, h_{i,a}) = \sigma_\omega^2$ has on the dispersion of $h_{i,a-1}$.

6 Self-employment

Reference adults can be allowed to move into and out of self-employment in the model, where transitions from one year to the next are uncertain, and depend upon age and whether the reference adult was self-employed in the preceding period. In the first year that a reference adult switches into or out of self-employment, their earnings are drawn from a log normal distribution that depends only on their age, year, relationship status, and whether they are identified as self-employed. In all subsequent years, self-employed individuals are modelled in the same way as employees, subject to three exceptions. First, self-employed and employees are subject to different parameters for the dynamic wage process defined by equation (6). Secondly, if private pensions are considered for analysis, then the self-employed do not benefit from employer contributions and can be considered to contribute a different share of their earnings to their pension asset, relative to employees (conditional on pension participation). Thirdly, a separate "own business" asset can be included in the model.

If the own-business asset is included, then the value of this asset is set to zero when a reference adult first enters self-employment. The own-business asset accumulates a fixed fraction of family earnings during their period of self-employment, which are paid prior to the evaluation of income tax. The balance of the own-business asset for each self-employed person is illiquid during their period of self employment and attracts a rate of return, which can be considered uncertain. This asset is converted into liquid wealth when a reference adult moves out of self-employment, in which case the balance is subject to capital gains taxes.

7 Private Pensions

Private pensions are modelled at the family level, and are Defined Contribution in the sense that every family is assigned an account into which their respective pension contributions are (notionally) deposited. Up to five private pension schemes can be run in parallel, which differ from one another in their required rates of personal pension contributions, employer pension contributions, management costs, and default options. In each period a family can be considered to be eligible to participate in a single private pension scheme. Where only one private pension is considered for analysis, then the family can choose both whether to make fresh contributions to their eligible scheme, and how much to contribute. Pension contributions are specified as a percentage of employment income, so that they can only be made by working families. Where multiple private pensions are allowed, then the pension decision is restricted to the extensive margin (whether or not to contribute).[9] The timing of pension receipt can also be specified as an endogenous decision.

Up to four separate state variables are required to reflect pensions as they are described above. One state reflects the value of a family's accumulate pension wealth. The second state reflects which private pension scheme a family is eligible for at any given time. The third state reflects a family's default decision regarding private pension contributions to their eligible pension scheme. And the fourth state reflects whether a family has previously accessed their private pension wealth. A separate state is typically not required where the rate of return to pension wealth is assumed uncertain, as this state is assumed to be co-linear with the return to risky investments.

This structure is explicitly designed to focus on key issues that have been raised in the contemporary debate regarding pension provisions in the UK. It is sufficiently flexible to reflect the multiplicity of pension arrangements that is a feature of private pensions in the UK. It can reflect the influence of

[9] This is one limitation of the model that we hope to relax in the near future.

management costs, and the risk-return trade-off on decision making. It can also reflect the influence that default options regarding pension participation and contribution rates may have, as a result of investor inertia (e.g. Carroll et al. (2009), McKay (2006)). It is consequently possible, for example, to use the model to consider the capacity for auto-enrolment to increase private sector retirement savings. We begin by describing simulation of the pension accrual phase, before discussing pension dispersals.

7.1 Pension accruals

The private pension scheme to which a family is eligible in any given time period is determined by a stochastic process. Suppose, for example, that the model allows for three alternative private pension schemes, A, B, and C. Further, suppose that a family chooses to work and to contribute to their eligible pension B at age 35. For any family 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 families that choose not to contribute to their eligible pension. Assume here that this probability is set to 85%, where the remaining 15% can be thought to represent the probability of job change from one period to the next. The model then identifies into which of three income bands the latent wages of the family falls. We refer to latent wages here because the model considers for this purpose a measure of earnings potential, and not actual earned income (on which see Section 5). Each income band has a separate set of probabilities describing the likelihood of eligibility to the alternative pension schemes. Suppose that our family 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 family will be eligible to 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 family through time using standard Monte Carlo methods.

Decisions regarding pension contributions over both the intensive and extensive margins can be influenced by assumed 'default options', which may vary between alternative pension schemes. This is achieved by allowing for 'decision costs' that discontinuously reduce welfare when a family decision departs from the pre-assigned default. Defaults regarding the scale of pension contributions are defined as fixed percentages of labour income. Defaults regarding whether or not to contribute to a pension are defined when a family first becomes eligible to a given scheme (i.e. auto-enrolment or active opt-in), and thereafter track a family'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.

Furthermore, pension contributions are subject to four series of constraints. First, contributions to private pensions can only be made if employment income exceeds a minimum threshold, π_l^P. Upper bounds are imposed on per-period 'concessional contributions' that are subtracted from earnings before calculating a family's tax burden. Upper bounds are also imposed on per-period 'non-concessional contributions' that do not enjoy relief from the prevailing period's income tax, but do benefit from the tax shielding of investment returns that is assumed within a pension fund. The limits on both concessional and non-concessional contributions can be varied with respect to three age bands. And fourth, an upper bound is placed on the aggregate value of a family's private pension wealth. Any private pension wealth in excess of this maximum is assumed to accrue to the government.

Pension balances are assumed to be perfectly portable between schemes, and each family is assumed to hold all of their pension wealth in the scheme for which they are eligible at the given point in time (which evolves stochastically as described above). At the start of each period, each pension scheme is assumed to accrue the same (post-tax) rate of return, r_t^P, which can be specified as uncertain. When returns to private pensions are assumed to be uncertain, then they are perfectly correlated with the returns to the risky liquid asset (r_t^r in Section 4). Hence, accrued pension rights do not hedge against uncertainty in the liquid asset portfolio. Management charges are also levied at the beginning of each period, and these can vary between pension schemes, so that the net return to each scheme may differ (equal to r_t^P minus the management charge specific to the respective scheme, mc^P). Intertemporal accrual of pension wealth, w^P, during the accrual phase for a family that chooses to contribute to its eligible pension scheme is described by equation (8).

$$
\begin{aligned}
w_{i,a}^P &= \left(r_{t-1}^P - mc^P\right) w_{i,a-1}^P + \left(\pi^P + \pi_{ec}^P\right) g_{i,a-1} \\
\ln\left(r_t^P\right) &\sim N\left(\mu_P - \frac{\sigma_P^2}{2}, \sigma_P^2\right), corr(r_t^P, r_t^r) = 1
\end{aligned}
\tag{8}
$$

where π^P is the private contribution rate of the family's eligible scheme, π_{ec}^P is the employer's contribution rate, g represents (pre-tax) labour income (as in Section 5) and $corr(.)$ is the correlation coefficient. As with liquid wealth, the only departure from equation (8) is following relationship transitions, when pension wealth is adjusted by a fixed factor to reflect the proportion of family assets accruing to the spouse.

7.2 Pension dispersals

The timing of pension dispersals can be determined either endogenously or exogenously in the model. At the time that a pension is taken-up, a fixed fraction of the family's accrued pension wealth is received as a lump-sum cash payment, and the remainder converted into an inflation adjusted life annuity. The annuity rates assumed for analysis are calculated with reference to the survival rates assumed for individual birth cohorts, an assumed return to capital, and an assumed transaction cost levied at the 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 families are assumed to access their pension wealth at their respective state pension ages (which may vary by birth cohort). When the timing of pension dispersals is endogenously determined, then each family is free to decide when to access its accumulated pension rights. Furthermore, limits can be imposed on a family'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 families).

7.3 Simulating pensions backward through time

The objective to simulate family characteristics backward through time raises additional complications to those described in Section 4.1 for variable family characteristics that exhibit substantial persistence; the decision of when to start drawing down pension wealth is one example. Suppose, for example, that the model is projecting back through time the circumstances of a family that is identified as already having started to draw down their pension wealth. The model must identify precisely when the family first started to access their pension wealth, which is complicated by the existence of multiple ages at which take-up of private pension income will be optimal, given the assumption that pension income was not taken up at some preceding age.

LINDA addresses this problem by projecting back in time two separate life histories for each family 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 family 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 family accesses its pension wealth at the latest possible opportunity that has not been shown to be inconsistent with the solution to the dynamic programming problem. 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 family unit that is identified as having previously accessed its pension wealth, with a reference adult who is aged 72 in the initial reference population. Life history A proceeds on the assumption that pension wealth was first accessed by this family 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 family characteristics at age 70. In this case, life history B assumes that the family 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 family, 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 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.

8 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 strict upper limits. And thirdly, ISAs impose no limits on when accumulated funds are 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 family 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 family 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. In common with the approach taken to simulate pensions, uncertain returns to ISAs are assumed to be perfectly correlated with the returns to the risky liquid asset (r_t^r in Section 4). In most periods, wealth held in an ISA, w^{ISA}, is assumed to vary intertemporally as described by the equation:

$$\begin{aligned} w_{i,a}^{ISA} &= r_{t-1}^{ISA} w_{i,a-1}^{ISA} + k_{i,a}^{ISA} \\ \ln\left(r_t^{ISA}\right) &\sim N\left(\mu_{ISA} - \frac{\sigma_{ISA}^2}{2}, \sigma_{ISA}^2\right), corr(r_t^{ISA}, r_t^r) = 1 \end{aligned} \tag{9}$$

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 (9) is when the relationship status of a reference adult is identified as changing, in which case a fixed factor adjustment is used to alter ISA wealth to reflect the proportion of the account accruing to the reference adult's partner.

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 family's ISA is made.

9 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 2006/7, each in a form that is sufficiently flexible to capture reforms that were implemented in the 2006 Pensions White Paper.

9.1 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 family, 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 family 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 5). In most years prior to state pension age, the number of years of accreditation to National Insurance contributions is defined by:

$$yr_{i,a}^{BSP} = yr_{i,a-1}^{BSP} + k_{i,a-1}^{BSP} \tag{10a}$$

where $k_{i,a-1}^{BSP}$ are the additional contributions accredited to family i at age $a-1$. The only exception to equation (10a) is when the relationship status of a reference adult is identified as changing, in which case a fixed factor adjustment is used to alter $yr_{i,a}^{BSP}$ to reflect the proportion of the contributions history accruing to the reference adult's partner.

Each family is assumed to draw its basic State Pension from state pension age, and this public transfer is added to non-employment income for tax purposes. The value of the BSP payable to each family depends upon the contributions history of the family, the value of the full BSP assumed for the reference year, a growth rate applied until the time when the reference adult of the family 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 family is paid the greater of the single allowance, paid in respect of the number of complete contribution histories accrued by all adult family members, and the couple allowance, paid in respect of a single adult's contribution history. 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.

9.2 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 family level, are accumulated prior to state pension age, and are associated with a (taxable) income stream from state pension age. Unlike the BSP, the model tracks rights to the S2P in the form of the associated annuity income. The annuity to which a family 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 family with earnings in excess of the Lower Earnings Threshold, LET_t, is assumed to gain a flat-rate increase in their S2P rights, θ_t. This flat-rate benefit can also be defined to accrue in respect of child care and involuntary unemployment, as described above for the BSP. Family 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 θ_2^{S2P}: $\theta_2^{S2P}\left(g_{i,a} - LEL_t\right)$. Family earnings in excess of the UAP_t have no bearing upon S2P rights.

In most periods, rights to the S2P follow:

$$y_{i,a}^{S2P} = r_{r/w}^{S2P} y_{i,a-1}^{S2P} + k_{i,a-1}^{S2P} \tag{11}$$

where $k_{i,a-1}$ denotes the additional rights to the S2P derived by the family'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 a fixed factor adjustment is used to alter $yr_{i,a}^{S2P}$ to reflect the proportion of the contributions history accruing to the reference adult's partner.

10 Allowing for Family Demographics

Family demographics in LINDA depend upon the survival probabilities of reference adults, the intertemporal evolution of the relationship status of reference adults, and the birth of dependent children to reference adults. All three of these characteristics are considered to be stochastic events, and LINDA assumes that agent expectations are fully consistent with the underlying transition probabilities (as discussed in Section 2.2). We describe the approach taken to simulate each characteristic in turn.

10.1 Modelling survival

The model focusses upon survival with respect to reference adults only; the mortality of the spouses of reference adults is aggregated with divorce to obtain the probabilities of a relationship dissolution that are discussed below. Survival in the model is governed by age and year specific mortality rates which are commonly reported as components of official life-tables by national statistics authorities. 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.

10.2 Modelling 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 considered to be uncertain. The transition probabilities that govern relationship transitions depend upon a reference adult's existing relationship status, age and birth year, and may also vary with respect to their education status or the presence of dependent children. These probabilities 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 and whether there are dependent children in the family.

The model does not distinguish family units on the basis of characteristics that are specific to the partners of reference adults. This approach assumes that any difference between families that is due to spouse-specific attributes is captured by other aspects of the model. The age distribution of partners of reference adults who are currently 70 years old, for example, will be implicit in the assumed probabilities that 70 year old reference adults experience a marital dissolution (due to the associated influence that it has on the probability of spousal mortality). Similarly, the wage dispersion of spouses will be reflected in the variance of the innovation term assumed for the latent wages of couples in the model ($\sigma^2_{\omega,n_{i,a-1}}$ in relation to equation 6).

10.3 Modelling 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 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 family, children are assumed to remain dependants until an exogenously defined age of maturity. A child may, however, depart the modelled family 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 family 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).[10] 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 family 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 22, 26, or 33 years of age.

Capturing realistic family sizes in the 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 family 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 the model 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 Table 1. 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 considered for analysis.[11] 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

[10] This assumes an age of maturity of 17.
[11] Sensitivity analysis suggested that a simulated population size of 1,000,000 family units obtains reliable fertility transition probabilities in an applied context considered for the UK.

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		

transition probabilities for each birth age are then calculated by aggregating up the data calculated for the simulated population within the respective age band (as illustrated in Table 1). Further detail regarding this aspect of the model can be obtained from the authors upon request.

10.4 Simulating dependent children backward through time

As noted in Section 7.3, simulating family characteristics backward through time raises additional complications for variable family characteristics that exhibit substantial persistence. Dependent children clearly fall into this category. It would be conceptually straight-forward to apply the same techniques as described for the retirement state to simulate dependent children backward through time. The exogenous and stochastic nature of the approach adopted to simulate relationship status and children, however, motivated the adoption of a ‘trial-and-error’ approach. Evolution of the family demographics of reference adults depends entirely upon the assumed random draws that are used to evaluate uncertainty when simulating the reference population through time. As the processes involved in simulating population demographics are relatively simple, LINDA takes repeated trials of random draws for the life-course to the age at which each reference adult first enters the model (their age in the reference population) until a life-course profile is identified that is consistent with the relationship status and number and age of dependent children that they are recorded as having in the reference population.

11 Details of Solution Routines 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 output, and for methods to determine the degree of numerical accuracy obtained.

11.1 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 during the last 10 years. 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 128 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.

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 a minimum of variation to 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 families 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 rates of participation look sensible.

11.2 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, the interpolation methods can be altered, as can the numerical search routines.

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.[12] 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.[13] 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 being 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 interpolation for analysis; otherwise, it conducts a linear interpolation.[14] It is of note that the cubic interpolation, 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 to the lifetime decision problem. Two of these routines focus on value function calls, and one is based upon the associated Euler conditions. Running the model using each of these alternative solution procedures, but otherwise holding the parameters of the model fixed, provides a method for validating the solutions that are obtained.

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

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

[14] 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.

All three search routines are based upon the same set of procedures for selecting utility maximising decisions beyond consumption. The optimisation procedure is comprised of three discrete loops. In the outer loop, each feasible labour supply choice is independently tested, and a selection is based upon the (optimised) expected lifetime utility that is evaluated for each. The intermediate loop searches over feasible combinations of pension contributions, ISA contributions, and portfolio allocations between safe and risky liquid assets. Where this intermediate loop is conducted over a two dimensional domain, then the search is undertaken using Powell's method. Where the search is conducted over a single dimension, then Brent's method with or without reference to derivatives can be employed.[15] The inner-most loop searches over consumption, based upon one of three alternative options that are discussed at further length below.

- *A brute-force solution*

Following French (2005), a routine is included with the model that performs a comprehensive search of the entire (bounded) domain for consumption that is defined by the imposed budget constraint. Here, the search domain is explored by the following loop:

1. the search domain, $[lw, up]$, is divided into nn equally spaced points, where: $nn = \min(50, (up - lw) * 4 + 1)$
2. the value function is evaluated at each of the nn points
3. the top mm points are identified, where: $mm = nint(nn/5)$
4. the limits of the domain defined by the mm points are identified
5. the proportional reduction in the search space is calculated
6. if the proportional reduction in the search space is less than 60%, then multiple local maxima are assumed to have been identified for the value function. In this case, the search space is divided in half, and the search continues over each sub-space independently
7. when the search domain is reduced to less than £5.0 (so that the difference between nn points in the preceding iteration is no greater than £0.50 per week), then the search over each identified subspace is concluded, and the maximum evaluation is selected.

This approach has the advantage that it does not assume any form of regularity regarding the nature of the value function. It's disadvantage is that it can be computationally demanding, and therefore time consuming.

[15]The three search routines referred to here are based on Press et al. (1986).

- *Systematic search over the value function*

The search over feasible consumption alternatives can be undertaken without reference to derivatives, based upon Brent's method. This approach combines parabolic interpolation with a golden section search, and has been found to be efficient, particularly where the surface over which the search is conducted is reasonably well behaved. Although this approach tends to solve the utility maximisation problem much faster than the brute-force approach that is described above, it is not explicitly designed to distinguish between local and global optima. To account for the possibility of multiple local optima, the model includes a *supplementary routine* that conducts tests about identified optima. In this case, the model will explore a localised grid above and below an identified optimum for a preferred level of consumption, based upon value function calls. If an alternative value of consumption is identified by this supplementary routine as strictly preferred to the original local maximum, then the routine will search recursively for any further solutions above and below. This process is repeated until no further solutions are found. Of all feasible solutions, the one that maximises the value function is selected.

- *Systematic search based on Euler conditions*

In some cases, the value function may be sufficiently well-behaved to search for solutions to the decision problem with reference to Euler conditions. Using the Euler condition to solve the utility maximisation problem has two advantages: first, it represents a discrete alternative to the two search routines that are described above, which are both based upon repeated evaluations of expectations over the value function; and second, numerical routines that search for a zero can be faster than those that search for a maximum. The model consequently allows the optimisation problem to be solved using Euler conditions, in which case the Bus & Dekker (1975) bisection algorithm is applied. The supplementary routine that is discussed above can also be applied here to guard against the identification of undesirable local optima.

12 Conclusions

The hurdles to accessing current best practice economic methods of intertemporal decision making have limited the use of such methods to just a few specialist practitioners. Two principal uses can be identified for these methods of analysis. The first is to explore empirically alternative assumptions regarding agent behaviour. The second is to explore quantitatively the behavioural implications of policy counterfactuals. Although the former of these research fields may have suffered little from restricted access to specialist practitioners, the same cannot be said for the second. Further, it may be claimed that the whole point of obtaining a better understanding of the decisions that people make is to improve the evidence base for anticipating how behaviour will evolve in the future. In this regard, we must hope that the substantial research effort that has been spent on improving our understanding of intertemporal decision making will soon deliver considerable benefit to those charged with anticipating future behavioural trends, given the sparse nature of the benefits that have been derived to date.

LINDA is a model that is designed specifically to facilitate this objective. By packaging away technical details, providing easy access to a range of policy-relevant parameters, and generating descriptive data for a population cross-section that are easily accessible, it is hoped that this model will facilitate the use of current best practice economic methods of behavioural analysis by policy makers in Whitehall.

LINDA simulates the evolving circumstances for the families of a population cross-section of reference adults forward and backward 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, student status, education status, employment status (including an allowance for the self-employed), labour income, liquid wealth, savings in Individual Savings Accounts (ISAs), pension eligibility, timing of access to pension wealth, and time of death. Decisions that are endogenous to the model include (non-durable) consumption, 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, 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 health and disability status, migration, and housing. Furthermore, the model does not include endogenous decisions regarding a range of important factors, from relationship status and fertility, through to education and occupational classification. 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.

The restrictions imposed on the model, including those listed above, are motivated by two key factors: the limitations of contemporary computing technology, and the development of our technical understanding. Both of these factors are likely to recede with time, and LINDA's development is designed to take account of this in two important ways. First, the technical specification of the model takes a modular structure that is designed to facilitate the introduction of new characteristics as computing power is enhanced. Secondly, in addition to offering access to current best practice economic methods of intertemporal behavioural analysis to policy makers, LINDA is also a state-of-the-art research tool. The principal focus of our prospective research agenda is centred around use of LINDA 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, and making the results of the associated research rapidly accessible to policy makers are the principal motivations driving LINDA's development.

References

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

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

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

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

Bus, J. C. P. & Dekker, T. (1975), 'Two efficient algorithms with guaranteed convergence for finding a zero of a function', *ACM Transactions on Mathematical Software* **1**, 330–345.

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. & 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.

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.

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.

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.

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

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

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

Lucchino, P. & van de Ven, J. (2013), Modelling the dynamic effects of transfer policy: the LINDA policy analysis tool. National Institute Discussion Paper 405.

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, T. & Rabin, M. (1999), Procrastination in preparing for retirement, *in* H. J. Aaron, ed., 'Behavioral Dimensions of Retirement Economics', Brookings Institute, Washington.

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

Rust, J. (2008), Dynamic Programming, *in* S. N. Durlauf & L. E. Blume, eds, 'New Palgrave Dictorionary 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.

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.

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