Health, Disability Insurance and Labour Supply: Evidence from a Dynamic Structural Model

Chiara Dal Bianco*

19th October 2017

Abstract

This paper provides a quantitative assessment of the labour market and welfare effects of the disability insurance (DI) benefit. To this end, I develop a life-cycle model in which individuals choose consumption, labour supply and whether to claim for DI. The effects revealed by counterfactual policy simulations are largely heterogeneous by health; low-health and poor individuals place a higher value on DI. If there is a reduction in the benefit amount, only half of those who leave the benefit return to work. Policies that reduce the cost of re-entering the labour market by 10% increase the labour supply by 5.3 percentage points among DI recipients, without decreasing welfare.

Keywords: disability insurance, labour supply, older workers, health
JEL-Codes: J14,J26,I1

**Department of Economics and Management, University of Padova, Via del Santo 33, 35123 Padova, Italy (e-mail:*chiara.dalbianco@unipd.it*).*
For their useful comments and suggestions, I thank James Banks, Mariacristia De Nardi, Andrea Moro, Stephen Pudney, Giacomo Pasini, Guglielmo Weber and Francesca Zantomio. I also thank those who participated in the 2016 VIU Summer Institute on Ageing, the IBEO 2016 Workshop (Institution Individual Behavior and Economic Outcome), the 32nd European Economic Association meeting and the 29th Conference of the European Association of Labour Economist. Data from The English Longitudinal Study of Ageing obtained through the UK Data Archive have been used with permission. Funding from the Farmafactoring Foundation is gratefully acknowledged. All responsibility for the analysis and interpretation of the results lies with the author.

1 Introduction

Recent studies have shown that disability benefits play an important role in the departure of older workers from the labour market (Wise, 2016). In OECD countries, public spending on disability amounted to 2.1% of GDP in 2013 and has been quite stable over time, but the rules governing disability programmes have changed considerably in recent decades with more stringent medical criteria and an increase in integration policies, such as the provision of employment support and vocational rehabilitation (OECD, 2010).

One of the policymakers' objectives is to design cost-effective programmes that provide income support to the disabled without reducing the labour supply of those capable of working. To achieve this goal, it is important to understand what factors affect the persistence of disability benefits enrollment and how individuals' labour supply responds to programme eligibility requirements, transfer amounts and policies that reduce the cost of work.

In this paper, I provide a quantitative assessment of these effects for individuals approaching retirement age. To this end, I adopt a structural approach, in the spirit of the structural labour supply and retirement literature. This methodology allows me to account for the full dynamic effects of policies on agents' choices and to evaluate the welfare effects of the alternative disability insurance (DI) structures.

I develop and estimate a model of labour supply and savings behaviour for males living with a partner at the end of their working life using UK data. In the model, individuals choose whether to participate in the labour market and how many hours to work. Moreover, if their health level is below an estimated threshold they can enter the DI programme and remain in the programme in the following years without the need for health reassessment.

The model allows for uncertainty about wage realisation, health development and life expectancy. In developing the model, special attention has been devoted to the measure of health and the evolution of health over time. I construct a continuous health index using a set of objective health indicators collected in the English Longitudinal Study of Ageing (ELSA) replicating the same health conditions covered by the health assessment used to determine eligibility for DI benefits. The specified process for health has a deterministic component that depends on age and a stochastic component allowing both persistent and transitory shocks. Health status enters the deterministic component of the exogenous wage process (productivity channel) and the probability of surviving to the next period; moreover, there is a time cost of being in poor health that affects utility through leisure.

I contribute to the existing literature in several ways. I focus on older workers and propose a richer model than in previous structural research. Labour supply decision (both the extensive and the intensive margins), DI application, private and public pension accrual and pension claiming decision are modelled in a unified framework to shed light on the role of DI as a potential retirement pathway. Additionally, the model assumes a continuous, more comprehensive measure of health than previous models, which have greatly simplified the treatment of this variable for computational reasons or data limitations. This allows to evaluate the policy effects by health level.

I estimate the model using ELSA data from 2002 to 2008, a period in which DI policies and parameters were relatively stable. The model parameters are estimated in two steps. First, the parameters of the exogenous health and wage processes are estimated using standard minimum distance techniques. Second, the remaining parameters are estimated using the Method of Simulated Moments to match profiles generated by the dynamic model with data life-cycle profiles of assets, labour supply participation, hours worked and DI participation. The model is able to replicate the main patterns observed in the data and heterogeneity by age and health in decision profiles quite well. I use the model to simulate alternative reform scenarios that are considered in the political debate: altering the health requirements needed to qualify for DI benefits, changing the benefit amount and introducing policies that promote claimants' return to work.

I document a number of relevant findings. There is large heterogeneity by health in the effects of policy interventions, suggesting the importance of targeting DI benefits to truly needy individuals. In particular, individuals with fewer assets and poorer health place a higher value on DI; individuals with assets and health in the first quartile of the distribution are willing to pay 36% of their assets at age 50 to be compensated from benefit removal. The value is close to zero for those with health and assets in the fourth quartile. This heterogeneity in the welfare effect derives in part from heterogeneous labour supply responses to changes in the DI structure. For example, the elasticity of labour supply non-participation to benefit generosity is about zero for individuals in very poor health and is 0.45 for individuals with health close to the DI eligibility threshold. This supports the design of policies that distinguish between those with any working capacity left and those in need of permanent support but also highlights that of individuals leaving the benefit programme due to a reduction in the benefit amount only half return to work. A policy intervention that reduced the cost of work by 10% (such as policies that improve non-discrimination and accessibility in the workplace to foster labour supply inclusion of individuals with disabilities) is promising, as it would increase labour supply by 5.3 percentage points and reduce DI participation by the same amount, without decreasing welfare.

The remainder of the paper is structured as follows. Section 2 reviews the literature on the effects of DI on labour supply and welfare and describes the UK DI programme. Section 3 presents the model of lifetime decision making. Section 4 introduces the data and the health measure. Section 5 presents the estimation strategy and estimation results. Section 6 shows the model's ability to replicate main patterns observed in the data. Section 7 presents results from several policy experiments that investigate the effects of alternative DI structures on DI rate, labour supply and welfare. Section 8 consists of a summary of the main results.

2 Related literature and institutional context

The literature on DI programmes has mainly focused on the effects of DI on the labour supply. In his 1989 seminal paper, Bound compares rejected and allowed applicants in the US to show what would have been the labour supply of DI beneficiaries had they not received the benefit. He shows that 30 to 50% of denied applicants are working, and this represents an upper bound of the negative labour supply effect of DI. von Wachter *et al.* (2011) using detailed administrative data enrich Bound's findings and document the presence of an even higher unused working capacity among younger claimants applying for mental health and musculoskeletal conditions. Bound's results have been proved robust to the use of more causal identification strategies, such as a discontinuity in the eligibility criteria for a relevant sub-sample of applicants (Chen and van der Klaauw, 2008) and the random assignment of examiners to applications (Maestas *et al.*, 2013; French and Song, 2014).

Even though much of the literature focuses on the US institutional context, the role of DI as one of the biggest social insurance programmes is critical in most OECD countries (OECD, 2010). The large heterogeneity in the benefit designs and the many reforms in the last decades offer a natural laboratory to explore the effect of alternative benefit structures on labour supply. The generosity of the benefit amount, the stringency of the screening process and various incentives to remain in or return to work have often been used as instruments to reduce benefit inflow and moral hazard problems, control benefit duration and increase benefit outflow. Several contributions in the literature have exploited the exogenous variations provided by these reforms to estimate the labour supply responses to changes in DI.

Gruber (2000) uses a regional variation in benefit generosity in Canada and estimates a labour supply elasticity to benefit amount between 0.28 and 0.36 for older workers. Mullen and Staubli (2015) use the exogenous variations generated by a series of reforms to DI and the old age pension system during the 1990s and 2000s in Austria and estimate a DI participation elasticity to benefit amount between 0.7 and 1.2, with younger and older individuals in low-skill jobs being less respondent to benefit changes.

de Jong *et al.* (2011) use data from a controlled experiment in the Netherlands in which examiners were instructed to screen applications more strictly in a subset of regions. They find that stricter screening procedures reduce the attractiveness of the DI programme. Disney *et al.* (2006) and Ball and Low (2014) look at the introduction of a stricter health assessment in the UK disability programme in 1995. They provide evidence of a better targeting of those in need and reject the hypothesis that the reform increases the severity of the health shock required to claim disability. In 1996 in Austria, the age at which conditions to receive DI are relaxed was increased from age 55 to age 57, Staubli (2011) explores this exogenous variation and finds that the resulting increase in labour supply is only half the decrease in DI participation. Karlström *et al.* (2008)'s findings for Sweden go in the same direction; the abolition of more generous DI application rules for older claimants in 1997 do not lead to an increase in labour supply.

Campolieti and Riddell (2012) show that the introduction of an annual earnings disregard in 2001 in Canada has a sizable positive effect on the probability of doing any paid work while receiving benefits for both men and women. Using Austrian data, Ruh and Staubli (2015) find that the earnings cap has a negative effect on labour supply among DI beneficiaries. Kostøl and Mogstad (2014) find a positive effect on labour supply of a policy introduced in 2005 in Norway that allows DI recipients to keep a portion of their benefit if they return to work; however, no effect is found for DI recipients approaching retirement age.

In their authoritative survey paper, Bound and Burkhauser (1999) emphasise the prevalent focus of DI empirical literature on the DI effect on labour force attachment, whereas less attention has been devoted to the social and welfare value of DI programmes in providing protection against the economic consequences of the onset of a disability. Very few papers in the literature have tried to measure the value of DI; using longitudinal data, Meyer and Mok (2013) for the US and Ball and Low (2014) for the UK quantify the value of DI looking at consumption drops that follow a deterioration in the health of individuals.

Recent papers have addressed the importance of considering the insurance and the incentive aspects of DI programmes jointly in a life-cycle framework to quantify the welfare value of DI and the labour supply and welfare effects of alternative benefit structures. Bound *et al.* (2010) specify a discrete choice dynamic programming model for the behaviour of older workers in the US. They account for the endogeneity of health status to labour market behaviour and model health as a latent variable having self-reported disability as an indicator. They find that health has an important role in explaining earlier exits; however, changes in the DI programme structure do not have large effects on the probability of applying for the benefit. Benítez-Silva *et al.* (2011) focus on older workers in the US and simulate behavioural responses to a reduction of $1 in the DI benefit for every $2 in earnings instead of the current 100% tax for DI beneficiaries with earnings above the permitted threshold. They estimate a small effect of the reform. Low and Pistaferri (2015) recognise that disability applications in the US are also increasing among younger individuals. They develop a life-cycle model in which the DI screening process is carefully modelled and agents face several sources of risk—health shocks, productivity shocks and labour market frictions. They use self-reported disability to measure health and focus on welfare and behavioural effects of several DI reformed scenarios. They find that the increase in welfare generated by more generous programmes exceeds the negative effect on incentives.

The extensive reduced form literature surveyed in this section suggests the importance of considering the heterogeneity by health and age of the DI incentive effects and highlights the lower responsiveness of individuals approaching retirement age to DI benefit changes. It also provides evidence of the effectiveness in terms of labour supply response of altering specific parameters of the DI programme but with little attention to the welfare effects of such changes. In this paper, I focus on individuals approaching retirement age and build on the structural retirement literature that accounts for both financial incentives and the role of health to explain labour market transitions (French, 2005; French and Jones, 2011), considering state provided DI among individual possible choices. Previous structural work focusing on older workers (Bound *et al.*, 2010; Benítez-Silva *et al.*, 2011) does not consider the heterogeneity in the labour supply by health level. Low and Pistaferri (2015) focus on DI application over the entire life cycle and therefore do not model retirement decision and financial retirement incentives.

Much of the structural literature on retirement and DI has been developed for the US institutional context. However, institutional features vary significantly across countries and considering different institutional settings offers the opportunity to enrich our understanding of the incentive/insurance trade-off. In this paper, I consider the UK institutional setting, characterised by a liberal welfare state such as the US. The DI benefit amount in the UK is flat and not earnings-related as in the US, providing a different source of variation for identification. Moreover, eligibility conditions do not require people to have been unemployed for several months before applying and there are no long waiting periods to examine applications, which simplifies the analysis.

A long season of reforms begun in the mid-nineties has reduced the generosity of the benefit and introduced stronger working incentives to reduce public spending on DI. The most recent reform in 2008 replaced the Incapacity Benefit (IB) with Employment and Support Allowance (ESA). Additional changes to the benefit structure are currently under discussion, and a benefit cut has been introduced for a specific group of claimants in 2017.[1]

The ESA programme introduced in 2008 is in line with recent reforms in OECD countries to achieve a new balance between income security and labour market integration for people with disabilities. A Work Capacity Assessment, stricter than the previous health test, determines eligibility for the benefit and classifies claimants into two groups—the Support group and the Work-related Activity (WRA) group. If classified as able to follow work-related activities, individuals have to attend the Pathways to Work programme[2], while those in the Support group are entitled to the benefit without additional requirements. From 2011 to 2014, existing IB claimants were reassessed and those who were eligible moved to ESA. The effects of this most recent reform are less clear cut. Banks *et al.* (2015) use administrative data, the Labour Force Survey and ELSA to sum-

[1] See Appendix B for details on benefits targeted to disabled individuals in the UK.

[2] Adam *et al.* (2010) investigate the economic impact of the Pathways to Work pilot programme implementated in 2003. They find that the positive impact is mostly on individuals who would have left the benefit in the first year in any case but who would have stayed out of the labour market with a higher probability in the absence of the programme.

marise recent trends and heterogeneity in disability benefit receipt, focusing on the effect of the 2008 reform. They show evidence of an improved targeting of the benefit to those more in need, but this is true mainly for females. Comparing disability benefit receipt, employment rate and unemployment benefit receipt between 2008 and 2012 for different levels of disability, the authors find no strong and conclusive evidence that the decline in disability benefit receipt is associated with increased employment and/or unemployment.

3 The model

In this section, I present the life-cycle model of savings, labour supply, private pension and IB claiming decisions for males aged 50 or above living with a partner. The model accounts for several crucial institutional features affecting such decisions. The financial incentives provided by the pension systems are considered by modelling accrual in both public (earnings related) and private pensions. Along with the contributory health-related benefit (IB), financial incentives provided by non-contributory and means-tested benefits are included in the model by reproducing eligibility for Disability Living Allowance and Attendance Allowance as well as Income Support (including the premium for low-income households containing at least one disabled individual) and Working Tax Credits (which has a supplement for disabled workers). The agents face uncertainty regarding health developments, wage offer realisations and life expectancy. Assets are defined at the household level, and spousal income is modelled as a deterministic function of the male's characteristics. For sample size reasons (in ELSA data 70%, of males above 50 live with a partner) and to investigate the behaviour of a homogenous sample of individuals, I focus on couples who share the same insurance channels in facing health and income shocks.

In the model, a household head seeks to maximise his expected lifetime utility of the form:

$$U(c_t, l_t) + E_t \left[\sum_{j=t+1}^{T+1} \beta^j \Pi^s(j-1,t) \left(\pi_j^s U(c_j, l_j) + (1-\pi_j^s) b(a_j) \right) \right], \qquad (3.1)$$

where $t = 1, \ldots, T$. In each period t, the individual receives utility U_t from consumption c_t and leisure l_t. When he dies, he values bequest of assets according to a bequest function $b(a_t)$ with a_t assets at time t. Let β be the discount factor, $\Pi^s(j,t)$ be the probability of living to age j conditional on being alive at age t

and π_t^s be the probability of being alive at time t conditional on being alive at time $t-1$. The household head maximises Equation (3.1) by choosing consumption c_t, hours worked h_t, whether to apply for IB d_t and whether to claim private pension p_t, subject to a budget constraint.

The within-period utility function is a CRRA, non separable in consumption and leisure, of the form:

$$U(c_t, l_t) = \frac{1}{1-\nu}(c_t^{\gamma} l_t^{1-\gamma})^{1-\nu}. \tag{3.2}$$

The parameter γ represents the consumption weight; the lower the γ the greater the weight placed on leisure. The parameter ν represents the relative risk aversion coefficient and the elasticity of intertemporal substitution of the consumption and leisure composite good, for which a Cobb–Douglas aggregator is used. The elasticity of intertemporal substitution of consumption, holding the labour supply fixed, is given by $1/(\gamma * (\nu - 1) + 1)$. Hours of leisure l_t are defined as follows:

$$l_t = L - h_t - \phi_H(\hat{H} - H_t) - \phi_{P_t}\mathbb{1}(h_t > 0) - \phi_{d_t}\mathbb{1}(d_t = 1), \tag{3.3}$$

where L is the time endowment, ϕ_H is the time cost of being sick and $(\hat{H} - H_t)$ is a measure of sickness, obtained as the highest possible level of health $(\hat{H})$ minus the current level of health of the individual (H_t)[3]. If the health status (H_t) worsens and ϕ_H is positive (as expected), leisure (l_t) will decrease and because leisure is a normal good, the marginal utility of leisure will increase. Following French (2005) and French and Jones (2011), the cost of being in poor health enters the utility as a time cost, and the same is true for the cost of participating in the labour market ϕ_{P_t} and the cost of receiving disability benefit ϕ_{d_t}. Both ϕ_{P_t} and ϕ_{d_t} are allowed to vary with age[4], such that $\phi_{P_t} = \phi_{P_0} + \phi_{P_1} t$ and $\phi_{d_t} = \phi_{d_0} + \phi_{d_1} t$.

The bequest function is specified following De Nardi (2004):

$$b(a_t) = \phi_B \frac{(a_t + K)^{(1-\nu)\gamma}}{1-\nu}. \tag{3.4}$$

The parameter K, which is positive, regulates the curvature of the bequest func-

[3] I do not consider the medical expenditures channel which is particularly important in other institutional contexts (De Nardi *et al.*, 2010, 2016), because in the UK universal healthcare is provided. Medical expenses should not be so relevant, at least between age 50 and 70 before the costs of institutionalisation arise (they are not covered by the Health Care system).

[4] The fixed cost of work varies with age in a number of studies, including French and Jones (2011) and Rust and Phelan (1997).

tion and allows the utility of a zero bequest to be finite. The parameter ϕ_B represents the intensity of bequest motives.

Health evolves according to a stochastic process with a deterministic component, $\omega_H(age_t)$, which depends on age, a persistent component (the autoregressive component θ_t) and a transitory component (the iid shock η_t):

$$\log H_t = \omega_H(age_t) + \theta_t + \eta_t$$
$$\theta_t = \rho_H \theta_{t-1} + \nu_t^H, \qquad \nu_t^H \sim N(0, \sigma^2_{\nu_H}), \qquad \eta_t \sim N(0, \sigma^2_\eta). \tag{3.5}$$

The process for wages has a deterministic component, $\omega_w(H_t, age_t)$, which depends on health and age. Persistence in wages is captured by the autoregressive component ϵ_t.

$$\log w_t = \omega_w(H_t, age_t) + \epsilon_t$$
$$\epsilon_t = \rho_w \epsilon_{t-1} + \nu_t^w, \qquad \nu_t^w \sim N(0, \sigma^2_{\nu^w}). \tag{3.6}$$

I assume that at time $t-1$ the individual knows θ_{t-1} and ϵ_{t-1}, but he only knows the distribution of the innovations ν_t^i and of the temporary shock η_t.

Following French (2005), I do not directly model the joint decisions of the couple, but I account for the presence of the partner by including in the head of household's budget constraint the spousal income ys_t as a function of the individual's age, after tax labour and pension income, $ys_t = ys(income_t, age_t)$. I assume that marital status does not change over the period considered, either for separation or for death of the partner.[5]

The probability of surviving to period $t+1$ given that the individual is alive in period t, π^s_{t+1}, is a function of age and health in period t. I assume that the probability of surviving to age T+1 conditional on being alive at age T is zero ($\pi^s_{T+1} = 0$), and I set T equal to 90.

Finally, the asset accumulation equation is of the form:

$$a_{t+1} = a_t + y(w_t h_t, ra_t, ib_t, pb_t, sb_t; \tau) + ys_t + tr_t - c_t$$

where $y(\cdot, \tau)$ is after tax income; r is the real interest rate; ib_t is the IB amount,

[5]This parametrisation has the advantage of keeping the model simple while accounting for the fact that the head of the household does not rely only on his own income. However, this simplification does not allow exploring and fully accounting for the insurance role of female labour supply within a couple (Attanasio *et al.*, 2005).

received if the individual has health below the threshold $\bar{H}_d$ and claims for it ($d_t = 1$); pb_t is private pension benefit, received from the year in which the individual claims for it; sb_t is state pension benefit and tr_t represents non taxable transfers (such as Income Support, Pension Credit and Disability Living Allowance dla received if health is below the threshold $\bar{H}$ (with $\bar{H} < \bar{H}_d$)). The tax function, τ, and the modelled benefits are described in great detail in Appendix G. I assume that individuals cannot borrow against future pension income and means-tested benefits ($a_t \geq 0, \forall t$).

Each individual can be endowed with a private retirement plan. Two different plans are modelled—a Defined Contributions private plan ($pen = DC$) and no private plan ($pen = NO$).[6] In each period t, in addition to their private plan endowment, individuals observe their age, health status H_t, the private pension claiming status p_{t-1}, the disability benefit claiming status d_{t-1}, the amount saved in the DC fund (q_t^{DC}), the state pension accrual (q_t^{SP}) and the wage offer (w_t). They then choose whether to claim for private pension ($p_t = 1$), hours to work ($h_t \in [0, \bar{h}]$), whether to apply for IB ($d_t = 1$) and consumption (c_t). IB can be claimed up to the state pension age (SPA), which is 65 for males in the period considered. I assume that at age 70 everyone is retired.

3.1 Disability benefits

At each age between 50 and 64, individuals with health below a certain level ($\hat{H}_d$) can decide to claim for IB, a state-provided disability insurance in force between 1995 and 2008. If they were already claiming the benefit in $t-1$ they can continue to receive the benefit irrespective of their health status.

According to the rules, IB eligibility is conditional on having paid enough contributions in the three years before the start of incapacity. However, if the condition is not met, the applicant can still qualify for a means-tested benefit (Income Support) of equal amount. I therefore assume that contributory requirements are always satisfied. Moreover, even if recipients might do some type of work not exceeding the limits on weekly hours and earned income, I assume work is not allowed while receiving the benefit. The benefit amount is lower in the first 28 weeks and higher after having passed the 'suitable work test'. In the model, the decision period is one year; I therefore assume for simplicity that the annual

[6]In principle, the pension plan can be endogenous as individuals self-select into a particular type of job offering specific benefits (such as occupational pension funds). Given that I consider individuals at the end of their working life for whom private pension membership is mainly a predetermined characteristic, the type of plan is assumed to be exogenous.

benefit amount is fixed.

The last set of assumptions concerns the application process. I assume that (i) claiming for IB entails a time cost ϕ_{d_t} capturing the stigma of receiving disability benefit, the cost of filling out administrative forms and the cost of completing the health assessment; (ii) qualifying individuals definitely receive the benefit (i.e. the rejection rate is assumed to be zero)[7] and (iii) health is measured without error in the examination process. I assume perfectly observed health in the examination process due to data limitations that make it difficult to reasonably identify the error made by the examiner in assessing the applicants' health. One important limitation of this assumption is that it does not allow assessing the effectiveness of the screening process but only changes in individual behaviour under different health eligibility thresholds.

In addition to IB, among the non-contributory benefits I include Disability Living Allowance and Assistance Allowance. These benefits are aimed at covering additional costs due to personal care or mobility needs for individuals aged below or above 65, respectively. I assume that the benefit amount is the same for both benefits (in the model formulation I named these benefits dla). Qualifying individuals also receive means-tested benefits, such as Income Support, Pension Credit and Working Tax Credit. I assume that each entitled individual claims for the benefit. Further details on welfare benefits and their model implementation are reported in Appendix G.

3.2 State and private pensions

The state pension provision is of two different types, the Basic State Pension (BSP) that is received if individuals have payed National Insurance contributions for at least a quarter of their working life and Second Tier State Pension (STSP) that is related to earnings history. For simplicity, I assume that everyone is entitled to the full BSP. For STSP, at age 50 individuals start with an initial level of benefit entitlement[8]. The amount is then updated according to a function of earnings reported in Appendix C. I assume that each individual starts to receive the pension

[7]This captures the UK context quite well—at least before the ESA introduction—compared to the US where the award rate after two years from initial application is 0.53 and after 10 years is 0.67 according to French and Song (2014).

[8]In the UK system, workers can decide to contract out the contributions paid to STSP and to contribute instead to a private pension plan (often an occupational DB plan). This is not modelled. Individuals are endowed with an initial accrual in STSP and DC pension, and I assume they pay contributions to each plan according to their earnings.

benefit at SPA.[9]

In the UK, there are two main sources of private pension benefits—defined contribution (DC) and defined benefit (DB). A progressive shift from DB to DC has been observed in recent years; therefore, I assume that private pension wealth can only be in the form of DC pension funds. The amount in the fund depends on workers' (c_w) and employers' (c_e) contributions and on the rate of return (φ) of the fund. I assume the DC pension wealth evolves as follows: $q_{t+1}^{DC} = (1 + \varphi)(q_t^{DC} + (c_w + c_e)w_t h_t)$. Pension amount pb^{DC} depends on the accrued amount in the fund, the lump-sum amount the individual decides to receive when he first claims for the benefit and the annuity rate at the time of the annuitisation (r_{DC}). I assume that individuals purchase an annuity fixed in nominal terms, which is the most commonly bought. In addition, when the individual claims the benefit he gets a fraction ls of the fund as a tax-free lump sum and annuitises the rest $(1-ls)$, net of administrative costs l. The tax-free lump sum can reach up to 25% of the pension pot. In the model, I set minimum age to claim a private pension at 55[10].

The rate of return of the fund is assumed to be deterministic[11], the annuity rate does not vary with age, the lump-sum payment is a constant fraction of the amount in the fund and the amount to save in the retirement account is assumed to be a fixed fraction of annual earned income. The benefit at claiming age is given by $pb_t^{DC} = r_{DC}(1 - ls)q_t^{DC}(1 - l)$.

3.3 Dynamic programming problem and model solution

In what follows, I formalise the dynamic programming problem that individuals solve at each time period t. Let $V(X_t)$ be the value function at time t, with the vector of state variables defined as $X_t = (a_t, w_t, H_t, d_{t-1}, q_t^{SP})$ if $pen = NO$ and as $X_t = (a_t, w_t, H_t, d_{t-1}, p_{t-1}, q_t^{DC}, q_t^{SP})$ if $pen = DC$, where p_{t-1} is a dummy variable equal to one if in $t-1$ the individual has already annuitised the private pension fund and d_{t-1} takes the value of one if in $t-1$ the individual has received IB. I can write the value function as $V(X_t) = \max\{V^i(X_t)\}$, where the index i denotes the six possible discrete choice options: $i = 1$ if the individual is active in

[9]There are in fact few people observed claiming the benefit after SPA even if there are no penalties and modest incentives in terms of benefit amount in postponing benefit receipt.

[10]As a result of the Finance Act 2004, the minimum pension age from which DC fund can be annuitised has been increased from 50 to 55.

[11]The assumption of a deterministic rate of return implies that the model abstracts from an important source of uncertainty that is likely to affect individual retirement decision.

the labour market ($h_t > 0$), $i = 2$ if he is active and he claims for private pension ($h_t > 0$ and $p_t = 1$), $i = 3$ if the individual is neither active nor claiming a benefit ($h_t = 0$ and $p_t = 0$), $i = 4$ if he claims for IB ($d_t = 1$), $i = 5$ if he claims for private pension ($p_t = 1$) and $i = 6$ if he claims for both private pension and IB ($p_t = 1$ and $d_t = 1$). Depending on t and on the type of private plan, the set of choice variables differs. When $i = 1$:

$$V^1(X_t) = \max_{a_{t'}, h_t} \left\{ U(c_t, l_t) + \beta \pi^s_{t'} \iint_{\substack{H_{t'}, \\ w_{t'}}} V(X_{t'}|X_t) dF(X_{t'}|X_t) + \beta(1 - \pi^s_{t'}) b(a_{t'}) \right\}$$

with $l_t = L - h_t - \phi_H(\hat{H} - H_t) - \phi_{P_t}$ and budget constraint $a_{t'} = a_t + y(.;\tau) + ys_t + tr_t - c_t$. The budget constraint differs among pension types because DC plan holders are assumed to contribute a fixed fraction of their salary, c_w, to the fund. Therefore $y(.;\tau)$ equals $y(w_t h_t, ra_t, sp_t \mathbb{1}(age \geq \text{SPA}); \tau)$ if $pen = NO$ and $y((1 - c_w) w_t h_t, ra_t, sp_t \mathbb{1}(age \geq \text{SPA}); \tau)$ if $pen = DC$. With $i = 2$, there is no accrual in private pension, which means that c_w is zero and the amount of the private pension benefit is added to taxable income. Furthermore, when $i = 3, 4$ or 6, h_t is set to zero and the maximisation is only with respect to savings. From age 70 there is no uncertainty on future wages but only on future health, because I assume that individuals exit the labour market by age 70. For those having a private retirement account, the budget constraint at retirement age ($p_t = 1$ and $p_{t-1} = 0$) is slightly different from the one after retirement age. There is in fact the possibility of withdrawing up to 25% (lump sum - *ls*) of the amount in the account free of taxes, such that: $a_{t+1} = a_t + y(ra_t, pb_t, sb_t; \tau) + ys_t + ls * q^{DC}_t (1 - l) + tr_t - c_t$.

Individuals are heterogeneous with respect to state variables X_{it}. Wages and health status will differ across individuals given different realisation of wage and health shocks. However, given the same age, wage, health status, asset level, retirement decision and pension accrual, different individuals will make the same decisions. I denote the vector of preference parameters with ϑ, $\vartheta = \left\{\beta, \nu, \gamma, L, \phi_H, \phi_{P0}, \phi_{P1}, \phi_{d0}, \phi_{d1}, \phi_B, K, \bar{H}_d\right\}$ and the parameters that determine the data generating process for the state variables with χ:

$$\chi = \left\{ r, \omega_H(age_t), \sigma_{\nu_H}, \sigma_\eta, \rho_H, \omega_w(H_t, age_t), \sigma_{\nu_w}, \rho_w, \left\{ ys_t, \pi^s_t, pb^{DC}_t, sb_t \right\}^T_{t=1} \right\}.$$

The model is solved backward starting from period T and computing the solution in each period by assuming that agents form expectations about future realisations of the state variables according to the transition probabilities assumed by the model. The state variables are discretised into a finite number of points on a grid, and the value function is evaluated at each point of the state space. I take expectations with respect to shocks in health and wages and with respect to mortality risk.[12]

4 Data

I use data from ELSA, a biennial longitudinal survey, representative of English private household population aged 50 and over that started in 2002. ELSA contains detailed information on assets, both financial and property wealth; pension fund membership and accrued rights to private pensions; out-of work benefit receipt and earnings. It also contains detailed information on health status, both subjective and objective.

The information I need to compare model simulations with the data are labour market participation decisions, hours worked, pension and non-pension assets and health and IB rates for males living with a partner. Pension wealth and particularly accrual in state and private pensions are not directly reported by the respondents, who are only asked about the amount in the DC funds they are currently contributing to. To recover a comprehensive measure of private and state pension wealth, I use the pension wealth derived variables released together with raw data for each wave of ELSA. The derivation of pension wealth is consistent across waves starting from the second wave of data collection.[13]

4.1 Measuring health

The continuous measure of health that I use in the model is constructed using the rich set of health indicators available in ELSA. In particular, I follow Poterba *et al.* (2013) and apply principal component analysis to a set of dummy variables covering several dimensions of individuals' health.[14] To construct the index, I

[12]I integrate the value function with respect to the transitory component of health, η_t , using three-node Gauss–Hermite quadrature (see Judd, 1998). To capture uncertainty over the persistent components of health and wages, I convert θ_t and ϵ_t into discrete Markov chains, following the approach of Tauchen (1986).

[13]In the first wave, there is no distinction between DB and DC pensions. The derivation of accruals is described in Appendix C.

[14]Several continuous measures that use a set of indicators to recover 'true' latent health have been proposed, see in particular Meijer *et al.* (2011) , Jürges (2007) and Poterba *et al.* (2013). Kapteyn and Meijer (2014) and Venti (2014) discuss the main characteristics of these three health

select only objective indicators of health, even if self-reported, that should be less sensitive to measurement error (Crossley and Kennedy, 2002), heterogeneity in health perception (Lindeboom and van Doorslaer, 2004) and justification bias (Bound *et al.*, 1999) with respect to self-reported general health.[15] The selected health indicators replicate the set of health dimensions asked about in the Work Capacity Assessment (WCA), which is the Department of Work and Pensions's method of determining a person's ability to perform any type of work. The WCA is a measure of the extent to which a person is incapable of performing certain specified everyday activities laid down in legislation. These activities cover physical and sensory functions and mental function. The latter consists of performing four activities: daily living activities, completion of tasks, coping with pressure and interaction with other people. To account for mental health in the health index, I include the score obtained by the interviewees on the Center for Epidemiologic Studies Depression (CESD) scale.

The complete list of indicators used to construct the index is reported in Table A.1 in Appendix A, together with descriptive statistics for each indicator using pooled cross section data for males from wave 1 to 6. In addition to indicators capturing physical, sensory and mental functions (tested in the WCA), I add a set of variables relevant to better capturing the severity of limitations, such as limitations with Activities of Daily Living (ADL) and Instrumental ADL (IADL) (two dummy variables that take the value of one if at least one limitation is reported), any pain (one if the individual reports suffering pain) and if the individual is receiving help. Finally, given that in the model I am interested in capturing not only disability but a comprehensive measure of health, in the spirit of the Poterba *et al.* (2013) index, I add a set of indicators for diagnosed conditions.

The health index seems to capture a comprehensive measure of health well. As expected, it decreases with age and its distribution is remarkably different between males receiving and not receiving IB before SPA (see Figure A.1 in Appendix A).

---

indices. What seems to be important is the set of indicators used to construct the index more than the statistical technique implemented. The items' selection depends on the research question and on which aspect of health is of interest.

[15]However, these measures are not immune to biases, even if they are arguably more objective than self-reported general health. For example, van Ooijen *et al.* (2015) show that self-reported diagnosed conditions are under-reported when compared with administrative hospitalisation data, particularly for the mental health domain.

5 Model estimation

I estimate the model parameters in two steps, as in Gourinchas and Parker (2002), French (2005) and French and Jones (2011). First, I fix some parameters to values estimated in the literature and I estimate exogenous processes of health, wages and survival probability. Second, I estimate the remaining structural parameters using the Method of Simulated Moments.

5.1 First step: Estimation of exogenous processes

The estimation of the exogenous processes for health, wages and mortality is carried out using the first six waves of data from 2002 to 2012. The underlying assumption is that the introduction of ESA in 2008 does not effect health developments, wage offers and mortality risk.

Health process

The parameters of the health process to be estimated are the parameters of the deterministic component ($\omega_H(age)$), the variance of the persistent component ($\sigma^2_{\nu^H}$), the autoregressive coefficient (ρ_H) and the variance of the transitory component (σ^2_η). I first estimate the fixed effect regression in Equation 5.1 to obtain an estimate of the age parameters ($\hat{\pi}_i^H$), controlling for time effects and family size effects:

$$\log H_{it} = \sum_{j=1}^{3} \pi_j^H age_{it}^j + \sum_{k=1}^{K} \delta_k^H 1\left\{size_{it} = k\right\} + \mu^H U_t + \zeta_{it}^H, \tag{5.1}$$

where the error term is $\zeta_{it}^H = f_i + \theta_{it} + \eta_{it}$. I define the 'adjusted error term' as $g_{it}^H = \Delta\zeta_{it}^H = \Delta\theta_{it} + \Delta\eta_{it}$. The three parameters of interest are identified by the variance, lag one and lag two covariances of the adjusted error term (see Appendix D for details on moments' derivation) and are estimated using standard minimum distance techniques[16].

To estimate the process, I select males aged from 50 to 90 for whom the health index is non-missing, which means that the entire set of questions used to construct the index have to be non-missing. I end up with 22,088 individual-year observations for 6,587 distinct respondents.

In the first column of Table 5.1, the parameter estimates of the third order polynomial in age are reported. Health is decreasing with age, and the declin-

[16] See, for example, Low *et al.* (2010).

Table 5.1: Parameters of the deterministic component of health and wage processes. (Standard errors in parentheses.)

Process	Health (1)	Wage (2)
age	-0.518***	0.034**
	(0.126)	(0.012)
$age^2/100$	0.792***	-0.026**
	(0.184)	(0.010)
$age^3/100$	-0.004***	
	(0.001)	
health		0.026
		(0.024)
$health^2$		-0.002
		(0.004)
Observations	22,088	13,144

ing trend becomes steeper after age 70. The error component model specified for health has a persistent AR(1) component and a transitory component. The autoregressive parameter estimate is 0.977, suggesting high persistence of the process (see Table 5.2).

Table 5.2: Parameters of the stochastic component for health and wage processes. (Standard errors in parentheses.)

	Health process (1)	Wage process (2)
ρ	0.977	0.864
	(0.049)	(0.112)
σ_ν^2	0.167	0.012
	(0.025)	(0.004)
σ_η^2	0.372	0.023
	(0.027)	(0.004)

My estimates are close to recent estimates using similar specifications. For example, van Ooijen *et al.* (2015) propose and estimate a health measurement model in which the error component has a specification similar to the one I propose. They use self-reported health corrected by means of objective health measures collected in hospitalisation data as a health measure and find a high persistence process with an autoregressive parameter of 0.88. Using ELSA data, Blundell *et al.* (2016a) estimate a dynamic model of health and find that the sum of a transitory white noise process and a permanent AR(1) process is a good representation of health, with estimated values of the autoregressive parameter ranging from 0.90 to 1.06.

I include family size in the fixed effect regression, to control for potential changes in health status due to changes in the family structure and unemployment rate to control for time effects[17]. For unobserved heterogeneity captured by the fixed effect, I recover the fixed effects $\hat{f}_i$ and divide the distribution of $\hat{f}_i$ in two parts—below and above the first quartile, identifying low versus middle and high health levels at age 50, conditional on cohabiting with a partner. This heterogeneity in the fixed effect captures the part of health that is assumed to be predetermined at age 50 and that does not vary with age. I solve the model for these two different 'types' of individuals, meaning that this binary variable enters the state space.

The estimated process for health is able to replicate the observed evolution of health quite well[18]. Using the sample of males living with a partner and born between 1946 and 1955 as initial conditions, I simulate 20,000 histories of shocks for health and mortality risk and compute health at each age.

Wage process

The wage process is specified with a deterministic component that depends on age and health status and a stochastic term ζ_{it}^{w} that is the sum of an individual fixed effect (f_i), a persistent (ϵ_{it}) and a transitory components (ξ_{it}).

$$\log w_{it} = \sum_{j=1}^{2}(\pi_j^w age_{it}^j + \alpha_j^w H_{it}^j) + \sum_{k=1}^{K} \delta_k^w 1\left\{size_{it} = k\right\} + \mu^w U_t + \zeta_{it}^w. \quad (5.2)$$

I assume the ξ_{it} represents measurement error. The fixed effect estimation of Equation 5.2 provides estimates for age (π_j^w) and health (α_j^w) effects on productivity. Even if fixed-effects estimation allows getting rid of unobserved heterogeneity, there might be a selection bias if wage growth differs between workers and non-workers, given that only accepted wages are observed. To account for selection into participation, I estimate Equation 5.2 using both accepted (observed) and offered (unobserved) wages, where offered wages for those not observed working are imputed as described in Appendix D.

[17] When I simulate from the estimated processes for health and wages and from data profiles for decision variables presented in Section 5.2.2, I fix family size at two and set the unemployment rate at 4.9%, which is the 2004 annual unemployment rate for males in England (Source: Labour Force Survey. ILO unemployment rate.)

[18] In Appendix A, Figure A.2a reports the mean observed and simulated health, whereas Figure A.2b reports the observed and simulated fraction of individuals with health below the first quintile of the unconditional health distribution.

The second column of Table 5.1 reports parameter estimates for the wage process. Wages are increasing in health up to the first health quartile (3.5) when the relationship becomes almost flat. They are increasing in age up to age 67 and decreasing afterwards. The estimated autoregressive parameter can be compared with wage process estimates obtained using ELSA linked with National Insurance data in Crawford and O'Dea (2016). They estimate couple wage processes for three different education groups, and their autoregressive parameter ranges from 0.87 for low educated couples to 0.95 for high educated couples, where the education level is that of the male. The parameter value I obtain using only ELSA data is 0.864, at the lower bound of their range of estimations (see Table 5.2)[19]. When using these estimates in model simulation the fixed effect is set equal to the average fixed effect for individuals of the reference cohort (those born between 1946 and 1955).

Mortality risk

I assume that the probability at time t of dying by $t + 1$ is a function of age and health status in t. I first compute from the data the probability of dying by $t+1$ conditional on having a certain health level H in t, that is $Pr(death_{t+1}|H_t)$, controlling for cohort effects[20]. To do that, I discretise the health measure in four categories, below the 10th percentile ($i = 1$), between the 10th and the 20th percentiles ($i = 2$), between the 20th percentile and the median ($i = 3$) and above the median ($i = 4$). The unconditional probability of dying by $t + 1$, $Pr(death_{t+1})$, is then obtained as $Pr(death_{t+1}) = \sum_{i=1}^{4} Pr(death_{t+1}|H_t = i) * Pr(H_t = i)$, where the probability of health level i, $Pr(H_t = i)$, is computed in the data controlling for cohort effects, as for the conditional probabilities.

Mortality rates computed using ELSA data are lower than comparable mortality rates from the life tables (see Figure A.3 in the Appendix A). This might be due to non-random attrition and/or initial selection into participation; older and unhealthier individuals might be more likely to exit the panel and healthier individuals might be more likely to enter the panel. I assume that mortality risks perceived by the individuals are consistent with the life tables, and I correct mortality rates estimated from ELSA data by rescaling mortality in each health-age group in order to match the life tables' mortality rates (see Appendix D for

[19]Crawford and O'Dea (2016) do not account for selection; however, having earnings history data the problem of selection should be less severe. Moreover, wage persistence can be different at the end of working life, that is, after age 50.

[20]The strategy to control for cohort effects is the same as is explained in Section 5.2.2.

details).

5.2 Second step: Method of simulated moments

Among preference parameters ϑ, I fix relative risk aversion ν, discount factor β and time endowment L. I set relative risk aversion of the composite good consumption-leisure ν to 2. Holding labour supply fixed and assuming that consumption weight γ varies between 0.4 and 1 (French and Jones, 2011), the coefficient of relative risk aversion for consumption is given by $\gamma(\nu - 1) + 1$, which ranges from 1.4 to 2, in line with values estimated in the literature (Blundell *et al.*, 1994; Attanasio and Weber, 1995; Banks *et al.*, 2001). The discount factor β is set to 0.9756 as in Low and Pistaferri (2015), who use the central values of estimates from Gourinchas and Parker (2002) and Cagetti (2003). Time endowment L is set to 4,466 hours as in French (2005).

The remaining parameters are estimated using the Method of Simulated Moments to minimise a weighted distance (Generalised Method of Moments (GMM) function) between simulated life-cycle profiles and data life-cycle profiles.Table 5.3 reports the complete list of matched moments. I include the fraction receiving IB and the average health status for those receiving IB from age 50 to age 64, which leads to 15×2 moment conditions. Moreover, I match the fraction receiving IB in t given that they were receiving IB in $t - 2$ (due to the biannual nature of the data set used) from age 52 to age 64, which leads to 13 moment conditions. These moments, together with labour supply conditional on health, identify the health threshold ($\bar{H}_d$), that is, the level of health below which individuals can claim for IB, and the stigma cost (ϕ_{d_t}), which varies with age. In particular, the stigma cost is identified by the fraction of individuals not claiming IB even if they are in very bad health.

Furthermore, I match labour supply participation from age 50 to age 69 (assuming that at age 70 everyone is retired) conditional on four health intervals, leading to 20×4 moment conditions, and annual hours worked from age 50 to age 69, which generates 20 additional moments[21]. Individuals in worse health work less than individuals in good health, and this heterogeneity in labour supply by health is used to identify the time cost of being in bad health (ϕ_H). As in previ-

[21] In the data, hours worked refer to the usual weekly hours in the current job, and no information is collected on past jobs for those currently out of work. I derive annual hours by assuming that individuals have worked the entire year, and when information was available I correct this measure to account for unemployment or sickness periods. However, the resulting measure shows low variability by health level. I therefore include only unconditional moments for hours worked.

ous studies, hours of work are used to pin down the fixed cost of work (ϕ_{P0}), and the decrease in hours worked with age (together with the decrease in participation with age) helps identify the slope of the fixed cost of work (ϕ_{P1}).

Finally, I include in the set of moments mean assets and three asset quantiles, first tertile (π_1), median (π_2) and second tertile (π_3) from age 50 to age 70, resulting in 21×4 moments. Asset profiles, together with labour supply profiles, identify the consumption weight parameter (γ). If death were not stochastic, assets at the time of death would identify the bequest function parameters (ϕ_B and K). When time of death is uncertain, people save both for precautionary and bequest motives; therefore, the identification of these parameters is weak but guaranteed because risk aversion has been fixed. To derive initial conditions for

Table 5.3: List of matched moments.

Moment	age span	number of moments
Fraction receiving IB	50–64	15
Average health when in IB	50–64	15
Persistence in IB receipt	52–64	13
Labour supply participation by health	50–69	20×4
Hours worked	50–69	20
Mean assets	50–70	21
Median assets	50–70	21
First and second assets tertiles	50–70	21×2
Total		227

model simulation, I select the cohort of males living with a partner born between 1946 and 1955, excluding self-employed individuals, and assume that marriage status does not change over time. I consider individuals interviewed in wave 2 and individuals entering the survey in wave 3.[22] The final sample used to estimate the joint initial distribution of assets, accrual, health and wages consists of 657 individuals. Each of the simulated individuals receives a draw of assets, accrual, health and wages from the estimated initial distribution.

The procedure to estimate the parameter vector ϑ involves the following steps. First, the data generating processes for health, wages and mortality are estimated as described in Section 5.1. Second, the data moments to be matched are computed by estimating life-cycle profiles accounting for cohort and health effects, as explained in Section 5.2.2. Third, initial conditions for model simulation are derived from the joint distribution of state variables X_{ij}, and histories of wage,

[22] I do not consider wave 1 because the derivation of pension wealth variables is standardised only starting from wave 2.

health and mortality shocks are generated from the estimated data generating processes. Fourth, given an initial arbitrary parameter vector ϑ_n, the model is solved and decision rules are computed. Fifth, using initial conditions and histories of shocks from the second step, I simulate S life-cycle profiles for the decision variables. Sixth, I compute the difference between data and simulated profiles. The moment conditions are stacked in a K-elements ($K = 227$) vector denoted by $\theta(\vartheta, \chi_0)$. Seventh, the moment conditions are weighted up according to a GMM objective function (Equation 5.3). A new parameter vector is selected, and the process is repeated from step four to step seven up to convergence of the GMM objective function to the minimum:

$$\arg\min_{\vartheta} \frac{I}{1+\tau}\hat{\theta}(\vartheta, \chi_0)'\hat{W}_I\hat{\theta}(\vartheta, \chi_0), \tag{5.3}$$

where I is the number of observations, τ is the ratio of I to the number of simulations S, $\hat{\theta}(\cdot)$ is the sample analogue of $\theta(\cdot)$ and $\hat{W}_I$ is the $K \times K$ weighting matrix. Details regarding the choice of the weighting matrix and the distribution of $\hat{\vartheta}$ are presented in Appendix E.

5.2.1 Calibration of other parameters

I set the amount for IB and DLA equal to the average amount received by males aged between 50 and 69 in 2004, that is £3,460 for IB and £3,000 for DLA[23]. The parameters of DC pension are calibrated as follows: contribution rates (c_w, c_e) are set at 6%; the rate of return of the fund (φ) is assumed deterministic and set at 7%[24]; the annuity rate r_{DC} is set at 4%, in line with rates reported for males aged 65 in the UK compulsory market by Cannon and Tonks (2011); administrative costs are set at 10%, as in Crawford and O'Dea (2016); and the lump-sum payment ls is assumed equal to 15% of the pension pot. The rate of return on the safe asset r is set at 0.029, the average real return on UK Government liability in 2002–2008 (Barclays, 2013).

5.2.2 Estimation of data life-cycle profiles

The life-cycle profiles for assets, participation, hours worked, IB claimants and IB persistence are estimated accounting for cohort and health effects. The procedure is similar to the one implemented in French (2005). To increase sample size, I

[23]Source: amount data from Department of Work and Pensions tabulation tool.

[24]Crawford and O'Dea (2016) compute mean and standard deviation of DC fund returns between 1994 and 2010. The mean is 3.97% and the standard deviation is 13.8%.

use information of both singles and couples. Taking as an example hours profile, I regress log hours, $log(h_{it})$, on an individual specific effect f_i, age dummies, a full set of family size dummies $size_{it}$, a dummy for having a cohabiting partner $couple_{it}$ and unemployment rate U_t, proxying for aggregate time effects. When considering labour supply participation as the outcome, age dummies are interacted with health. This specification allows estimation of age parameters (and in case of participation age parameters conditional on a certain level of health), accounting for individual fixed effect, time effect and family size effect.

I derive an estimate of f_i and regress the $\hat{f}_i$ on a set of ten-year cohort dummies[25] and the couple dummy to get the conditional expectation of f_i for males of a specific cohort cohabiting with a partner, $E\left[f_i|cohort = c, couple = 1\right]$. When simulating the data profile I replace the individual effect f_i with $\tilde{f}_i = f_i - E\left[f_i|cohort_i, couple_i\right] + E\left[f_i|cohort = c, couple = 1\right]$. The reference cohort c comprises those born between 1946 and 1955. This results in data profiles that are representative of the same group of individuals used to set up initial conditions for model simulations.

To estimate the decision profiles, I consider only the first four waves of ELSA, that is data from 2002 to 2008, a period in which DI policies and parameters were relatively stable. The DI programme within the model reproduces the main features of IB. In particular, the health eligibility rule to receive DI is estimated to match the fraction of IB recipients and thus represents DI eligibility conditions before the introduction of ESA.

Assets profiles

Assets are defined at the couple level and include net financial and housing wealth. In the period considered, that is between 2002 and 2008, house prices have grown rapidly. Blundell *et al.* (2016b) report real house price movements in England from 2002 to 2013. Between the first two waves of ELSA, in 2002 and in 2004 respectively, house prices increase by 40%. A life-cycle profile using an asset measure that combines housing and non-housing wealth might be largely influenced by this positive shock in illiquid wealth. Moreover, if the assets profile used to calibrate the model is not corrected for house price changes, then the assets increase observed in the data would be explained by savings in pension and non-pension wealth.

I assume that the house price increase and the resulting wealth increase for

[25] In the case of the participation equation, among the regressors I add health categories.

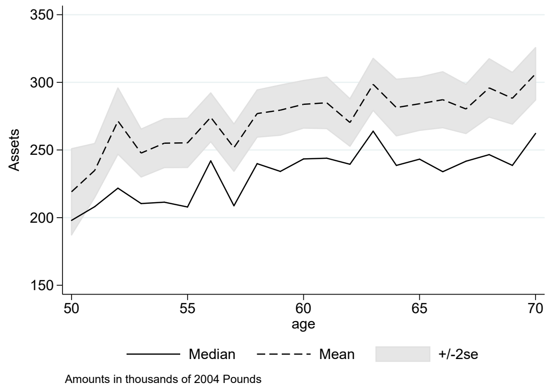


Figure 5.1: Mean and median asset profiles.

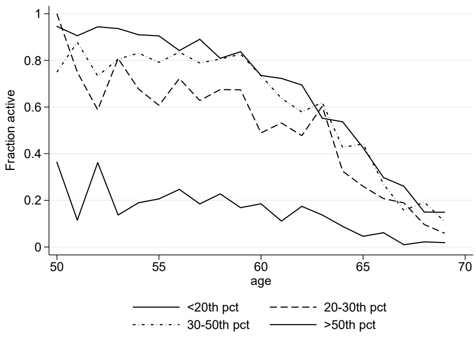
homeowners do not affect individual decisions in terms of consumption, retirement and labour supply. Therefore, I strip out house price changes by dividing net primary housing wealth by the house price index, using as reference year 2004, and I assume a price increase equal to the real rate of return on other financial assets. The corrected net primary housing wealth is added up to net non-housing wealth and used to estimate the asset life-cycle profile corrected for cohort effects.

In estimating the assets profile, I use 9,347 individual-year observations for 4,225 males aged from 50 to 70, interviewed in waves from 1 to 4. The resulting assets profile is reported in Figure 5.1. Assets are slightly increasing from age 50 to 60 and are almost flat after age 60.

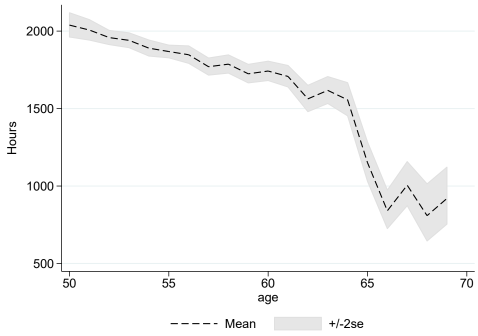
Labour market participation profiles

Figure 5.2a shows participation in the labour market conditional on health. The estimated profiles are obtained using 11,860 individual-year observations for 5,124 males between age 50 and 75observed in the first four waves of ELSA. For each age, the graph shows participation rates among individuals with health below the 20th quantile, between the 20th and the 30th quantile, between the 30th quantile and the median and above the median. Health quantiles refer to the unconditional health distribution for individuals aged 50 to 90 in the data. As expected, participation decreases with age and is lower for individuals having a lower level of health.

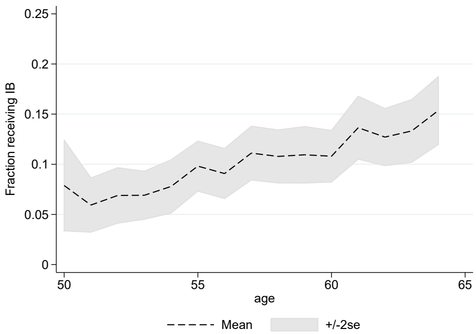
The aggregate hours profile for those active in the labour market is reported in Figure 5.2b. To estimate the profile, I use 4,904 individual-year observations for 2,388 males aged between 50 and 75, interviewed in waves form 1 to 4. Mean hours is almost flat at about 2,000 hours up to age 65 when it drops to 1,000 hours.



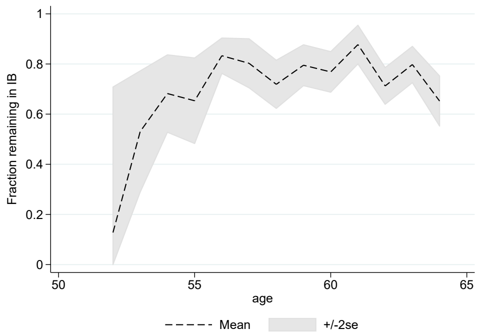
(a) Labour market participation



(b) Hours



(c) IB participation



(d) IB persistence

Figure 5.2: Estimated data life-cycle profiles.

DI participation profiles

Figure 5.2c shows the fraction of those claiming for IB. The estimation is performed using 6,853 individual-year observations for 3,243 males aged between 50 and 64. The profile is quite noisy up to age 52; it is increasing with age and for the cohort considered it reaches about 15% for those aged 64. After age 64, having reached SPA, it is not possible to claim IB anymore. Finally, Figure 5.2d shows the fraction of those receiving IB in t given that they were receiving IB in $t-2$. The IB receipt is highly persistent, at about 80%. There is some evidence of an increasing persistence with age; however, the profile is noisily estimated for those younger than 55. In the estimation, 3,511 individual-year observations for 1,828 males participating in the first four waves of ELSA have been used.

5.2.3 Estimation results

Table 5.4 reports the preference parameter estimates. The estimated health threshold to receive IB corresponds to the 20th percentile of the unconditional health distribution for individuals aged 50 to 90 in the data. The fixed cost of

Table 5.4: Structural parameter estimates.

Parameter	Description	Value	SE
γ	consumption weight	0.552	(0.002)
ϕ_H	cost of being in poor health (hours)	732	(3)
ϕ_{P0}	fixed cost of work at age 50 (hours)	698	(5)
ϕ_{P1}	age trend of the fixed cost of work (hours)	42	(0.3)
ϕ_B	bequest weight	5.66	(0.04)
K	bequest function curvature (£)	62,030	(507)
$\bar{H}_d$	eligibility threshold for IB (health level)	3.35	(0.01)
ϕ_{d0}	stigma cost of receiving IB (hours)	38	(0.5)
ϕ_{d1}	age trend of the stigma cost (hours)	37	(0.3)

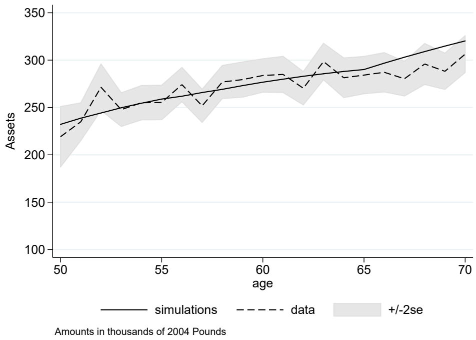
work is 698 hours at age 50 and 1,496 hours at age 69, similar to the estimated value in the baseline specification of French (2005), which is 1,315 hours at age 69. The time cost of being in poor health is high; for individuals with health at the eligibility threshold for receiving IB the cost of sickness is 1,122 hours, that is 25% of the time endowment. Holding labour supply fixed, the implied coefficient of relative risk aversion for consumption is 1.55. The marginal propensity to bequeath out of an extra pound is 0.45, and the bequest motive becomes operative at 49,000 pounds.

6 Model fit

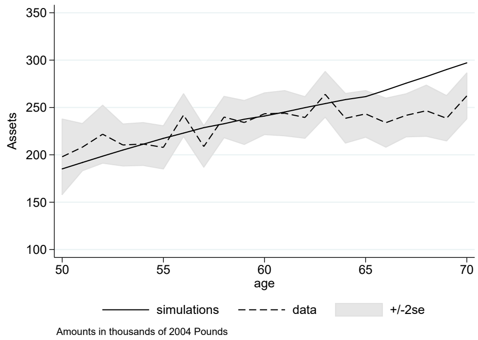
The estimated model replicates the main facts observed in the data quite well, such as age profiles for assets, participation, participation by health, IB rate and IB persistence, for both matched moments and moments that are not directly targeted in the estimation procedure.

Figure 6.1 reports simulated versus data profiles for assets. Mean and median assets are almost always within the confidence intervals but are slightly overestimated above age 65. The first tertile is overestimated, particularly after SPA, whereas the simulated second tertile replicates the data profile well (see Figure A.4a in Appendix A). Additional assets moments not included in the GMM criterion function but well replicated by the model are the first and second tertiles of assets conditional on health (above and below the median). Simulated and data profiles for conditional assets moments are shown in Figure A.5 in Appendix A.

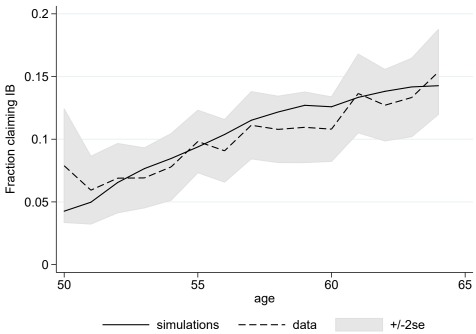
The fraction of individuals receiving IB in the simulation is very close to that in the data (see Figure 6.1c). In Figure 6.1d, I report the estimated probability of receiving the benefit in t conditional on having received the benefit in $t-2$ in the data and in the simulations. The model is able to replicate the persistence ob-



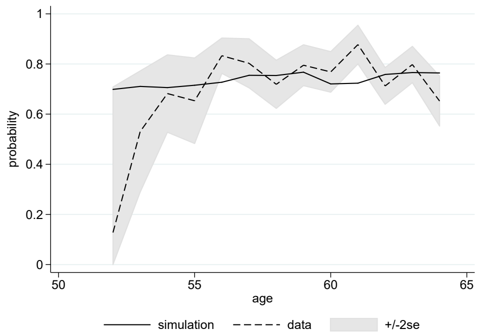
(a) Mean assets



(b) Median assets



(c) Incapacity Benefit



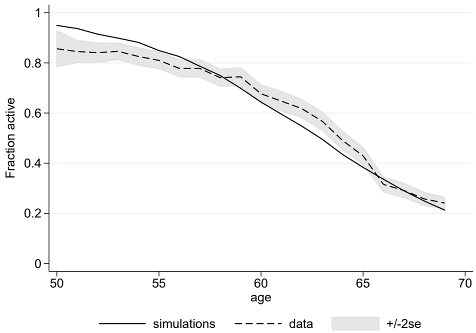
(d) $Pr(d_t = 1 | d_{t-2} = 1)$

Figure 6.1: Life-cycle profiles. Simulations versus data.

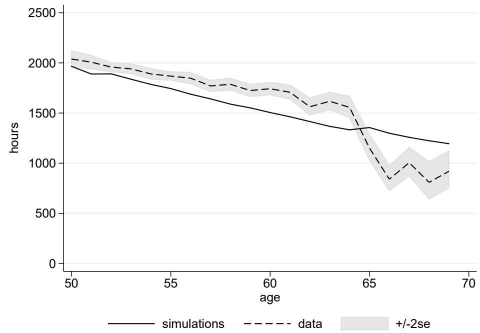
served in the data quite well. The increase in persistence with age is less evident in the simulations, but the simulated profile is generally within the confidence interval. The average level of health of individuals claiming IB is reported in Figure A.4b in Appendix A; the simulated profile lies within the confidence interval.

In Figure 6.2, I report the participation and hours worked profiles. Aggregate participation is slightly higher than that observed in the data for those younger than 60 (Figure 6.2a). This is also evident in Figure 6.2 (bottom panel) in which participation is disaggregated by health levels. The model captures differences in participation for individuals with different levels of health quite well. The model replicates the profile for hours worked less precisely (Figure 6.2b); the average number of hours worked declines linearly with age from 2,000 hours at age 50 to 1,200 at age 69, whereas in the data at age 65 it sharply drops at 1,000 hours per year. Financial incentives provided by state and private pensions are not enough to explain the sharp drop in hours worked at SPA, perhaps due to non-pecuniary motives not captured by the model.

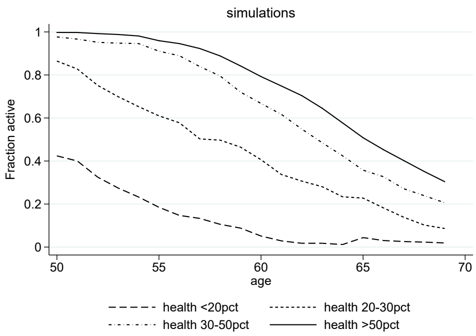
Finally, in the model individuals are endowed with a private pension fund



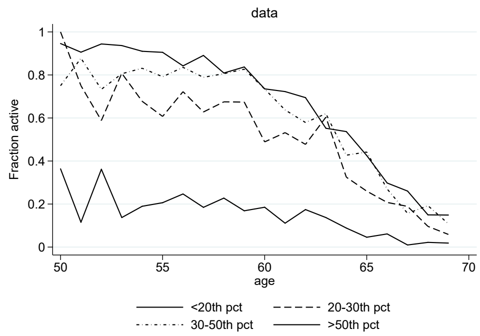
(a) Participation



(b) Mean hours



(c) Participation by health



(d) Participation by health

Figure 6.2: Life-cycle profiles. Simulations versus data.

and decide when to annuitise the amount in the fund. While in the model only accrual in DC plans is considered, in the data accrual can be both in DC and DB funds. In the initial distribution of the state variables, DB accrual is converted into DC accrual[26]. According to Figure A.6 in Appendix A, the model generates a distribution of claiming age close to that observed in the data when merging DC and DB fund holders, even if the private pension claiming age is not in the set of matched moments.

7 Policy experiments

The goal of this section is to understand how changes in the structure of the DI programme that replicate recent and proposed reforms affect labour supply decisions, DI benefit claiming decisions and welfare. First, I simulate individual responses to increasing or decreasing the health eligibility threshold to receive the benefit (Section 7.1). An extreme case of this experiment is the complete elimination of IB, which allows measuring the willingness to pay to have IB in the

[26]Details are reported in Appendix C.

tax and benefit system. Second, I look at behavioural responses to changes in the amount of IB (Section 7.2). This allows computing non-participation and IB rate elasticities to benefit generosity and comparing them with estimates from other institutional contexts. Third, I investigate the effectiveness of a policy intervention promoting the labour market participation of workers with health conditions limiting their working capacity (Section 7.3).

7.1 Varying the health eligibility rule

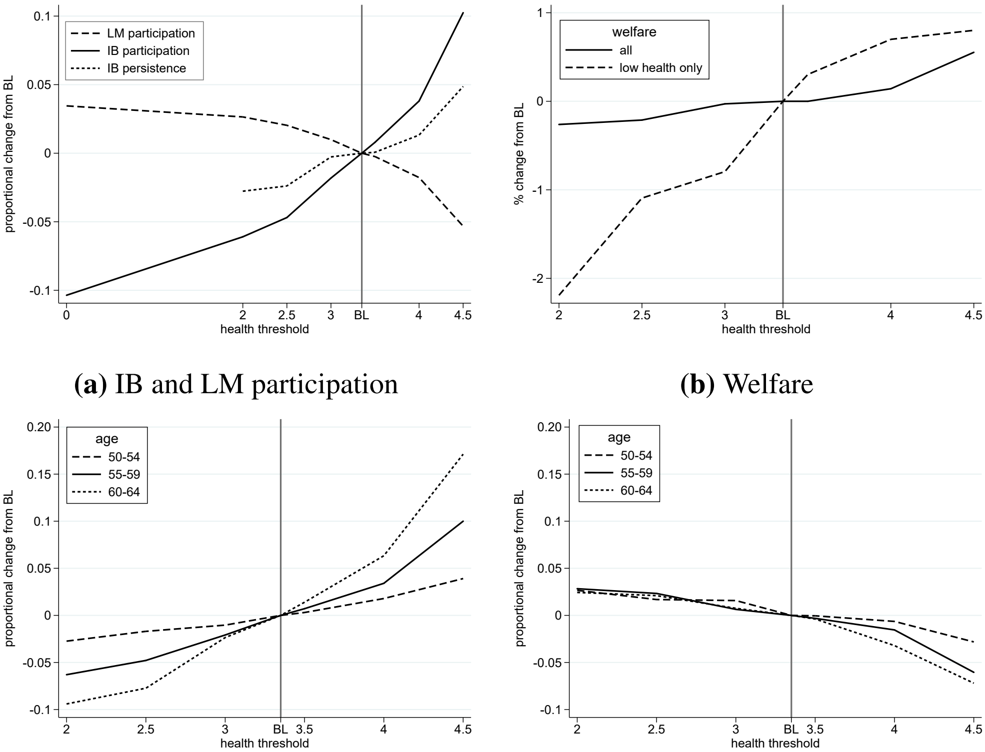
In the model, the health eligibility threshold to receive IB is represented by the parameter $\bar{H}_d$, estimated to be equal to 3.35. To assess the effects of changes in the stringency of the health assessment, I simulate the model outcomes under several health thresholds. Figure 7.1a shows that a significant fraction of individuals losing the benefit when the requirement becomes more stringent are induced to enter the labour market. For example, reducing the threshold from 3.35 to 2.5 (corresponding to a 5% reduction in eligible individuals) reduces IB participation by 5 percentage points but increases labour market participation by 2 percentage points. Persistence in benefit receipt is reduced when the threshold is made more stringent, but the effect is small. This is not surprising because of the absence of a health reassessment after entering the benefit; reducing the health threshold reduces benefit inflows and improves targeting but affects persistence less.

Figures 7.1c and 7.1d explore heterogeneity by age in IB and labour market participation, respectively. The reduction in benefit inflow due to a more stringent health assessment is higher for older males (age 60–64), but the reduction in benefit participation does not translate into an equally remarkable increase in labour market participation for this group. When the health threshold is increased, IB (LM) participation increases (decreases) with age, meaning that a more generous scheme generates a greater disincentive to work for older people.

Table 7.1: Median welfare compensation.

	Total assets at age 50 (quartiles)			
health at age 50 (quartiles)	1st < £122,000	2nd -£213,000	3rd -£350,000	4th -£350,000
1st	0.36	0.06	0.04	0.05
2nd	0.04	0.01	0.02	0.01
3rd	0.03	0.01	0.01	<0.01
4th	0.01	<0.01	<0.01	<0.01

To evaluate the welfare effect of altering the DI programme, I calculate the



(a) IB and LM participation

(b) Welfare

(c) IB participation

(d) Labour market participation

Figure 7.1: Varying the health eligibility rule. (BL=baseline).

amount of assets that an agent needs to receive at age 50 to be indifferent between the reformed and baseline scenarios. To make amounts comparable across asset levels, the compensating variation is expressed as a fraction of total liquid assets at age 50.[27] In Figure 7.1b, the welfare effect is reported averaging over the entire sample (solid line) and for those with a low initial level of health (dashed line)[28]. A more stringent threshold has a small effect on welfare in aggregate terms, but when considering those who lose much in terms of insurance, that is, those with a higher probability of being in poor health, the effect is significantly larger.

The extreme case of complete benefit removal computed by setting the threshold

[27]I assume that revenue costs, whether positive or negative, resulting from policy changes are offset by fiscal policies that affect only individuals younger than 50, which are not considered by the model. The model setup allows for the evaluation of revenue-neutral policies, but I chose to ignore revenue considerations in the counterfactuals beacuse policy reforms are hardly ever funded only by (or savings are hardly ever distributed only to) the policy beneficiaries. Moreover, DI accounts for less than 2% of GDP in the UK; therefore, even an extreme reform has small effects on general taxation.

[28]The low-health individuals are defined as those with an individual effect (f_i) in the first quartile of the fixed effects distribution, obtained from the estimation of the fixed effect health regression (see Section 5.1).

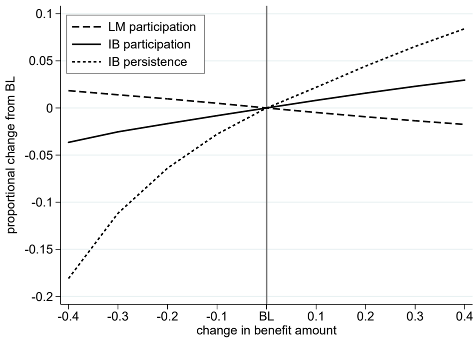
at zero, not included in Figure 7.1b, is considered in detail in Table 7.1. The welfare effect when the benefit is removed is a measure of the value of IB. The removal of IB is a loss for everyone in expected terms but has important distributional consequences that I illustrate in reporting the results for different combinations of health and wealth levels. The fractions of total assets reported in Table 7.1 suggest the presence of a large heterogeneity in the value of IB for individuals having different levels of health and wealth at age 50. On average, individuals with fewer assets and poorer health place a higher value on DI. In particular, individuals with assets and health in the first quartile of the distribution at age 50 would need 36% of their initial assets (a median value of about £6,000) to be compensated for benefit removal, a fraction significantly higher with respect to that needed by individuals with the same wealth level but with health in the second quartile (4%). These results highlight the importance of targeting the DI benefit to truly needy individuals.

7.2 Varying the benefit amount

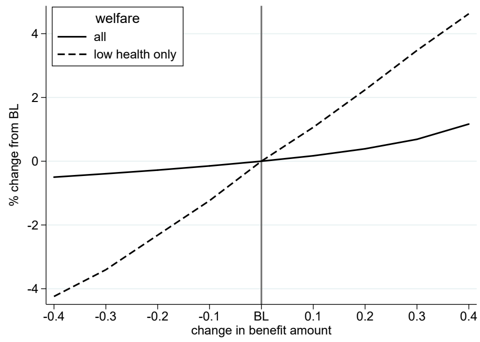
In this policy experiment, I compute the effects of varying the benefit amount from a 40% reduction to a 40% increase relative to baseline on labour supply decisions, IB receipt and welfare. To facilitate the comparison with existing literature regarding the effect of DI on labour supply, I compute the elasticity of non-participation in the labour market and the elasticity of IB participation to benefit generosity.

As expected, Figure 7.2a shows that IB participation is positively correlated with benefit generosity and negatively correlated with labour market participation. The effect of changing the benefit amount on IB persistence is particularly remarkable. When the benefit is made less generous a relatively large fraction of agents decide to exit the programme. This effect is stronger for those in relatively good health (not reported in the figure).

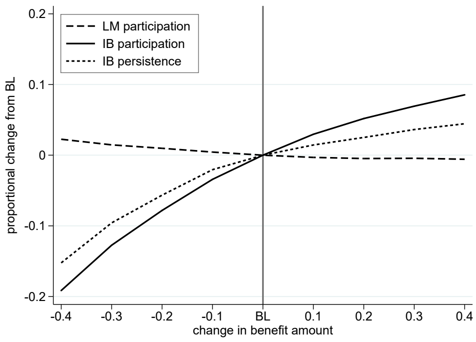
Figure 7.2b reports how welfare changes from baseline when the benefit amount changed. Welfare is increasing in benefit generosity, but the size of the effect is bigger for those with low health at 50. Focusing on health heterogeneous effects, Figures 7.2c and 7.2d replicate Figure 7.2a for those in very poor health and for those with health close to the eligibility threshold, respectively. The change in IB participation is smaller for those in worse health, but the most remarkable difference is in labour market participation. Those in very poor health do not change their labour supply behaviour when the benefit amount is made more or less gen-



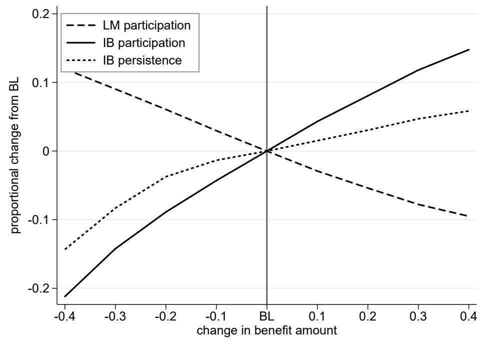
(a) IB and LM participation



(b) Welfare



(c) Low health



(d) Health close to the threshold

Figure 7.2: Varying the benefit amount.

erous, whereas those just below the eligibility threshold adjust their labour supply behaviour in response to changes in benefit generosity.

To compare model predictions with results from the existing literature, I compute the elasticity of non-participation rate to benefit generosity. The resulting elasticity is 0.20 at the lower bound of the estimates surveyed by Bound and Burkhauser (1999) using data from the US $(0.2 - 1)$. Two reasons might explain the lower elasticity found for the UK. First, the elasticities I compute refer to males aged between 50 and 64. If considering only individuals aged 50 to 54 the elasticity of non-participation to the labour market is about 0.55, which is a value closer to earlier estimates and is consistent with previous findings, suggesting that elasticities are decreasing as age increases[29]. Second, the peculiar aspect of IB of providing a low flat rate benefit amount results in mainly low-skilled workers

[29]From Table 13 in Bound and Burkhauser (1999), the elasticities estimated for males between 45 and 59 are far lower than those estimated for males aged below 50. Mullen and Staubli (2015) compute the elasticities for workers aged between 35 and 59 and find that the elasticity of the DI rate with respect to benefit generosity is highest for workers between 45 and 49, and it decreases for workers above age 50.

with lower employment opportunities entering the benefit. Therefore, given the characteristics of the target population one might expect the demand for IB to be rather inelastic to marginal changes in the benefit amount.[30]

Finally, I compute the elasticity of IB participation rate (0.79) and IB inflow rate (0.08) to benefit generosity.[31] Confirming the results shown in Figures 7.2c and 7.2d, elasticities are heterogeneous by health level; agents with health below the 10th percentile show an IB participation elasticity of 0.47, and those with health close to the eligibility threshold (between the 10th and the 20th percentiles) show an IB participation elasticity of 0.65; as expected, IB inflow elasticity is higher for those close to the eligibility threshold (0.8). Regarding labour force non-participation, those at the bottom and at the top of the health distribution show an elasticity close to zero (about 0.05). The more responsive individuals are close to the eligibility threshold, with an elasticity of 0.45.

7.3 Reducing the fixed cost of work when at risk of IB entry

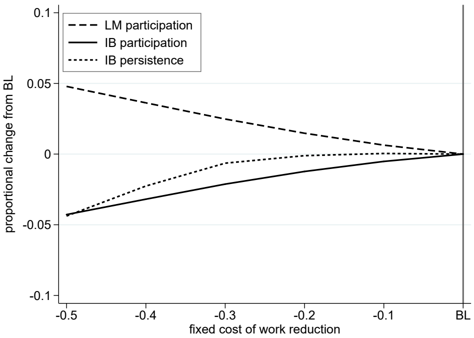
Recent reforms have also introduced measures to promote labour market inclusion of people with disabilities. To simulate the effect of these measures, I assume that the mechanism through which this policy affects individuals is by reducing their fixed cost of work, which in the model is specified as a time cost linearly increasing with age (ϕ_{P_t}). Specifically, I assume that if an individual has health below the IB eligibility threshold, his fixed cost of work is proportionally reduced by 10 to 50% in the reformed scenario.[32]

Reducing the fixed cost of work for those at risk of IB entry increases participation and reduces IB receipt almost linearly and with the same slopes (Figure 7.3a). IB persistence is unaffected up to a 30% reduction in the fixed cost and starts to decrease sharply afterward. On average, welfare slightly increases; however, the effect is much higher when focusing on individuals most at risk of IB entry (Figure 7.3b).

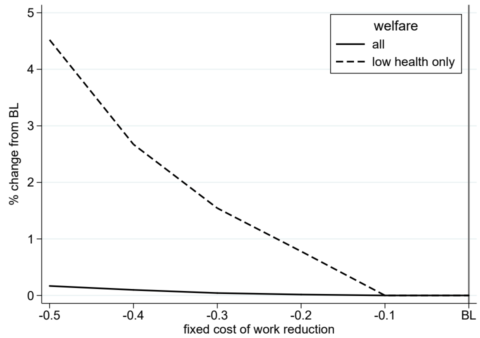
[30] This is consistent with the findings of Mullen and Staubli (2015), according to whom the elasticity of DI participation rate to benefit generosity in Austria is lower for low-skilled and poorer workers.

[31] Clear differences in the DI application process and awards in the US and the UK make a comparison between elasticities of DI application and awards to benefit generosity computed for the US (Bound and Burkhauser, 1999) and the DI participation and inflow rate elasticities computed for the UK uninformative.

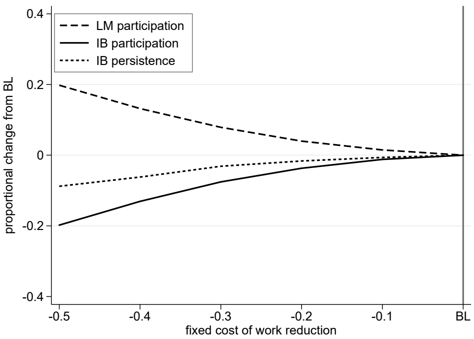
[32] This means that the fixed cost reduction does not apply only to those receiving IB and with health below the eligibility threshold but applies to all individuals that are at risk of entering IB. Moreover, the proportional reduction implies that the absolute reduction in the cost of work is increasing with age.



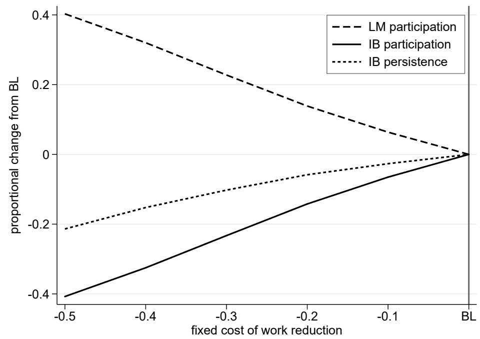
(a) IB and LM participation



(b) welfare



(c) Low health



(d) Health close to the threshold

Figure 7.3: Reducing the cost of work.

In Figures 7.3c and 7.3d, the same statistics in Figure 7.3a are reported but for those in very poor health and those with health close to the eligibility threshold, respectively. The reduction in the fixed cost of work is 'beneficial' for both groups but has a higher impact in terms of IB reduction and labour market participation increase for the latter. Among the policy interventions considered in this section, this reform scenario is the only one that increases labour supply and reduces IB participation without reducing welfare. This feature makes it particularly appealing and interesting.

Adam *et al.* (2010) investigate the effect of the Pathways to Work pilot implementation in 2003, a set of measures providing both financial and non-financial support to IB claimants to increase their chances of returning to work. Using a difference-in-differences approach, they quantify a 5.8 percentage point increase in employment of those exposed to the programme with respect to the control group. [33] It is not a priori clear how policies enhancing labour market inclusion

[33] The programme, with some changes with respect to its initial design, was implemented within the ESA in 2008; those placed in the WRA group are mandated to the Pathways to Work.

quantitatively affect the fixed cost of work, but under the assumption of a direct effect of policy interventions aimed at improving the work readiness of disabled individuals on the fixed cost of work ϕ_{P_t}, I can use the estimates of Adam *et al.* (2010) to give a sense of the ϕ_{P_t} reduction caused by such an intervention.

In the model, all those potentially eligible for IB (with health below the eligibility threshold) are exposed to the intervention, that is, they face a reduction in ϕ_{P_t}. To make the sample comparable to the one of Adam *et al.* (2010), I consider only those receiving IB in the baseline scenario and find that a 10% reduction in ϕ_{P_t} increases labour supply participation by 5.3 percentage points, a value comparable in magnitude to the effect of Pathways to Work estimated by Adam *et al.* (2010). If this is the order of magnitude of current interventions, a 10% reduction in ϕ_{P_t} has no effect on welfare and has mainly behavioural effects on those close to the eligibility threshold (see Figure 7.3d), supporting the ESA policy design that implements the Pathways to Work programme only for the WRA group.[34]

8 Conclusions

In this paper, I develop and estimate a life-cycle model of labour supply, DI claiming and saving behaviour for the UK. I model the decisions of males living with a partner and approaching retirement age facing uncertainty regarding wage realisation, health and life expectancy. The model is able to replicate quite well asset profiles, labour market participation and its heterogeneity by health level, as well as the fraction receiving DI by age and DI persistence over time. Health is measured on a continuous scale and is based on a large set of objective indicators collected in ELSA covering the health domains measured in the health assessment to receive DI. Both the mean and the distribution of health evolution over time are well reproduced by the specified process for health.

Having a continuous measure of health allows the investigation of heterogeneous effects of alternative policy interventions for individuals with different levels of health. I document that individuals with lower assets and lower health place a higher value on DI. As a consequence, on average, strengthening the health assessment to receive DI or reducing the benefit amount has small negative effects on welfare but significantly larger effects for those with a higher probability of experiencing a decrease in health.

According to model simulations, individuals in relatively better health are

[34]Additional evidence on the effects of the ESA introduction predicted by the model are discussed in Appendix F.

more responsive to changes in DI. Moreover, the model allows quantifying the effects of several policy interventions. For example, I show that of the individuals leaving the benefit due to a reduction in the benefit amount only half return to work. When focusing on those in relatively better health a 10% decrease in benefit amount results in a 4 percentage point decrease in DI participation and, among those exiting the benefit, 7 out of 10 return to work. Finally, a policy intervention that reduces the fixed cost of work by 10% (for example by improving work readiness and reducing the search costs) increases labour market participation by about 6 percentage points for those receiving DI in the baseline scenario without decreasing welfare.

The results in this paper are limited to cohabiting males aged above 50. Couples tend to be wealthier, less at risk of under-saving, in better health and have lower mortality rates than singles. This suggests that the results could not be easily extended to singles. In the model, I abstract from couple joint decisions; however, it is important to note that the presence of the partner and her economic status might influence the couple's decision to participate in the labour market and to claim DI (Blau, 1998; Blau and Gilleskie, 2006; Borella *et al.*, 2017).

The model assumes that health is perfectly observed, and this prevents the evaluation of changes in the effectiveness of the screening process. To include this aspect in the model, one would need data on the application process, the examiner's evaluation and the final decision. The model also assumes that health is exogenous. It is possible that changes in DI benefit induce individuals to modify their behaviours, thus affecting their health. The investigation of heterogeneous effects by household type and the potential endogenous role of health are left to future research.

References

Adam, S., Bozio, A. and Emmerson, C. (2010) Reforming disability insurance in the UK: Evaluation of the Pathways to Work Programme, (mimeo).

Attanasio, O., Low, H. and Sánchez-Marcos, V. (2005) Female labor supply as insurance against idiosyncratic risk, *Journal of the European Economic Association*, **3**, 755–764.

Attanasio, O. P. and Weber, G. (1995) Is consumption growth consistent with intertemporal optimization? Evidence from the consumer expenditure survey, *Journal of Political Economy*, **103**, 1121–1157.

Ball, S. and Low, H. (2014) Do self-insurance and disability insurance prevent consumption loss on disability?, *Economics*, **81**, 468–490.

Banks, J., Blundell, R. and Brugiavini, A. (2001) Risk pooling, precautionary saving and consumption growth, *Review of Economic Studies*, **68**, 757–779.

Banks, J., Blundell, R. and Emmerson, C. (2005a) The balance between Defined Benefit, Defined Contribution and State Pension provision, *Journal of the European Economic Association*, **3**, 466–476.

Banks, J., Blundell, R. and Emmerson, C. (2015) Disability benefit receipt and reform: Reconciling trends in the United Kingdom, *Journal of of Economic Perspectives*, **29**, 173–190.

Banks, J., Emmerson, C. and Tetlow, G. (2005b) Estimating pension wealth of ELSA respondents, Working Paper W05/09, Institute for Fiscal Studies.

Barclays, C. (2013) The equity gilt study 2013, Technical report, Barclays Capital, London.

Benítez-Silva, H., Buchinsky, M. and Rust, J. (2011) Induced entry effects of a $1 for $2 offset in SSDI benefits, (mimeo).

Blau, D. M. (1998) Labor force dynamics of older married couples, *Journal of Labor Econometrics*, **16**, 595–629.

Blau, D. M. and Gilleskie, D. B. (2006) Health insurance and retirement of married couples, *Journal of Applied Econometrics*, **21**, 935–953.

Blundell, R., Britton, J., Dias, M. C. and French, E. (2016a) The dynamic effects of health on the employment of older workers, Working Paper WP 2016-348, University of Michigan Retirement Research Center.

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

Blundell, R., Crawford, C. and Jin, W. (2014) What can wages and employment tell us about the uk's productivity puzzle?, *Economic Journal*, **124**, 377–407.

Blundell, R., Crawford, R., French, E. and Tetlow, G. (2016b) Comparing retirement wealth trajectories on both sides of the Pond, *Fiscal Studies*, **37**, 105–130.

Borella, M., De Nardi, M. and Yang, F. (2017) The aggregate implications of gender and marriage, *The Journal of the Economics of Ageing*, forthcoming.

Bound, J. and Burkhauser, R. V. (1999) Economic analysis of transfer programs targeted on people with disabilities, *Handbook of Labor Economics*, **3C**, 3417–3528.

Bound, J., Schoenbaum, M., Stinebrickner, T. R. and Waidmann, T. (1999) The dynamic effects of health on the labor force transitions of older workers, *Labour Economics*, **6**, 179–202.

Bound, J., Stinebrickner, T. and Waidmann, T. (2010) Health, economic resources and the work decisions of older men, *Journal of Econometrics*, **156**, 106–129.

Cagetti, M. (2003) Wealth accumulation over the life cycle and precautionary savings, *Journal of Business and Economic Statistics*, **21**, 339–353.

Campolieti, M. and Riddell, C. (2012) Disability policy and the labor market: Evidence from a natural experiment in Canada, 1998-2006, *Journal of Public Economics*, **96**, 306–316.

Cannon, E. and Tonks, I. (2011) Annuity markets: Welfare, money's worth and policy implications, Working Paper 24, Netspar.

Chen, S. and van der Klaauw, W. (2008) The work disincentive effects of the disability insurance program in the 1990s, *Journal of Econometrics*, **142**, 757–784.

Crawford, R. (2012) ELSA pension wealth derived variables (waves 2 to 5): Methodology, Tech. rep., Institute for Fiscal Studies.

Crawford, R. and O'Dea, C. (2016) Household portfolios and financial preparedness for retirement, (mimeo).

Crossley, T. F. and Kennedy, S. (2002) The reliability of self-assessed health status, *Journal of Health Economics*, **21**, 643–658.

de Jong, P., Lindeboom, M. and van der Klaauw, B. (2011) Screening disability insurance applications, *Journal of the European Economic Association*, **9**, 106–129.

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

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

De Nardi, M., French, E. and Jones, J. B. (2016) Medicaid insurance in old age, *American Economic Review*, **106**, 3480–3520.

Disney, R., Emmerson, C. and Wakefield, M. (2006) Ill health and retirement in Britain: A panel data-based analysis, *Journal of Health Economics*, **25**, 621–649.

Duffie, D. and Singleton, K. (1993) Simulated moments estimation of markov models of asset prices, *Econometrica*, **61**, 929–952.

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

French, E. and Jones, J. B. (2011) The effects of health insurance and self-insurance on retirement behaviour, *Econometrica*, **79**, 693–732.

French, E. and Song, J. (2014) The effect of disability insurance receipt on labor supply, *American Economic Journal: Economic Policy*, **6**, 291–337.

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

Gruber, J. (2000) Disability insurance benefits and labor supply, *Journal of Political Economy*, **108**, 1162–1183.

Heckman, J. (1979) Sample selection bias as a specification error, *Econometrica*, **47**, 153–161.

Judd, K. L. (1998) *Numerical Methods in Economics*, MIT Press.

Jürges, H. (2007) True health vs response styles: exploring cross-country differences in self-reported health, *Health Economics*, **16**, 163–178.

Kapteyn, A. and Meijer, E. (2014) A comparison of different measures of health and their relation to labor force transitions at older ages, in *Discoveries in the Economics of Aging* (Ed.) D. A. Wise, University of Chicago Press, pp. 115–150.

Karlström, A., Palmeb, M. and Svensson, I. (2008) The employment effect of stricter rules for eligibility for DI: Evidence from a natural experiment in Sweden, *Journal of Public Economics*, **92**, 2071–2082.

Kostøl, A. R. and Mogstad, M. (2014) How financial incentives induce disability insurance recipients to return to work, *American Economic Review*, **104**, 624–655.

Lindeboom, M. and van Doorslaer, E. (2004) Cut-point shift and index shift in self-reported health, *Journal of Health Economics*, **23**, 1083–1099.

Low, H., Meghir, C. and Pistaferri, L. (2010) Wage risk and employment risk over the life cycle, *American Economic Review*, **100**, 1432–67.

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

Maestas, N., Mullen, K. J. and Strand, A. (2013) Does disability insurance receipt discourage work? Using examiner assignment to estimate causal effects of SSDI receipt, *American Economic Review*, **103**, 1797–1829.

Meijer, E., Kapteyn, A. and Andreyeva, T. (2011) Internationally comparable health indices, *Health Economics*, **20**, 600–619.

Meyer, B. D. and Mok, W. K. (2013) Disability, earnings, income and consumption, Working Paper 18869, National Bureau of Economic Research.

Mullen, K. J. and Staubli, S. (2015) Disability benefit generosity and labor force withdrawal, Working Paper WR-1132, RAND.

Newey, W. (1985) Generalized method of moments specification testing, *Journal of Econometrics*, **29**, 229–256.

Newey, W. and McFadden, D. (1994) Large sample estimation and hypothesis testing, Handbook of Econometrics, North Holland, Amsterdam.

OECD (2010) *Sickness, Disability and Work: Breaking the Barriers*, OECD Publishing.

Pakes, A. and Pollard, D. (1989) Simulation and the asymptotics of optimization estimators, *Econometrica*, **57**, 1027–1057.

Pischke, J.-S. (1995) Measurement error and earnings dynamics: Some estimates from the psid validation study, *Journal of Business and Economic Statistics*, **13**, 305–314.

Poterba, J. M., Venti, S. F. and Wise, D. A. (2013) Health, education, and the postretirement evolution of household assets, *Journal of Human Capital*, **7**, 297–339.

Powell, J. L. (1994) Estimation of semiparametric models, Handbook of Econometrics, North Holland, Amsterdam.

Ruh, P. and Staubli, S. (2015) The impact of the Substantial Gainful Activity Cap on disability insurance recipients' labor supply, Working Paper NB 15-09, National Bureau of Economic Research.

Rust, J. and Phelan, C. (1997) How Social Security and Medicare affect retirement behavior in a world of incomplete markets, *Econometrica*, **65**, 781–831.

Staubli, S. (2011) The impact of stricter criteria for disability insurance on labor force participation, *Journal of Public Economics*, **95**, 1223–1235.

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

van Ooijen, R., Alessie, R. and Knoef, M. (2015) Health status over the life cycle, Working Paper DP 10/2015-062, Netspar.

Venti, S. F. (2014) Comment on "A comparison of different measures of health and their relation to labor force transitions at older ages", in *Discoveries in the Economics of Aging* (Ed.) D. A. Wise, University of Chicago Press, pp. 151–156.

von Wachter, T., Song, J. and Manchester, J. (2011) Trends in employment and earnings of allowed and rejected applicants to the Social Security Disability Insurance program, *American Economic Review*, **101**, 3308–3329.

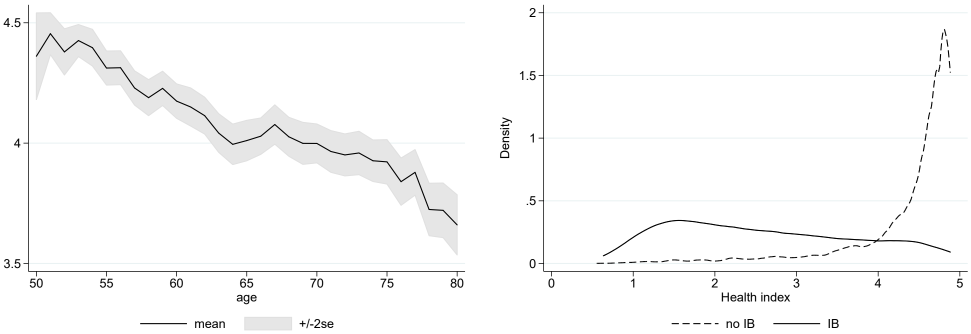
Wise, D. A. (Ed.) (2016) *Social Security Programs and Retirement around the World: Disability Insurance Programs and Retirement*, University of Chicago Press.

Appendix

A Additional Tables and Figures

Table A.1: Variables used for the computation of the health index.

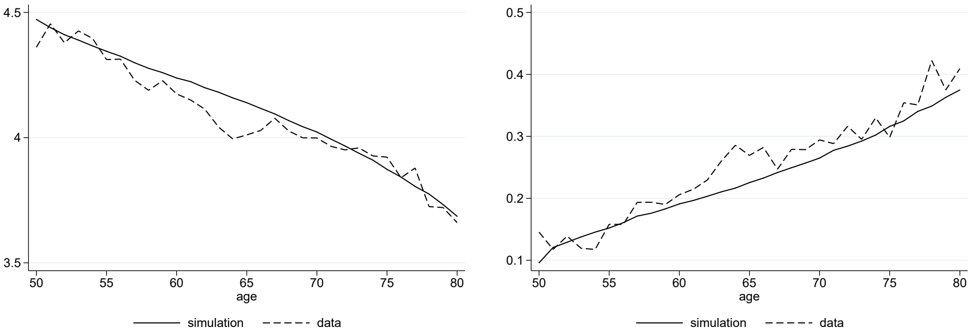
	ELSA variables	age$<$65	age$\geq$65
physical function	difficulty sitting 2 hours	0.12	0.12
	difficulty getting up from chair	0.17	0.26
	difficulty walking 100 yards	0.08	0.15
	difficulty climbing several flights stairs	0.19	0.37
	difficulty climbing one flight stairs	0.08	0.16
	difficulty stooping, kneeling or crouching	0.24	0.38
	difficulty reaching or extending arms	0.08	0.10
	difficulty pulling or pushing large objects	0.09	0.16
	difficulty lifting or carrying weights	0.11	0.19
	difficulty picking up 5p coin from table	0.04	0.06
sensory function	fair or poor eyesight	0.08	0.13
	fair or poor hearing	0.21	0.33
	problem of incontinence	0.05	0.10
mental health	Any emotional, nervous or psychiatric problems	0.10	0.06
	Depression (CESD scale)	0.24	0.27
limitations' intensity	at least one ADL	0.13	0.23
	at least one IADL	0.11	0.21
	Any pain	0.33	0.35
	Receiving care	0.12	0.23
diagnosed conditions	High blood pressure or hypertension	0.36	0.49
	Any heart problems	0.15	0.32
	A stroke (cerebral vascular disease)	0.02	0.08
	Diabetes or high blood sugar	0.09	0.14
	Chronic lung disease	0.05	0.09
	Asthma	0.11	0.12
	Arthritis	0.23	0.36
	Osteoporosis	0.01	0.03
	Cancer	0.04	0.11
	Parkinson's disease	<0.01	0.02
	Alzheimer's disease	<0.01	0.01
	Dementia	0.01	0.02



(a) Average health

(b) Health distribution by IB status

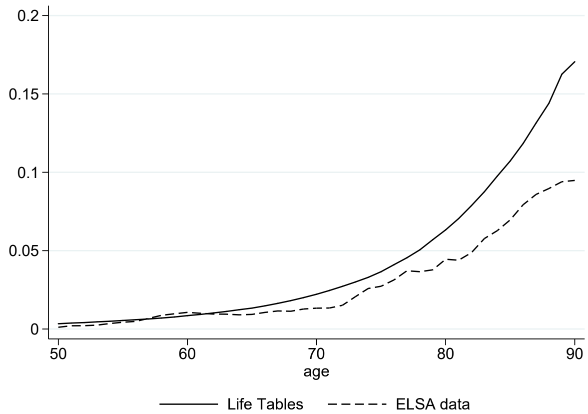
Figure A.1: Health index distribution by age and IB status.



(a)

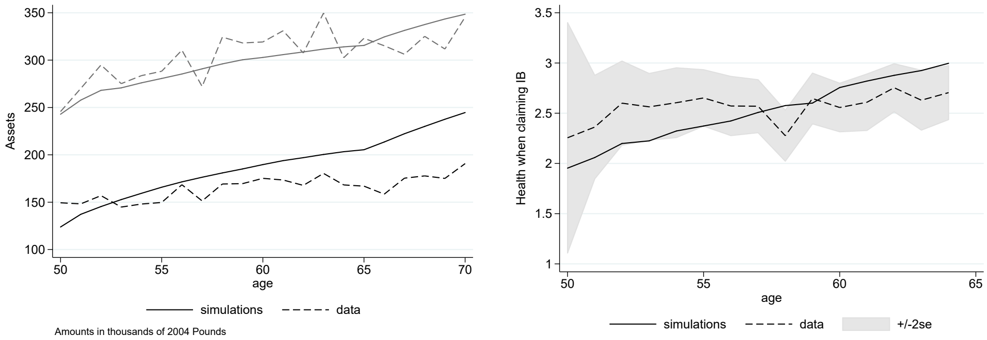
(b)

Figure A.2: Mean health distribution (a) and fraction of individuals with health below the first quintile of the unconditional health distribution (b). Simulations vs data.



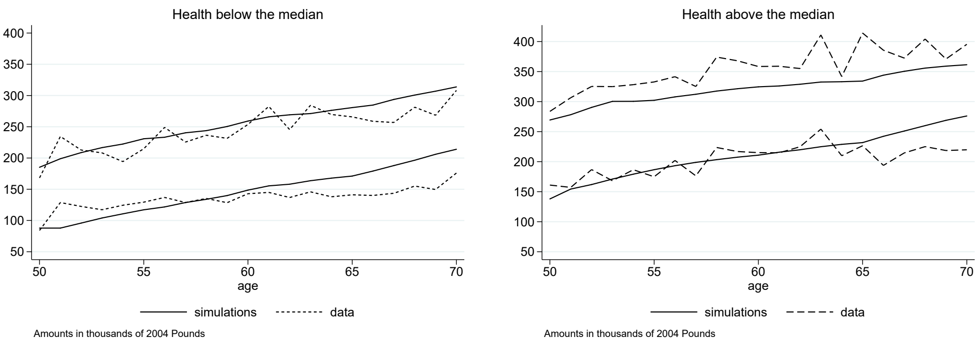
Note:ELSA mortality rates are computed considering the entire sample of males and not only cohabiting males.

Figure A.3: Comparison between data and Life Tables mortality rates.



(a) **(b)**

Figure A.4: Additional matched moments. First and second tertile of assets distribution by age (a), and Average health when receiving IB (b).



(a) **(b)**

Figure A.5: First and second tertile of assets distribution, for health below (a) or above (b) the median. Simulations versus data.

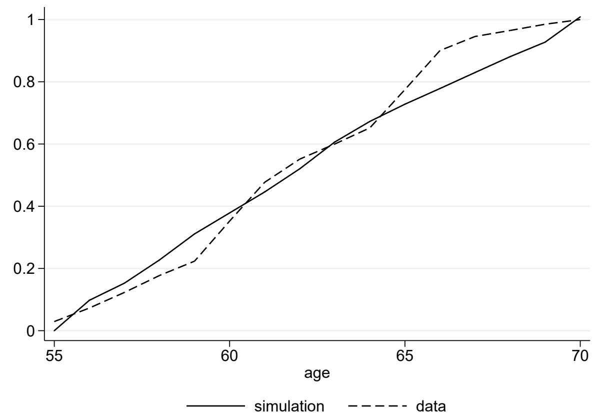


Figure A.6: Cumulative fraction claiming private pension.

B History of benefits targeted toward people with disability in the UK

Disability insurance has been introduced in 1948 under the name of *Sickness Benefit*. The benefit entitlement was linked to contributions whereas the benefit amount was a flat-rate not related to earnings. Benefit duration was unlimited and no distinctions were made between short- and long-term sickness. A medical assessment administered by personal doctors was required to get the benefit. With the introduction of the *Invalidity Benefit* in 1971 those who were receiving Sickness Benefit for more than 28 weeks were moved to this more generous benefit without the need of a new medical assessment. In 1983 *Statutory Sick Pay*, payed by employer, replaced the Sickness Benefit for the first 8 weeks (increased to 28 weeks in 1986). Sickness Benefit remained available for those not eligible for Statutory Sick Pay.

The sharp increase in public spending as well as the increase in the number of claimants were arrested with the 1995 reform which replaced Sickness Benefit and Invalidity Benefit with *Incapacity Benefit* (IB), taxable and paid up to state pension age. To qualify for the fist 28 weeks of benefit the medical assessment remained the same as for Sickness Benefit. A higher benefit was paid after the first 28 weeks, provided that the individual passed the 'suitable work test', administered at the regional level. Recipients may be able to do some types of work, called 'Permitted Work', within limits on weekly hours and earned income.

In line with the 1995 reform, the 1999 Welfare Reform and Pensions Act remarkably tightened eligibility conditions. Eligibility was tested with the Personal

Capability Assessment, a health test aimed at fostering return to work. In addition, contribution requirements referred only to contributions paid in the last three years before the start of incapacity. Finally, the reform introduced a benefit cut and a means-testing with regard to private pension income[35].

The *Pathways-To-Work* programme, started in 2003 as a pilot programme and then progressively extended in the following years, was instead aimed at facilitating IB claimants to move off benefit receipt and back into paid work. There are three main elements of the programme. The first one is a mandatory work-focused interview eight weeks after benefit claim if aged between 18 and 59, and other five monthly interviews for those remaining in the programme. The second element is the Return to Work Credit, a financial incentive to return to work paid to individuals who have received IB for at least 13 weeks and have found work, provided that they work at least 16 hours a week and they earn no more than £15,000 a year. The last element is a set of new and existing coaching activities, offered to those in receipt of IB, aimed at improving work readiness by helping individuals with job search and to manage health related problems within a work context (Adam *et al.*, 2010).

Finally, in 2008 *Employment and Support Allowance* (ESA) was introduced for new claimants in place of IB. A Work Capacity Assessment, stricter than the previous health test, determines eligibility to the benefit and classifies claimants into two groups: the support group and the work-related activities group. If classified as able to follow work related activities (WRA), individuals have to attend the Pathways-to-Work programme, those in the support group are instead entitled to the benefit without additional requirements. From 2011 to 2014, existing IB claimants had been reassessed and those eligible moved to ESA.

For individuals not eligible for contributory benefits, in the 1970s a set of benefits to compensate the extra cost endured by disabled individuals was introduced. The current benefits are the result of the 1992 reform which introduced *Disability Living Allowance* (DLA) for those starting to claim the benefit before age 65. For those aged over 65, *Attendance Allowance* (AA), introduced in 1971, remained available. Finally, means-tested benefits such as *Income Support* (IS) and *Working Tax Credits* (WTC) have specific premiums for disability. Details on these benefits are reported in the description of 2003/2004 Tax and Benefit system in Appendix G.

[35]For private pension income exceeding £85 a week the benefit amount was reduced 'by an amount equal to 50% of that excess'

C State and Private Pension benefits

The state pension provision is of two different types: Basic State Pension (BSP) that is received if individuals have payed National Insurance contributions for at least a quarter of their working life and Second Tier State Pension (STSP) which is related to earnings history. As stated in the model formulation, I assume that everyone is entitled to Basic State Pension full amount, whereas the initial endowment in earnings related state pension (Second Tier State Pension) is obtained from pension wealth derived variables provided in the released data, assuming the individual retires in the interview year. Pension wealth derived variables include the present discounted value of future or current pensions under some assumptions about earnings growth and participation decision up to state pension age. The amount of STSP is then updated each year according to the level of annual earnings (*earn*), provided that they are above the Lower Earnings Limit (LEL) and below the Upper Earnings Level (UEL):

$$STSP_{t+1} = \begin{cases} STSP_t + \frac{LET*0.4}{(SPA-16)} & \text{if } earn \in (LEL, LET] \\ STSP_t + \frac{LET*0.4+(UET-earn)*0.1}{(SPA-16)} & \text{if } earn \in (LET, UET] \\ STSP_t + \frac{(earn-LEL)*0.2}{(SPA-16)} & \text{if } earn \in (UET, UEL] \end{cases} \tag{C.1}$$

where LET and UET stand for Low and Upper Earnings Threshold. The threshold to compute Second Tier State pension accrual for 2003/2004 tax and benefit year are the following:

Table C.1: Threshold to compute STSP accrual (tax year 2003/2004).

Threshold	Description	Value for 2003/2004
LEL	Lower Earnings Limit	£3,900
LET	Low Earnings Threshold	£10,800
UET	Upper Earnings Threshold UET = 3xLET - 2xLEL	£24,600
UEL	Upper Earnings Limit	£30,420

At claiming age the state pension benefit is simply given by

$$sb_t = \begin{cases} STSP_t & \text{if} \quad age \geq SPA \\ 0 & \text{if} \quad age < SPA \end{cases} \tag{C.2}$$

In the model I assume that state pension claiming age is 65 for everyone: the state pension age for males in the UK is 65 and even if incentives to delay state pension claim have been increased since 2004, the majority of individuals claim state pension at age 65. Figure C.1 reports the age at which individuals are first observed receiving state pension using ELSA data. The picks at both age 65 and 66 is due to the biannual nature of the data and the type of information collected: we know whether they are receiving the benefit in the interview year but not in which year they started to receive the benefit. Complementing ELSA data with other sources of information

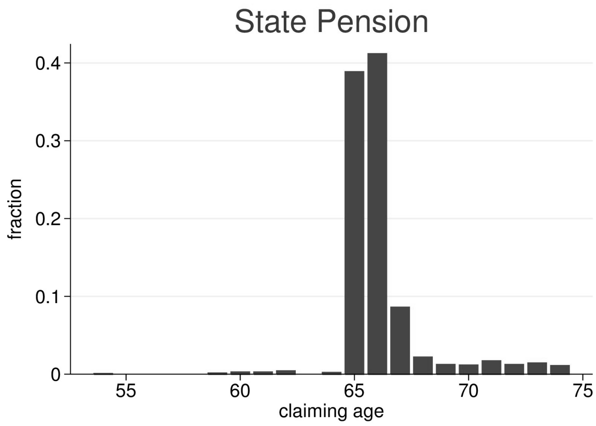
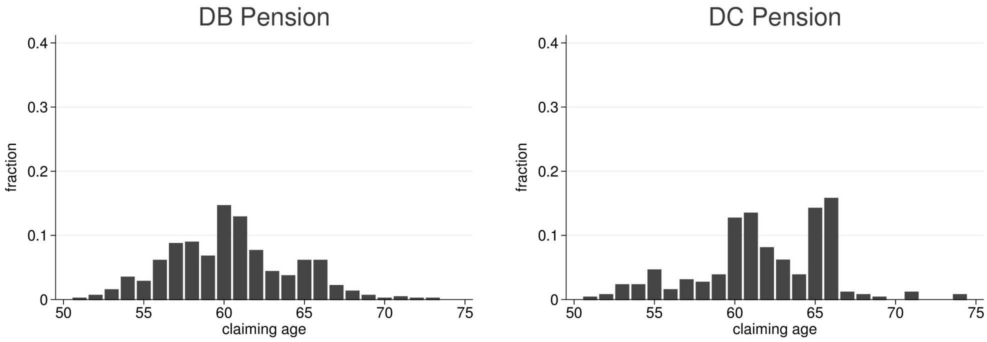


Figure C.1: Age at which individuals are first observed receiving state pension benefit. ELSA wave 1 to 6.

For what concerns private pensions, as in the US a progressive shift from defined benefit (DB) plans to defined contribution (DC) plans has been observed in the UK. Banks *et al.* (2005a) compare DB and DC plans with respect to financial risk, longevity risk, mobility and labour supply. While social security rules seem to be more relevant for the low wealth group, labour market incentives provided by DB and DC pension plans are crucial to understand the behaviour of the high wealth group. In particular the progressive shift from DB to DC could result in later retirement due to the fact that in a DC plan incentives are smoother across ages and accrual rates are higher at later ages than in a DB plan.

In Figure C.2, using ELSA data, I report the age at which I first observe individuals receiving a DB or a DC pension respectively. The distribution of age for DC pension recipients is shifted to the right with respect to the one for DB recipients as expected.



(a) **(b)**

Figure C.2: Age at which individuals are first observed receiving (a) DB pension (b) DC pension. ELSA wave 1 to 6.

In the model I consider only Defined Contribution Pension, and in particular I convert accrual in DB plan as if it was accrued in a DC fund (i.e. it gives rights to an equivalent benefit). I do not model the decision between DC and DB plan, individuals are assumed to be endowed with a DC plan (or no private plan) when entering the model at age 50.

Starting values for accrual endowments are obtained from pension wealth derived variables, available in the public released ELSA data. First, following backward the procedure explained in Crawford (2012) and Banks *et al.* (2005b), from pension wealth variables I compute the benefit amount for DC, DB and other private pensions to which the individual is currently contributing, from which he retains rights or from which he is currently receiving a pension. I sum up all the benefit amounts to which each individual is entitled and I assume that benefit amount is generated by the annuitization of a DC pension fund. To recover DC private pension accrual from benefit amount, I assume the following rule applies: $pb^{DC} = r_{DC}q^{DC}(1 - l)$, where pb^{DC} is the benefit amount, q^{DC} the accrual, r_{DC} the annuity rate which varies with age and l administrative costs set to 10%. Accrual measures derived with the described procedure inherit all the assumptions made in pension wealth computation regarding earnings history and employment decisions over the life-cycle.

D Exogenous processes

D.1 Moments derivation

Assuming for simplicity to have yearly data on health status, the equation to be estimated is

$$\log H_{it} = \pi_1^H age_{it} + \pi_2^H age_{it}^2 + \pi_3^H age_{it}^3 + \sum_{k=1}^{K} \delta_k^H 1\left\{size_{it} = k\right\} + \mu^H U_t + \zeta_{it}^H,$$

with $\zeta_{it}^H = f_i + \theta_{it} + \eta_{it}$.
The adjusted error term is defined as $g_{it} = \Delta\theta_{it} + \Delta\eta_{it}$.

The variance of g, the lag one covariance and the lag two covariance identify the three parameters of interest: $\sigma_{\nu_H}^2$, σ_η^2 and ρ_H.

$$Var(g_{it}) = \frac{2\sigma_{\nu_H}^2}{1+\rho} + 2\sigma_\eta^2$$
$$Cov(g_{it}, g_{it-1}) = \frac{\rho-1}{1+\rho}\sigma_{\nu_H}^2 - \sigma_\eta^2$$
$$Cov(g_{it}, g_{it-2}) = \frac{\rho(\rho-1)}{1+\rho}\sigma_{\nu_H}^2$$

Given that ELSA data are biannual, I assume that this does not affect the fixed part of the equation because age and health are contemporaneous and not lagged variables, however to get rid of the individual effect the first difference could only be computed between t and $t-2$. The adjusted error term becomes $\bar{g}_{it} = \Delta_2\theta_{it} + \Delta_2\eta_{it}$.

The moments that identify the parameters of interest are the variance of $\bar{g}_{it}$, the lagged two and lagged four covariances.

$$Var(\bar{g}_{it}) = 2\sigma_{\nu_H}^2 + 2\sigma_\eta^2$$
$$Cov(\bar{g}_{it}, \bar{g}_{it-2}) = (\rho^2 - 1)\sigma_{\nu_H}^2 - \sigma_\eta^2$$
$$Cov(\bar{g}_{it}, \bar{g}_{it-4}) = \rho^2(\rho^2 - 1)\sigma_{\nu_H}^2$$

I apply the same procedure and the same error specification to the wage process, assuming that the transitory component captures measurement error.

D.2 Selection in the wage profile

The data report only accepted wages, however individuals' decision to participate depends on the wage offer. Estimation of the wage process using only accepted wages might result in biased estimates if offered wages differ among those observed working and those remaining out of the labour market.

I solve this problem in two steps. First, I impute potential wages for non-workers estimating a regression model with selection by using full maximum likelihood (Heckman, 1979). Second, I estimate the wage process in Equation 5.2 using both observed and potential (imputed) wages.

In the first step I regress observed wages on a large set of individual characteristics: a polynomial in age, family size, educational level, time fixed-effects, a set of health controls[36], whether the individual is a smoker, whether he has a private pension plan, wealth quintiles and homeownership.

The exclusion restrictions in the selection equation cover two different aspects: financial incentives and family structure. The former is aimed at capturing institutional characteristics affecting labour supply decision and household financial constraints that might influence labour market attachment. The latter should capture family needs that affect participation decision (such as the presence of children and the health of the partner) and preferences of the couples to retire jointly, for example to spend time together.

Among controls for financial incentives affecting the decision to participate but not the wage offer I include the presence of a mortgage, whether the individual is above state pension age and whether he is above 55, which is the age from which individuals with a private pension plan can start to withdraw from the plan. For what concerns family structure I control for the presence of a partner, the presence of children, partner's health, whether the partner is above state pension age and whether the partner is above minimum age to withdraw from private pension plans.

The imputation model is used to obtain potential wages for non-workers. I then use offered wages (accepted and potential) to estimate the parameters of the wage process specified in Equation 5.2.

D.3 Survival probability

The derivation of survival probabilities conditional on health level is performed by assuming that the mortality risk perceived by individuals is consistent with the life tables.

I consider data mortality rates conditional on health for males born between 1946 and 1955 living with a partner. Individuals of that particular cohort are

[36]The health controls include dummies for limitations with (instrumental) activities of daily living, mobility limitations, heart problems, depression, eyesight, hearing, need of care and the presence of long standing illnesses.

observed up to age 67 and extrapolation, used to predict mortality rates up to age 90, over-predicts mortality rates with respect to the life tables.

I proceed as follow:

- I estimate the probability of being of health level i ($\hat{Pr}(H_t = i)$) and of dying by $t+1$ conditional on health level i ($\hat{Pr}(death^D_{t+1}|H_t = i)$) using all information for male respondents and controlling for cohort and family size effect;

- the probability of dying by $t+1$ at each age t is given by:

$$\hat{Pr}(death^D_{t+1}) = \sum_{i=1}^{4} \hat{Pr}(H_t = i) * \hat{Pr}(death^D_{t+1}|H_t = i);$$

- I compare the estimated probability with the life tables for each age t:

$$\frac{\hat{Pr}(death^{LT}_{t+1})}{\hat{Pr}(death^D_{t+1})} = \alpha_t$$

- I rescale each conditional probability in such a way that the unconditional probability matches the life tables:

$$\hat{Pr}(death^{LT}_{t+1}) = \sum_{i=1}^{4} \hat{Pr}(H_t = i) * \hat{Pr}(death^C_{t+1}|H_t = i)$$

  with $\hat{Pr}(death^C_{t+1}|H_t = i) = \alpha_t * \hat{Pr}(death^D_{t+1}|H_t = i)$.

E Moment conditions and asymptotic distribution of parameter estimates

Under regularity conditions (Pakes and Pollard (1989) and Duffie and Singleton (1993)), the MSM estimator $\hat{\vartheta}$ is both consistent and asymptotically normally distributed:

$$\sqrt{I}(\hat{\vartheta} - \vartheta_0) \sim N(0, \mathbf{V}) \quad \text{(E.1)}$$

with variance-covariance matrix $\mathbf{V} = (1+\tau)(\mathbf{D'WD})^{-1}\mathbf{D'WSWD}(\mathbf{D'WD})^{-1}$, where **S** is the variance-covariance matrix of the data, **D** is the Jacobian matrix of the population moment vector (Equation E.2) and **W** the plim of the weighting matrix $\hat{W}$.

$$D = \left. \frac{\partial \theta(\vartheta, \chi_0)}{\partial \vartheta'} \right|_{\vartheta=\vartheta_0} \tag{E.2}$$

I use a diagonal weighting matrix (Pischke, 1995) because the optimal weighting matrix ($\mathbf{W} = \mathbf{S}^{-1}$) is asymptotically efficient but can be severely biased in small sample. The variance-covariance matrix **S** and the weighting matrix **W** are estimated with their sample analogue. The partial derivatives in the Jacobian matrix **D** are straightforward to compute by taking the numerical derivatives of $\hat{\theta}_I(.)$, with the exception of asset quantiles for which the discontinuities in the moment function male it nondifferentiable at certain points. As in French and Jones (2011) (see their online Appendix D), I follow the approach for nonsmooth function described in Pakes and Pollard (1989), Newey and McFadden (1994) and Powell (1994) to derive the asset quantile components of **D**.

If the model is properly specified Newey (1985) shows that

$$\frac{I}{1+\tau} \hat{\theta}(\vartheta, \chi_0)' \mathbf{R}^{-1} \hat{\theta}(\vartheta, \chi_0) \sim \chi^2_{K-9} \tag{E.3}$$

where $\mathbf{R}^{-1}$ is the generalized inverse of **PSP**, with $\mathbf{P} = \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}$.

The χ^2 overidentification test rejects the model, this is not surprising because the profiles often lay outside the confidence intervals.

F The Employment and Support Allowance reform: evidence from model simulation.

The model can be used to investigate the effects of ESA introduction on labour supply and DI participation and to compare model predictions with the descriptive evidence presented in Banks *et al.* (2015) using ELSA data. Using a measure of disability that is the sum of the conditions reported by the respondent, Banks *et al.* (2015) find that both individuals with moderate and severe disability levels experienced about a 20 percentage point reduction in DI receipt after the introduction of ESA, whereas only moderately disabled individuals seem to have experienced a small increase in employment (3 percentage points)[37]. Reduced form results might be affected by confounding factors, such as other labour market trends. Blundell *et al.* (2014) show that labour supply has increased with the 2008 financial crisis among males aged 55 to 74, which can be partly explained by the drop in housing wealth that induced individuals to work longer. An advantage of the structural approach implemented in this paper is that no factors other than the DI benefit reform affect individual behaviour.

According to model predictions, a small decrease in the health threshold[38] reduces benefit participation among those receiving DI in the baseline scenario by 19 percentage points and increases labour market participation by 7 percentage points. This effect is almost entirely driven by those with health close to the eligibility threshold. When policies promoting the labour market participation of disabled individuals are considered, a 10% reduction in the fixed cost of work reduces DI participation by 5.3 percentage points among those receiving DI in the baseline scenario and increases participation by the same amount. These results support previous evidence of an important reduction in the number of benefit claimants after ESA introduction. They also suggest that an important role is likely to have been played by the stricter eligibility rules and that, switching off other confounding trends in labour supply, the effects on labour supply participation are likely to be higher than those found in previous research, at least when married males are considered.

Regarding the recent benefit cut for those in the WRA group, the simulations highlight the heterogeneity of the effects by health level and support the structure of the current benefit programme that has reduced the benefit amount for those with some working capacity (WRA group) and maintained a higher benefit for those in need of permanent support (Support group). One can speculate that individuals with health close to the eligibility threshold are more similar to those in the WRA group (under the ESA rules), whereas individuals with health far below the eligibility threshold are likely to be good proxies for those in the Support group (under the ESA rules). Focusing on those with health close to the eligibility threshold, the model can be used to quantify the expected affect of such a policy. The model predicts that a 10% reduction in the amount of the transfer reduces

[37]The comparison of disability and employment rates is performed between 2008 and 2012, with respect to changes observed for those classified without any disability according to their health measure.

[38]I consider a 10% reduction in the estimated threshold, that is, from a health level of 3.35 to a health level of 3.

IB participation by 4 percentage points, and among those leaving IB 7 out of 10 would return to work.

Finally, results on the effects of a reduction in the fixed cost of work inform the current debate about the extension of adviser activities (and thus some sort of work-related activities) to the Support group to reduce the employment gap between the disabled and the non-disabled. According to the simulations, individuals in very poor health are almost inelastic to marginal reductions in the fixed cost of work.[39]

G Tax and benefit system

The tax and benefit system considered is the one for 2003/04. The tax unit in the UK system is the individual. Three different types of social security benefits can be identified: contributory benefits (earnings-replacement benefits and pensions), non-contributory and non-means-tested benefits (they do not require contributions but they depend on some contingencies) and means-tested benefits (they depend both on contingencies and benefit unit income).

In the first category there are Jobseeker's Allowance (JSA), Incapacity Benefit (IB) and Retirement Pension. Contributory JSA is not included in the model in order to avoid strong assumptions on contribution requirements, income-based JSA[40] is implemented instead. IB is a benefit targeted to sick individuals with temporary or long-term inability to work. The benefit can be received up to state pension age. It is taxable and the amount of the benefit depends on the weeks of sickness: a lower short term rate up to week 28, a higher short-term rate for weeks 29-52 of sickness and finally a higher long-term rate until state pension age. In the model a flat rate is applied and it is assumed that the benefit is received for twelve months, given that decisions are taken annually. Retirement Pension can be received starting from state pension age (65 for males). If contribution conditions are met the pensioner receives a flat rate basic pension. In addition, if pensioners have contributed to the State Earnings Related Pension Scheme (SERPS) an earnings-related pension is also payable. Both components are taxable.

[39] In the 'Work, health and disability green paper' of October 2016, the government stated its commitment to reducing by 50% the employment gap between the disabled and the non-disabled. A proposed strategy to achieve this goal is to relax the strict distinction between the Support and WRA groups regarding the amount and type of work-related activities required of claimants. This means that work advisers would have discretion over required engagement in WRA for those in the Support group.

[40] Income-based Jobseeker's Allowance is presented in more details among means-tested benefits.

In the second category - non-contributory and non-means-tested benefits - those relevant for this analysis are Attendance Allowance (AA) and Disability Leaving Allowance (DLA). These two benefits target disable individuals. Assistance Allowance can be claimed after age 65 by individuals that due to illness or disability need care during the day and/or the night. Individuals younger than 65 with personal care or mobility needs due to disability can claim DLA. For both AA and DLA different rates apply depending on the care needed. They are not taxable. In the model I include DLA and AA as flat-rate benefits received when health follows below a calibrated threshold.

Finally, the third category includes income-based JSA , Income Support (IS), Pension Credit (PC) and Working Tax Credit (WTC).
For JSA and IS the unit of entitlement is the benefit unit, the claimants are unemployed and those not required to seek work (disable and pensioners) respectively. In addition of being exempt from looking for work, IS claimants need to be under 60. Additional rules that apply to both benefits are working less than 16 hours per week and having less than £8,000 in capital. The benefit tops up income to the 'weekly applicable needs' (IS/JSA=max(0,(NEEDS-INCOME))). The applicable amount is the sum of personal allowances, premiums and housing cost. In the implementation of the benefit I do not consider housing costs. Relevant allowances and premiums amounts are reported in Table G.1.

The disability premium can be received by those entitled to a disability benefit, such as AA, DLA or IB. The income measure used to determine the entitlement to IS and JSA includes gross income from employment and all other income sources except investment income, AA and DLA. To these amount contributions and income tax are deducted. For individuals entitled to disability premium an amount of £10 is disregarded, £5 are instead disregarded for all the others. Investment income does not enter directly in the income measure but a tariff income of £1 every £250 capital is calculated on financial capital between £3,000 and £8,000.

Since September 2003 a means-tested income support scheme very similar to the one presented above was available to people aged 60 and older (Minimum Income Guarantee), but starting from October 2003 it has been replaced with Pension Credit. The introduction of the programme aimed at increasing the take-up of income support among the pensioners. In the tax function implemented in the model I consider the post-reform scenario. Thus I present below the main characteristics of PC. The PC consists of two elements: the Guarantee Credit

(GC) meant to top up income to an 'appropriate minimum guarantee' and the Savings Credit (SC) meant to reward those who save for retirement. To be eligible to GC, individuals must be aged 60 or older and there are no capital limits. The tested income is the same as for IS with the exception that the tariff income is of £1 every £500 instead of every £250 and it is computed for capital above £6,000. The applicable needs are computed according to the basic allowance and the premium reported in Table G.1 (as for IS housing costs are not considered). Eligibility to SC requires being 65 or older and having means above the savings threshold, a reduced 40% taper rate applied to means above the threshold. The maximum amount receivable is reported in Table G.1. The income taken into account is the same as for GC except WTC that are deducted.

Table G.1: Income Support, income-based Jobseeker's Allowance, Pension Credit: allowances and premia.

	Single	Couple
IS - JSA		
Personal Allowance	54.65	85.75
Disability premium	23.30	33.25
Severe Disability premium	42.90	42.90
GC		
Personal Allowance	102.10	155.80
Severe Disability premium	42.90	42.90
SC		
Saving Credit threshold	77.35	123.80
Maximum amount	14.79	19.20

Finally, WTC are paid to low paid workers to top up their earnings. The means tested benefit is paid to working adults working at least 30 hours per week or working at least 16 hours per week and having a disability. The maximum amount of the benefit is given by the sum of a *basic element* and other additional elements (see Table G.2). I consider eligible for the disability element individuals whose health level is below the threshold for receiving DLA. The means are defined as earned income plus work related benefits before the deduction of taxes and social security contributions. If the means are below the threshold figure, the benefit is given by the maximum amount. If the relevant income is higher than the threshold, then the difference between the two amounts is tapered away at a 37% rate. The WTC is not taxable.

The income tax schedule is based on three bands.

Table G.2: Working tax credit

Basic element	30.17
Disability element	40.32
Severe Disability element	17.08
Income threshold	5060

Table G.3: Income tax schedule

Band	Rate on earned income	Rate on investment income
0-1960	0.1	0.2
1961-30500	0.22	0.2
30501-	0.4	0.4

The tax base includes earnings, private pensions, state pension, incapacity benefit and interest income (ra) net of personal tax-free allowances and other exemptions. The main tax allowances are listed in Table G.4.

For those aged less than SPA National Insurance payments are levied on earnings between a lower limit (£4,628) and the upper earnings limit (UEL £30,940) at a rate of 11%. Those having gross earning below the lower limit do not pay social insurance contributions, whereas those with earnings above UEL are subject to a rate of 1%. These rules apply to those who are contracted in.

Table G.4: Personal tax allowances and credits

Allowance/credit	Amount per year (£)
Single personal allowance: all individuals	£4,615
Age allowance: Age 65-74	£6,610 reduced to £4,615 (50% of income over £18,300)
Age allowance: Age 75+	£6,720 reduced to £4,615 (50% of income over £18,300)
Married Couples age allowance: Age 65-74	£5,565 reduced to £0 (50% of income over £18,300, less any reduction to personal age allowance)
Married Couples age allowance: Age 75+	£5,635 reduced to £0 (50% of income over £18,300, less any reduction to personal age allowance)