Disability Insurance and the Effects of Return-to-work Policies

Chiara Dal Bianco*

Abstract

I provide a quantitative assessment of the labor market and welfare effects of return-to-work policies targeted at disability insurance (DI) recipients. I do so by estimating a life-cycle model in which individuals with different health evolving over time choose consumption, labor supply, and DI application. I find that a wage subsidy incentivizing return to work is welfare improving, and the willingness to pay for such reform is increasing in sickness and decreasing in wealth. This policy increases labor force participation of DI beneficiaries by 4.6 percentage points, and decreases the DI rate by 5.7 percentage points. A policy mandating a 10% yearly eligibility reassessment would decrease the welfare of individuals in bad health and poor economic condition, and force about 30% of the beneficiaries to exit the program, 54% of whom would return to work.

Keywords: disability insurance, labor supply, older workers, health, return-to-work policies
JEL-Codes: J14,J26,I1

For their useful comments and suggestions, I thank the editor and the three anonymous referees, James Banks, Mariacristina De Nardi, Hamish Low, Eric French, my advisor Andrea Moro, Luigi Pistaferri, Stephen Pudney, Giacomo Pasini, Guglielmo Weber and Francesca Zantomio. I also thank the co-ordinating editor and the data editor for their support in the publication process. I 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 Labor 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.

*Department of Economics and Management, University of Padova - via del Santo 33 - 35123 Padova, Italy. *Tel.*: +39 049 8274048. ORCID: 0000-0002-5449-0645. *E-mail*: chiara.dalbianco@unipd.it

1 Introduction

In most OECD countries the rules governing disability insurance (DI) programs have changed considerably in the recent decades with a remarkable shift towards policies aimed at re-integrating people experiencing a DI episode into the labor market. These return-to-work policies, such as in-work benefits, provision of employment support and vocational rehabilitation are aimed at containing DI program expansion – in OECD countries in recent years public spending on disability amounted on average to 2.1% of GDP and it has proven to be rather stable over time – and reduce its work disincentive effects (OECD, 2010).

Despite the introduction of the return-to-work policies in many institutional contexts, there is heterogeneous evidence on their effects. What is the labor supply response to their introduction? Do they affect DI program enrollments? Which individuals are more affected? Are individuals better off after their introduction? The answers to these questions are crucial to effectively design return-to-work policies. In this paper, I address these questions by estimating a model of labor supply and savings behavior, which allows me to account for the full dynamic effects of alternative policies on agents' choices and welfare. In particular, I focus on two specific policies that are illustrative of the main changes in the DI program of many countries in the last decade: a wage subsidy received when returning to work after a DI episode and a continuous eligibility reassessment of DI beneficiaries.

I focus on individuals approaching retirement age for whom DI has been shown to play an important role in their departure from the labor market (Wise, 2016); in particular I consider men living with a partner. In the model, individuals choose whether to participate in the labor market and how many hours to work. Moreover, they can apply for DI. DI applicants are accepted with probability depending on health and age. The model allows for uncertainty about wage realization, health development, and life expectancy. In developing the model, I devote special attention 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 health conditions covered by the health assessment used to determine eligibility for DI benefits. Health depends on age and on 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 estimate the model using ELSA data from 2002–2008, a period in which DI policies and rules 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 age profiles of assets, labor force participation (LFP), 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 individual responses to two alternative policy scenarios. First, I introduce financial incentives to return to work in the form of a wage subsidy received in the first year of work after a DI episode that is proportional to the DI benefit amount. Second, I implement a yearly eligibility reassessment of DI beneficiaries which occurs with a certain probability. Because I use a more detailed measure of health than in previous works, I can better investigate heterogeneity in behavioral responses.

I document a number of relevant findings that can be summarized in three broad conclusions. First, my findings suggest the presence of a significant capacity to work among relatively younger claimants with less severe disabilities that can be brought out by return-to-work policies. If DI beneficiaries were allowed to receive the benefit for an additional year when returning to work, the DI rate for those receiving DI at least once, would decrease by 5.7 percentage points, and LFP would increase by 4.6 percentage points for the same group. Reassessing the eligibility of 10% of DI recipients every year at random would force about 30% of beneficiaries to exit the program. Of these, about 54% would return to work. Second, the one-year wage subsidy increases the program's generosity and DI enrollment for a specific group of individuals, namely those approaching retirement age that are at the margin of program entry (i.e. in relative good health and relatively good economic conditions), but this effect does not prevail. Third, the welfare effects are on average small but positive for the first policy scenario, negative and approximately zero for the second. In both cases, they exhibit large heterogeneity by health level. The willingness to pay for the introduction of a wage subsidy is positive for both low- and high-health individuals and it increases with the level of sickness, reaching is maximum for those with moderate disability. On the contrary, a policy introducing a continuous health reassessment would increase the welfare only among individuals in good health and strongly reduce the welfare of those in bad health.

Results proved to be robust to alternative model specifications regarding the relative risk aversion coefficient, the variance of the temporary health shocks, the fixed cost of work and the introduction of application costs.

The remainder of the paper is structured as follows. Section 2 presents a literature review and discusses my paper's contributions. Section 3 presents the UK DI system and Section 4 the model of lifetime decision making. Section 5 introduces the data and the health measure. Section 6 presents the estimation strategy and estimation results. Section 7 shows the model's ability to replicate the main patterns observed in the data. Section 8 presents the results of the policy reforms and their robustness to alternative model specifications and Section 9 concludes.

2 Related literature

This paper is primarily related to the recent literature on the effectiveness of return-to-work incentives targeted at DI beneficiaries. In particular I investigate the effect of financial and non-financial incentives to return to work. In both areas, the results in the literature are mixed.

Focusing on return-to-work reforms that exclusively involve changes in financial work incentives, Benítez-Silva *et al.* (2011) use a life-cycle framework to show that a proposed US policy that allows DI recipients to keep a portion of their benefit if they return to work has negligible effects on both DI inflows and LFP. Kostøl and Mogstad (2014) explore a similar reform implemented in Norway and find positive and sizable effects on LFP; however, they cannot estimate the level of induced DI entry. Campolieti and Riddell (2012) conclude that the introduction of an annual earnings disregard in Canada has substantial effects on LFP but no effects on DI entry or exit. Zaresani (2018); Zaresani *et al.* (2021) find sizeable effects on labor supply and earnings of return-to-work policies comparing two Canadian provinces with different clawback regimes. Ruh and Staubli (2018) explore a relaxation of the earnings restrictions in Austria and find an increase in LFP among DI beneficiaries that is entirely offset by the increase in new claimants.

For what concerns non-financial return-to-work incentives, mainly in the form of health reassessments during the DI episode, to the best of my knowledge there is no much evidence in the literature, apart from Autor and Duggan (2006) and Adam *et al.* (2010) that consider both financial and non-financial incentives together and find modest effects on LFP. The literature has mostly concentrated on the effect on LFP and DI rate of increased or decreased stringency of the initial health assessment to enter the program (Low and Pistaferri, 2015; Gruber and Kubik, 1997; Karlström *et al.*, 2008; Staubli, 2011; Borghans *et al.*, 2014; de Jong *et al.*, 2011). This literature mainly conclude that stricter screening processes are likely to reduce DI dependence and increase LFP, signaling the presence of substantial working capacity among beneficiaries. However, depending on the generosity of the welfare system, screening stringency might generate substitution effects with other benefits. Low and Pistaferri (2015) evaluate the welfare consequences of more stringent health assessment and find that more generous entry conditions increase welfare whereas more stringent one decrease it.

I contribute to these two strands of the literature by providing additional evidence regarding the dynamic incentives and the welfare consequences of return-to-work policies. The structural approach with respect to reduced form analyses allows the evaluation of the welfare consequences of policy reforms and the assessment of policies that have not been implemented or that are too costly to evaluate in an experimental setting. In particular, I consider a proportional wage subsidy paid for the first year of work after a DI episode (financial return-to-work incentive) and a continuous DI eligibility reassessment (non-financial return-to-work incentive) which has recently received attention both in the UK and in the US institutional contexts. For what concerns financial return-to-work incentive, with respect to Benítez-Silva *et al.* (2011), my model considers heterogeneity in LFP by level of health and therefore allows the investigation of differences in behavioral responses. Additionally, I document the two competing mechanisms at play when introducing financial return-to-work incentives–an increased incentive to exit the program induced by the wage subsidy and an increased benefit generosity with potential increase in DI entry–and their importance in the UK institutional context, with particular attention to the characteristics in terms of health, age and wealth of individuals responding differently to these incentives. Moreover, I investigate a rather unexplored policy scenario in which DI beneficiaries are reassessed with a certain probability on a yearly basis, accounting for the potential interaction with other welfare benefits.

Finally, this paper contributes to the growing literature on the welfare consequences of DI changes. Very few papers provide an estimate of the willingness to pay for policy changes. Low and Pistaferri (2015) use estimates from a dynamic life-cycle model of labor supply and DI application to show that a less strict screening process increases welfare because the effect of reduced false rejections offsets the effect of increased false applications. Moreover, welfare is improved by more generous payments because the positive effect of an increased level of insurance among the more disabled outweighs the negative effect of too high insurance among the less disabled. Looking at a return-to-work policy, Benítez-Silva *et al.* (2011) find instead that the possibility of working while retaining part of the benefit is welfare-improving among the DI beneficiaries, but the overall effect is negative because the costs for the other taxpayers are higher than their willingness to pay for having this opportunity. In contrast to Low and Pistaferri (2015) who model the behaviors of US men aged 23–62, I focus on older workers approaching retirement age. Moreover, I complement their analysis by focusing on policy reforms aimed at increasing program exit rather than controlling program entry and I use a richer health measure that allows me to better account for heterogeneous responses to the policy reforms. Unlike Benítez-Silva *et al.* (2011) I find a positive welfare effect of working while retaining the benefit not only for DI beneficiaries but also for individuals in good health. While both models predict modest effects on DI entry and discontinuous LFP for DI beneficiaries receiving the wage subsidy, there are some crucial differences in the institutional backgrounds and in the policy considered that might help explain the different results. In the UK, the DI program provides a low flat rate benefit amount as opposed to the earnings-related benefit of the US: the replacement rate of DI in the UK was about 15% of average earnings in the late 1990's (Banks *et al.*, 2015b), whereas it was between 23 to 74% in the US in the same period (Autor and Duggan, 2003). Therefore, the cost of the subsidy is likely to be higher in the US than in the UK. Moreover, I consider a policy that subsidizes only the first year of work after a DI spell and not all the subsequent years up to DI exit as in Benítez-Silva *et al.* (2011). The policy I consider is less generous but effective in inducing individuals on the margin of program entry to exit DI in the UK context.

3 Disability insurance programs in the UK

In the UK, DI was first introduced in 1948 under the name of *Sickness benefit*. The benefit entitlement was linked to contributions whereas the benefit amount was flat, unrelated to earnings and with unlimited duration. A medical assessment administered by personal doctors was required to get the benefit. In 1971, a more generous benefit (*Invalidity benefit*) was introduced.[1]

The sharp increase in public spending as well as the increase in the number of claimants were arrested with the 1995 reform (see Figure 3.1) which replaced Sickness benefit and Invalidity benefit with *Incapacity benefit*, taxable and paid up to state pension age. To qualify for the first 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. Contribution requirements referred only to contributions paid in the last three years before the start of incapacity[2] and a new instrument was introduced to assess eligibility, the *Personal capability assessment* (PCA). The PCA was a self-assessed health test aimed at fostering return to work. The PCA was integrated with documentation on diagnosed conditions sent by the applicant's doctor to the Department of Work and Pension and it might be followed by a medical examination.[3] Rejected applicants had the right to appeal within one month of the date of the final decision. Appeals were considered by a tribunal, likely to include a lawyer and a doctor. While the appeal was pending, unemployment benefit could be received.

The *Pathways-to-work* program, started in 2003 as a pilot program and then progressively extended in the following years, was instead aimed at facilitating Incapacity benefit beneficiaries to move off benefit receipt and back into paid work. There were three main elements of the program. The first one was a mandatory work-focused interview eight weeks

[1] In 1983 a major reform was introduced that transferred administration of sick pay claims to employers, who are mandated to pay Statutory Sick Pay for the first eight weeks of sickness (increased to 28 weeks in 1986). Sickness benefit remained available for those who would not qualify to Statutory Sick Pay.

[2] The reform introduced also a benefit cut and a means-testing with regard to private pension income: for private pension income exceeding £85 a week the benefit amount was reduced 'by an amount equal to 50% of that excess'.

[3] People who assessed 80% disabled, who already received other means-tested disability benefits targeted to severely disable individuals, or with severe mental illness documented by a doctor, are exempted from the PCA.

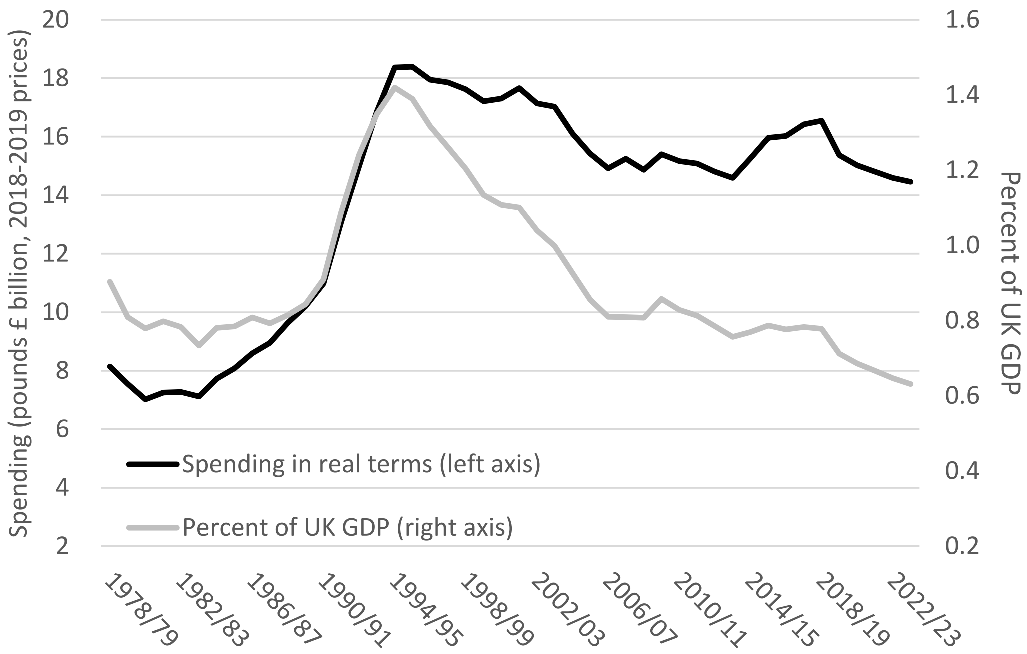


Figure 3.1: Total spending on Disability Benefits in the UK, 1978–79 to 2022–2023

Source: Department of Work and Pensions, Benefit expenditure and caseload tables, November 2018 (`https://www.gov.uk/government/publications/benefit-expenditure-and-caseload-tables-2018`.)
Notes: Spending on Sickness Benefit, Invalidity Benefit, Severe Disablement Allowance, Income Support on grounds of disability, Incapacity Benefit, and Employment and Support Allowance.

after benefit claim if aged between 18 and 59, and other five monthly interviews for those remaining in the program. The second element was the Return to Work Credit, a financial incentive to return to work paid to individuals who had received Incapacity benefit for at least 13 weeks and had found work, provided that they worked at least 16 hours a week and they earned no more than £15,000 a year. The last element was a set of new and existing coaching activities, offered to those in receipt of Incapacity benefit, 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* was introduced for new claimants in place of Incapacity benefit. A *Work capacity assessment*, stricter than the previous health test, determined eligibility to the benefit and classified claimants into two groups: the support group and the work-related activities group. If classified as able to follow work related activities, individuals had to attend the Pathways-to-work program, those in the support group were instead entitled to the benefit without additional requirements. From 2011 to 2014, existing Incapacity benefit claimants had been reassessed and those eligible moved to Employment and support allowance.

Figure 3.1 reveals that the 1995 reform has arrested the sharp increase in DI spending, whereas the introduction of Employment and Support Allowance did not reduce the spending further. This was mainly due to a lower than expected reduction in the number of claimants after the introduction of the new health assessment (Banks *et al.*, 2015a). Figure A.4 in Appendix A shows clearly that the numbers on DI benefits have continued to grow.

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, other means-tested benefits (*Income support* and *Working tax credits*) have specific premia for disability.[4]

The DI program rules implemented in the model mimic those of the Incapacity Benefit program in force between 1995 and 2008. Compared to the US, in the UK there is no need to be out of the labor market for several months in order to apply for DI, both employed and unemployed individuals can apply for the benefit, and there is not a long waiting period for the final examiner decision: in a few weeks after applications accepted applicants start to receive the benefit.

4 The model

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

$$U(c_t, l_t) + E_t \left[\sum_{j=t+1}^{T+1} \beta^j \Pi^s(j-1, t) \left(\pi_j^s U(c_j, l_j) + (1 - \pi_j^s) b(a_j) \right) \right], \tag{4.1}$$

where $t = 1, 2, \ldots, T$. 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 denoting 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 $t - 1$. Subject to the constraints outlined below, the household head maximizes Equation (4.1) by choosing consumption (c_t), hours worked (h_t), whether to apply for DI ($dapp_t$) or continue to claim it (d_t).

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

$$U(c_t, l_t) = \frac{1}{1-\nu} (c_t^\gamma l_t^{1-\gamma})^{1-\nu}. \tag{4.2}$$

The parameter γ represents the consumption weight, and the lower the γ the greater the weight placed on leisure. The parameter ν, given the CRRA utility function, represents the relative risk aversion coefficient and the elasticity of intertemporal substitution of the

[4] Details on these benefits are reported in the description of 2003/2004 Tax and Benefit system in Appendix I.

consumption and leisure composite good, for which a Cobb–Douglas aggregator is used. The elasticity of intertemporal substitution of consumption, holding the labor supply fixed, is given by $1/(\gamma * (\nu - 1) + 1)$. Hours of leisure l_t are defined as follows:

$$l_t = L - h_t - \phi_H(\hat{H} - H_t) - \phi_{P_t}\mathbb{1}(h_t > 0), \tag{4.3}$$

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).[5] 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 labor market ϕ_{P_t}. The fixed cost of work is allowed to vary with age[6], such that $\phi_{P_t} = \phi_{P_0} + \phi_{P_1}t$. The intercept of the fixed cost of work is allowed to vary at State Pension Age (SPA) for part-time workers, and the fixed cost becomes $\phi_{P_t} = \phi_{P_2} + \phi_{P_1}t$; this is because in the UK an increasing fraction of individuals remains active after SPA but reduces hours worked. This seems not to be driven by policy incentives or financial constraints but mainly to individuals declaring not be ready to stop working (ONS, 2015).[7]

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

$$b(a_t) = \phi_B \frac{(a_t + K)^{(1-\nu)\gamma}}{1 - \nu}. \tag{4.4}$$

The parameter K, which is positive, regulates the curvature of the bequest function 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):

$$\begin{gathered} H_t = \omega_H(age_t) + \theta_t + \eta_t \\ \theta_t = \rho_H \theta_{t-1} + \nu_t^H, \qquad \nu_t^H \sim N(0, \sigma^2_{\nu_H}), \qquad \eta_t \sim N(0, \sigma^2_\eta). \end{gathered} \tag{4.5}$$

[5] 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 and for that reason medical expenses should not be so relevant, at least between age 50 and 70 before the costs of institutionalization arise (they are not covered by the National Health Care system). Figure A.3 in Appendix A shows the marginal importance of out-of-pocket expenditures and voluntary payment schemes in the UK with respect to the US. (Source: OECD statistics.)

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

[7] In a robustness in Section 8.3 I impose that the intercept does not vary with age and investigate the robustness of the results to this alternative specification.

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 random walk component ϵ_t.

$$\log w_t = \omega_w(H_t, age_t) + \epsilon_t$$
$$\epsilon_t = \epsilon_{t-1} + \nu_t^w, \qquad \nu_t^w \sim N(0, \sigma^2_{\nu^w}). \tag{4.6}$$

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^H, ν_t^w) 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 labor 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.[8]

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

Finally, the asset accumulation equation is of the form:

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

where $y(\ldots, \tau)$ is after tax income; r is the real interest rate; di_t is the DI amount, received if the individual is allowed the benefit after application; pb_t is pension benefit, received from State Pension Age (SPA); and tr_t represents non-taxable transfers (such as Disability Living Allowance, Assistance Allowance, Income Support and Pension Credit). Each individual is endowed with a pension wealth that includes both public and private pension and, up to SPA, contributes to the fund a fixed fraction of his earnings. The tax function, τ, the modeled benefits and the pension wealth are described in detail in Appendix I. I assume that individuals cannot borrow against future pension income and means-tested benefits ($a_t \geq 0, \forall t$).

4.1 Disability benefits

The program rules implemented in the model mimic those of the Incapacity benefit program, the state-provided disability insurance program in force between 1995 and 2008 in the UK.[9] At each age between 50 and 64, individuals can decide to apply for DI. There is uncertainty

[8]This parametrization 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 labor supply within a couple. Female participation in the labor market has been shown to have an insurance role against permanent labor earnings risk within the household, and female participation rates increase as a result of additional uncertainty (Attanasio *et al.*, 2005).

[9]See Section 3 for details on the UK DI program and its evolution over time.

in the application process; I assume that the probability of being accepted is a decreasing function of health (H) and it is specified as a linear spline where the points to be interpolated are specific health percentiles. I set the acceptance probability to one when health takes its minimum value (zero) and at zero when health is above the median. The probability is α_1 at the 10th percentile, α_2 at the 20th percentile and α_3 at the 30th percentile of health.[10]

If an individual was already claiming the benefit in $t-1$ he can continue to receive the benefit irrespective of his health status. According to the legislation, DI 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. Even if DI is a contributory benefit, the amount of the benefit is flat; a lower amount is paid in the first 28 weeks and a higher amount is paid after having passed the Personal capability assessment. In the model, the decision period is one year; I therefore assume for simplicity that the annual benefit amount is fixed.

4.2 Dynamic programming problem and model solution

Timing of the model

In each period t, individuals observe the health and wage realizations (H_t and w_t), the disability benefit claiming status (d_{t-1}), their accrued pension wealth (q_t) and their assets (a_t). The vector of state variables is $X_t = (a_t, w_t, H_t, d_{t-1}, q_t)$. Then, they decide whether to participate in the labor market and whether to apply for DI. In the model, when individuals apply for DI, at the same time, they decide whether they would work if they were rejected. The uncertainty about the DI application is resolved at the beginning of the period, meaning that if the individual is rejected the benefit he can work, if this is his optimal choice. If instead he is allowed the benefit he will remain out of the labor market for that period. This is consistent with the UK DI system in which there is no need to be out of the labor market to apply for the benefit, both unemployed and working individuals can apply for DI. DI can be claimed up to SPA, which is 65 for men in the period considered. At SPA the pension wealth is annuitized. I assume that at age 70 everyone is retired and I set T equal to 90.

Value function

[10] An alternative specification of the acceptance probability as an inverse-logit ($p(H) = \frac{\exp(\alpha_0+\alpha_1 H)}{1+\exp(\alpha_0+\alpha_1 H)}$, where $p(H)$ is the acceptance probability and α_0 and α_1 the two parameters to be estimated) provides a very similar pattern of the acceptance probability, but a worse model fit, in particular of the fraction receiving DI. Estimates are available upon request.

Let $V(X_t)$ be the value function at time t. I can write the value function as

$$V(X_t) = \max \left\{ V^i(X_t) \right\},$$

where the superscript i denotes the possible discrete choice options. If the individual was not receiving DI in the previous period ($d_{t-1} = 0$) the possible choices are working ($h_t > 0$) or remaining inactive ($h_t = 0$), and applying for DI ($dapp_t = 1$) or not. If, instead, he received DI in the previous period ($d_{t-1} = 1$), he can continue to claim DI ($d_t = 1$), or exit DI and start working ($h_t > 0$). See Appendix C for details.

The model is solved backward starting from period T and computing the solution in each period by assuming that agents form expectations about future realizations of the state variables according to the transition probabilities assumed by the model.

5 Data

I use data from ELSA, a biennial longitudinal survey, representative of English private household population[11] aged 50 and over that started in 2002. It contains detailed information on 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.[12]

The first column of Table 5.1 reports descriptive statistics for male respondents living with a partner and aged 50-70 interviewed in waves 1–6 of ELSA (from 2002 to 2012). Men with a cohabiting partner represent 81% of the sample. The average age is 60.6, 55% are working and the average number of yearly hours worked is 1,835. Participation decreases with age: it is about 88% at age 50 and it declines to 30% around age 65 (the Statutory retirement age in the UK), however a non-negligible fraction of individuals remains active until age 70 (7-10%). Their wealth, which includes liquid and illiquid assets but not pension wealth, amounts to 274,495£ and their employment income to 16,670£ on average.[13] Among those younger than 65 (age at which DI in not available anymore) 10.2% receive DI (from 5% at age 50 to 17% at age 64). DI receipt exhibits high persistence over time: on average only 18.4% of those receiving DI are observed exiting the program in the following wave (in 2 years) and this percentage shrinks with age. The fraction of new recipients is 2.82% and it increases with age.

The second column of Table 5.1 reports descriptive statistics for a specific cohort of individuals born in 1946–1953. They are about 3 years younger than the overall sample of respondents, show higher participation rate (about 80% are working) and employment

[11] The initial sample does not include institutionalized individuals, however if a member of a sampled household moves to an institution in the following waves he is followed and interviewed.

[12] See Appendix A for additional details on ELSA data.

[13] All amounts are expressed in 2004 prices.

income. Their assets are 7% higher and the fraction receiving DI is substantially lower (7.77%), this can be a cohort effect (see Figure A.2 in Appendix A) and/or the effect of policy changes (Banks *et al.*, 2015a).

In estimating the model I use initial conditions and data moments computed for this cohort of individuals.

Table 5.1: Summary statistics.

	All sample	1946-1953 cohort
age	60.56	57.91
working (%)	54.80	71.87
hours | hours>0	1835	1873
wages (£) | hours>0	16,670	18,654
assets (£)	274,495	293,069
receiving DI (%)	10.22	7.77
outflow from DI in 2 years (%)	18.40	19.68
new recipients (%)	2.82	1.98

Notes: Due to the biennial nature of the data, the outflow from the benefit refers to a 2 year period. The same is true for the new recipients which are measured as the percentage observed in the benefit in year t but not in year $t-2$ (i.e. the previous wave).

5.1 Measuring health

Health status is measured on a continuous scale on the basis of the rich set of health indicators collected in ELSA. I follow Poterba *et al.* (2013) and apply principal component analysis to a set of variables covering several dimensions of individuals' health. In particular, given the categorical nature of the health questions, I use polychoric principal component analysis.[14] Differently from Poterba *et al.* (2013), I did not include self-reported general health to generate the health index. Among the rich set of information available in ELSA, I select health indicators that are less likely to suffer from measurement error (Crossley and Kennedy, 2002), heterogeneity in health perception (Lindeboom and van Doorslaer, 2004) and justification bias (Bound *et al.*, 1999). The full set of "objective" measures that I consider are reported in Table A.1 in Appendix A.[15] The choice of excluding self-reported general health, however, does not substantially affect the health indicator and its distribution over time as shown in Appendix B (Table B.2).

[14] Several continuous measures that use a set of (objective and subjective) indicators to recover 'true' latent health have been proposed in the literature, 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 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 aspects of health are of interest.

[15] 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 hospitalization data, particularly for the mental health domain.

In Table A.1, the health dimensions are classified into three categories: physical functions, mental functions and diagnosed conditions. The first two categories mimic the classification of the health descriptors asked in the Work Capability Assessment (WCA), which is the Department of Work and Pensions's current method of determining a person's ability to perform any type of work. This health test is in place since the ESA introduction in 2008. The WCA is a measure of the extent to which a person is incapable of performing certain specified everyday activities, covering both physical functions and mental, cognitive or intellectual functions.[16] The ELSA questions covering physical functions are very similar to those asked in the WCA. It is less the case for the mental functions. However, the dimensions investigated in the WCA (daily living activities, completion of tasks, coping with pressure and interaction with other people) are reasonably captured by the set of Instrumental Activities of Daily Living (IADL), the Center for Epidemiological Studies-Depression (CES-D) scale and specific questions on psychiatric problems and cognitive problems. The third set of indicators that I include is about diagnosed conditions that might compromise physical and mental functions and therefore complement previous health questions.

Table 5.2: Distribution of the health measure.

mean	10^{th} pct	20^{th} pct	30^{th} pct	50^{th} pct	90^{th} pct
0.813	0.533	0.707	0.799	0.884	0.952

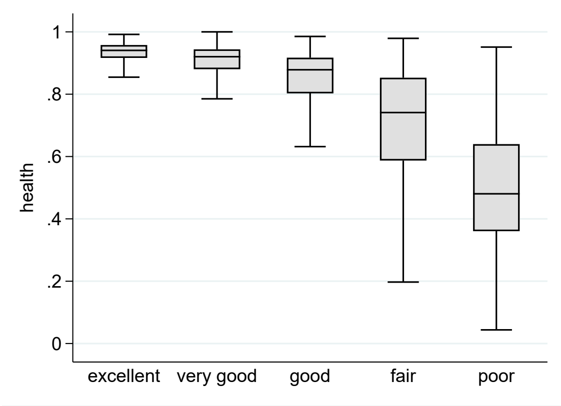
Note: Statistics computed for male ELSA respondents aged 50 to 80, interviewed from wave 1 to wave 6.

The health index varies between zero and one: higher values of the index denotes good health. Its distribution is reported in Table 5.2. It has mean 0.813 and median 0.884, it is left skewed and the left tail captures sickness severity with values below 0.533 (the 10th percentile) signalling very poor health.[17] Figure 5.1a reports the distribution of the index for each level of self-reported general health. The correlation between the two measures is high and the median values of the continuous index conditional on self-reported health differ substantially. However, the figure suggests that individuals reporting themselves in poor health might have high levels of health according to the continuous measure. The continuous index can also be compared with an overall health assessment that captures the presence of temporary or permanent health problems limiting the kind or amount of work the individual can do (a disability measure similar to the one used in Low and Pistaferri (2015)). Figure 5.1b shows the average health, according to the continuous index, for individuals reporting at least once health problems lasting for more than three months that limit their working capacity in the age range 50–65, compared with the average health of those reporting at most temporary health problems or no health problems at all. The level of health differs

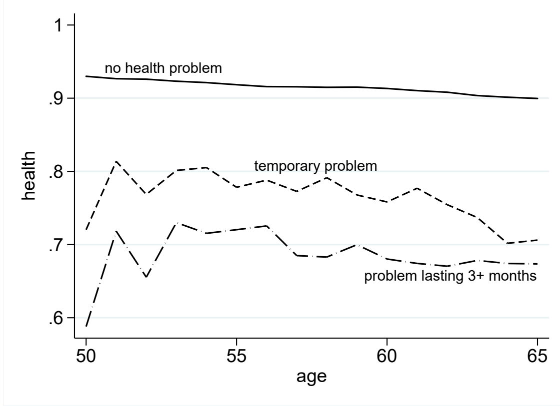
[16]See `https://www.disabilityrightsuk.org/wca-limited-capability-work-assessment-descriptors` for details.

[17]This distribution is consistent with the finding of Hosseini *et al.* (2021) on the life-cycle dynamic of a frailty index built using PSID data.

substantially in the three groups suggesting that my measure captures well the severity of the health limitations.



(a) Health distribution by self-reported health categories.



(b) Average health for different severity of health problems

Figure 5.1: Comparison of the generated continuous health index with self-reported general health (left panel) and with (temporary or permanent) health problems limiting the kind or amount of work the individual can do (right panel). ELSA wave 1 to 6.

Finally, the distribution of the index is substantially different between men receiving and not receiving DI (see Figure 5.2), with higher values of the index for individuals not receiving DI. The fact that there are individuals not receiving DI despite their very low health level and individuals receiving DI when in good health (which is confirmed when we consider alternative health indicators such as self-reported general health, see Figure A.1 in Appendix A) suggests the DI program does not perfectly target those in need. This can be due to, for example, individuals recovering from bad health shocks that remain into the program, or the inability of the application procedure to correctly classify those in need.

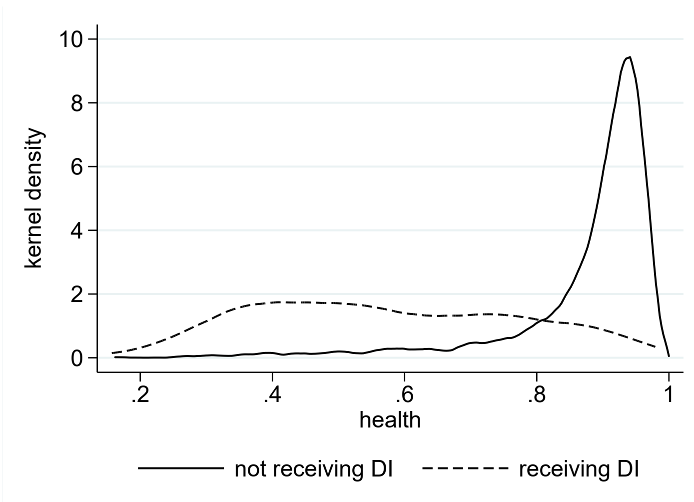


Figure 5.2: Health distribution by DI status. ELSA wave 1 to 6.

6 Model estimation

I estimate the model parameters in two steps.[18] 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 by minimizing a weighted distance (GMM function) between the simulated life-cycle profiles and the data life-cycle profiles.

6.1 Estimation of the 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 reform to the DI program implemented starting from 2008 for new DI claimants did not effect health developments, wage offers and mortality risk.

Health process

I estimate the health process parameters of Equation 4.5 using data for men aged 50 to 90.[19] Even if the interest is on individual decisions up to age 70, I include individuals aged above 70 because the model is solved and simulated up to age 90. I first estimate the fixed effect regression in Equation 6.1 to obtain an estimate of the age parameters ($\hat{\pi_1}^H$, $\hat{\pi_2}^H$ and $\hat{\pi_3}^H$). In the fixed effect regression I include family size to control for potential changes in health status due to changes in the family structure (i.e. the death of a spouse) and unemployment rate to control for time effects.[20] The real-valued dependent variable (H^*_{it}) is the logit transformation of the health measure which varies in the 0-1 interval (H_{it}).

$$H^*_{it} = f_i + \pi_1^H age_{it} + \pi_2^H age^2_{it} + \pi_3^H age^3_{it} + \sum_{k=1}^{K} \delta_k^H 1\left\{size_{it} = k\right\} + \mu^H U_t + \zeta^H_{it} \tag{6.1}$$

$$\zeta^H_{it} = \theta_{it} + \eta_{it}$$

To estimate the parameters of the random component, I use the 'adjusted residuals' ($\tilde{\zeta}^H_{it}$), that are the residuals from the regression in Equation 6.1 with the fixed effects added back and corrected to be representative of individuals born between 1946 and 1955. The three parameters of the random component ($\sigma^2_{\nu_H}$, σ^2_η and ρ_H) plus the initial variance at age 50 (σ^1_0)

[18] As in Gourinchas and Parker (2002), French (2005) and French and Jones (2011) among the others.

[19] I end up with 19,034 individual-year observations for 5,963 distinct respondents.

[20] When I simulate from the estimated processes for health and wages and from data profiles for decision variables presented in Appendix E, I fix family size at two and set the unemployment rate at 4.9%, which is the 2004 annual unemployment rate for men in England (Source: labor Force Survey. ILO unemployment rate. `http://www.ons.gov.uk/employmentandlabormarket/peoplenotinwork/unemployment/timeseries/ycol/lms`).

are identified by the variances and covariances of the adjusted residuals and are estimated using standard minimum distance techniques (see Appendix B for the details).

In the first column of Table 6.1, the parameter estimates of the third order polynomial in age are reported. As expected, health is decreasing with age, and the declining trend becomes steeper as age increases. The autoregressive parameter estimate is 0.967, suggesting high persistence of the process (see the second column of Table 6.1). Moreover, the initial variance for the stochastic component (σ_0^2) is large, indicating that a large part of the variability in health is predetermined at age 50. The estimated process is able to replicate well the average health, the health distribution, the variances and the covariances of the adjusted residuals over the life cycle (see Figure B.1, B.2 and Table B.1 in Appendix B).

Table 6.1: Parameters of the health process.

Deterministic component		Random component	
age	-0.503***	ρ	0.967***
	(0.126)		(0.011)
$age^2/10$	0.081***	σ_ν^2	0.063**
	(0.019)		(0.021)
$age^3/100$	-0.005***	σ_η^2	0.158***
	(0.001)		(0.028)
		σ_0^2	0.865***
			(0.056)
Observations	19,034		

These 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 as a health measure and find a high persistence process with an autoregressive parameter of 0.88 when self-reported health is corrected by means of objective health measures collected in hospitalization data. Using ELSA data, Blundell *et al.* (2016a) estimate a similar 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 the health dynamic, with estimated values of the autoregressive parameter ranging from 0.90 to 1.06. High persistence is also found by Hosseini *et al.* (2021) using a frailty index built from PSID data as measure of health.

Wage process

To recover the parameters of the wage process I estimate Equation 6.2, where the additional error term ξ_{it} is assumed to represent measurement error.

$$\log w_{it} = f_i + \pi_1^w age_{it} + \pi_2^w age_{it}^2 + \alpha_1^w H_{it} + \alpha_2^w H_{it}^2 + \sum_{k=1}^{K} \delta_k^w 1\{size_{it} = k\} + \mu^w U_t + \zeta_{it}^w \tag{6.2}$$

$$\zeta_{it}^w = \epsilon_{it} + \xi_{it}, \qquad \xi_{it} \sim N(0, \sigma_\xi^2)$$

Estimation of Equation 6.2 after first differencing the data allows to get rid of the unobserved heterogeneity that is likely to correlate with the health status. However, there might also be a selection bias issue if the wage growth differs between workers and non-workers, given that only accepted wages are observed. To account for selection into participation, I follow Low and Pistaferri (2015) and adopt a reduced form approach. I assume that financial incentives to participate in the labor market and the family structure serve as exclusion restrictions. To measure financial incentives, I use household financial constraints that influence the labor market attachment, and institutional characteristics that affect the labor supply decision. The family structure variables capture family needs that affect participation decision and the couples' preferences for joint retirement, for example to spend time together. Details are in Appendix B.

Using the parameter estimates from Equation 6.2, I derive the adjusted error term $\Delta(\epsilon_{it} + \xi_{it})$ and use its variance and lag-one convariance to identify the variances of the persistent and transitory random components of the wage process. See Appendix B for details.

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. Cross-sectional mortality rates computed using ELSA data are lower than comparable mortality rates from the life tables (see Figure B.4a in the Appendix B). 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 B for details on mortality rates derivation.

6.2 Method of simulated moments

Initial conditions

The model reproduces the behaviors of a specific cohort of individuals, those born between

1946 and 1955. To derive initial conditions for model simulation, I select men of this cohort living with a partner. I exclude self-employed individuals and I assume that marriage status does not change over time. I consider individuals interviewed in wave 1 and individuals entering the survey in wave 2. The final sample used to estimate the joint initial distribution of assets, pension wealth, DI claiming status, health and wages consists of 883 individuals. Each of the simulated individuals receives a draw of assets, pension wealth, DI claiming status, health and offered wage from the estimated initial distribution.

Calibrated parameters

Among preference parameters, I fix relative risk aversion ν, discount factor β, time endowment L and the parameters of the bequest function (ϕ_B and K). I set relative risk aversion of the composite good consumption-leisure ν to 2.[21] 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 4,880 hours.[22] Finally, I calibrate the curvature of the bequest function K and the bequest weight ϕ_B in such a way that the marginal propensity to bequeath is equal to 0.98 and the bequest motive becomes operative when the consumption value of total wealth exceeds $£8000$ (De Nardi *et al.*, 2010). These values are in line with those implied by the parameter estimates in French (2005). See Appendix D for details.

I set the amount of Incapacity benefit equal to the average amount received by men aged between 50 and 69 in 2004, that is $£3$,460.[23] The rate of return on the safe asset r is set at 0.029, the average real return on UK Government liability between 2002 and 2008 (Capital, 2013).

Estimated parameters and moment conditions

The remaining parameters are estimated using the Method of Simulated Moments to minimize a weighted distance (GMM function) between simulated life-cycle profiles and data life-cycle profiles (see Appendix E for details on data moments derivation). I use 177 moment conditions. I match the fraction receiving DI (15 moments) and the average health status for those receiving DI (15 moments) from age 50 to age 64. Moreover, I match the fraction receiving DI in t given that they were receiving DI in $t-2$ (13 moments) and the fraction entering DI (13 moments) from age 54 to age 64 . Furthermore, I match LFP from age 50 to age 69 (assuming that at age 70 everyone is retired) conditional on four health

[21] Subsection 8.3 shows that results are robust if the value of the relative risk aversion is set to 3.

[22] The time endowment is set by assuming 305 working days in a year and 16 hours per day to allocate between leisure and work.

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

intervals (80 moments), and annual hours worked from age 50 to age 69 (20 moments).[24] Finally, I match mean assets profile from age 50 to age 70 (21 moments).

In Appendix E I report a measure of the sensitivity of the estimator to perturbations of different moments conditions proposed by Andrews *et al.* (2017). This measure allows to more formally show which moments bear most heavily on which estimated parameters. The set of moments on DI (fraction receiving the benefit, inflow rate, persistence and average health), together with LFP for those with lower levels of health, mainly identify the parameters of the acceptance probability (α_i), see Figures E.6, E.7, E.8. Individuals in bad health work less than individuals in good health, and this heterogeneity in labor supply by health identifies the time cost of being in bad health (ϕ_H), see Figure E.2. As in previous studies, hours of work and LFP 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}) and how the fixed cost changes at age 65 (ϕ_{P2}), see Figures E.3, E.4, E.5. The labor supply profiles identify the consumption weight parameter (γ), see Figure E.1.[25]

Table 6.2: Structural parameter estimates.

Parameter	Description	Value	SE
γ	consumption weight	0.616	(0.002)
ϕ_H	cost of being in poor health (hours)	4,201	(30)
ϕ_{P0}	fixed cost of work at age 50 (hours)	909	(5)
ϕ_{P1}	age trend of the fixed cost of work (hours)	46	(0.5)
ϕ_{P2}	fixed cost of work at age 65 (hours)	467	(7)
α_1	acceptance probability – 10th pct of health	0.200	(0.010)
α_2	acceptance probability – 20th pct of health	0.124	(0.003)
α_3	acceptance probability – 30th pct of health	0.122	(0.003)

Table 6.2 reports the preference parameters' estimates. Holding labor supply fixed, the coefficient of relative risk aversion for consumption is given by $\gamma(\nu - 1) + 1$. The estimated consumption weight γ implies a coefficient of relative risk aversion for consumption equal to 1.616, in line with values estimated in the literature (Blundell *et al.*, 1994; Attanasio and Weber, 1995; Banks *et al.*, 2001). The estimated value for γ denotes quite strong preferences for work. The time cost of being in poor health is high. It is 4,201 hours for an individual with the worst level of health.

To compare this result with previous estimates that use categorical self-assessed overall health (good or bad), I compute the average health index for individuals in fair or poor health (bad health), and for individuals in good, very good or excellent health (good health). The

[24] 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. The resulting measure shows low variability across individuals and by health level. I therefore include only unconditional moments for hours worked.

[25] Details on the estimation procedure are presented in Appendix E.

average health index for the two groups is 0.64 and 0.88 respectively, which correspond to a time cost of 1,512 for those in bad health and 504 for those in good health. According to my model estimates, being in poor health with respect of being in good health entails a 21% reduction in the time endowment. This estimated value is at the upper bound of previous estimates. The same figure is 12% in French and Jones (2011) and 14-21% in Capatina (2015).

The fixed cost of work at age 50 is 909 hours and it increases by 46 hours each year. At age 65 the fixed cost of work reduces at 467 hours for those working part-time. It corresponds to a 19% (10%) reduction of the time endowment at age 50 (at age 65 if working part-time) in case of LFP. This figure is in line with French and Jones's (2011) estimates for older US individuals (20%), the value in Capatina (2015) is much higher instead (48–53%). Estimates of the acceptance probability parameters reveal that the uncertainty is concentrated in the first decile of the health distribution. The acceptance probability is set at one for the lowest level of health (zero) and at zero for health above the median. It is estimated to be 32% on average for health in the first decile and it remains rather constant at 13% for health in the second and the third deciles. These values imply a very large rejection probability for the unhealthy: 68% on average for those with health in the first decile and 87% on average for those with more modest levels of sickness - with health in the second and third deciles. Low and Pistaferri (2015), for the US, estimate a probability of being rejected of 57% for the unhealthy (those severely work limited) and 82% for those with moderate disability. Despite numerous stories of claimants finding it difficult to get support,[26] model assumptions might contribute to this difference in the rejection probability for the severely limited. In particular, the absence of direct or indirect costs of applying is likely to induce too many individuals to apply for DI. As a robustness check, in Subsection 8.3 I introduce a cost of applying for DI and discuss its potential implications.

7 Model fit

7.1 Targeted moments

The estimated model replicates the main facts observed in the data quite well. In Figure 7.1, I report the participation conditional on health level. 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.[27] As expected, participation decreases with age and is lower for individuals having a lower level of health. The model captures differences in participation for individuals with different

[26]Source: Institute for Fiscal Studies (`https://ifs.org.uk/publications/14011`)

[27]Health quantiles refer to the unconditional health distribution for individuals aged 50 to 90 in the data.

levels of health quite well. The profile for hours worked is reported in Figure 7.2a: in the data, the average number of hours worked declines linearly with age from about 2,000 hours at age 50 to 1,500 at age 65, when it sharply drops at 800 hours per year. Financial incentives provided by state and private pensions are not enough to explain the sharp drop in hours worked at SPA, allowing the fixed cost of work to change at age 65 for those working part-time (through the ϕ_{P2} parameter), increases the model ability in reproducing the discontinuity at age 65, at least partially.

Figure 7.2b reports simulated versus data profiles for average assets. In the data, assets are slightly increasing from age 50 to age 70, simulated mean assets are almost always within the confidence intervals.

Figure 7.3a shows the data and simulation profile for the fraction receiving DI. In the data, 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. The simulated fraction of DI claimants is very close to that in the data. In Figure 7.3b, I report the 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 well the high persistence of DI and its slight increase with age. The flow into DI is somewhat underestimated by the model (Figure 7.3c), as it is the average level of health of individuals claiming DI, in particular at younger ages (Figure 7.3d). However, in both cases, the simulated profiles lie within the confidence bars.

7.2 Moments not directly targeted in the estimation

The model can reproduce fairly well additional moments not directly targeted in the estimation procedure. Figure F.1 in Appendix F shows that simulated profiles for median, first and third terciles of assets match well the corresponding data profiles. Figure F.2 shows the match for earnings. Earnings reflects both the wage process and the labor supply decision (extensive and intensive margins). Also in this case, the simulated profiles match the data closely. Finally, Figure F.3 reports the fraction receiving DI by health level. Data profiles are noisy but the figure shows that the model is able to reproduce the observed heterogeneity. For what concerns other benefits included in the model, the fractions receiving Income Support and Pension Credit are well reproduced, at least up to SPA (Table F.1).

To facilitate the comparison with the existing literature and provide additional evidence on the model performance, I compute the elasticity of labor market non-participation and the elasticity of DI application to benefit generosity. To do that I simulate a revenue neutral policy reform (see Appendix H for details) in which the DI benefit amount is decreased (or increased) by 10%.[28] These figures are computed by means of a revenue-neutral policy

[28] I compute the elasticities using both reformed scenarios (a 10% increase or decrease in the benefit amount). The figures obtained largely coincide. In Table 7.1 I report the values obtained with a 10% reduction of the benefit amount. The first elasticity is computed as the ratio between the percentage change in DI application between the baseline and the reformed scenarios divided by the percentage change in DI benefit

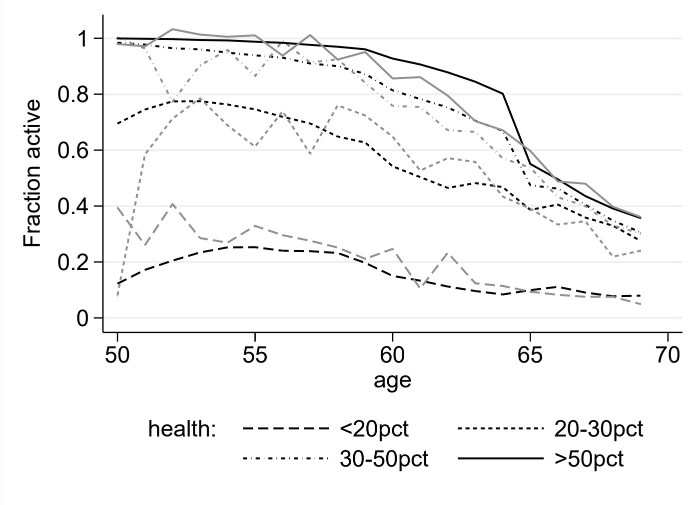
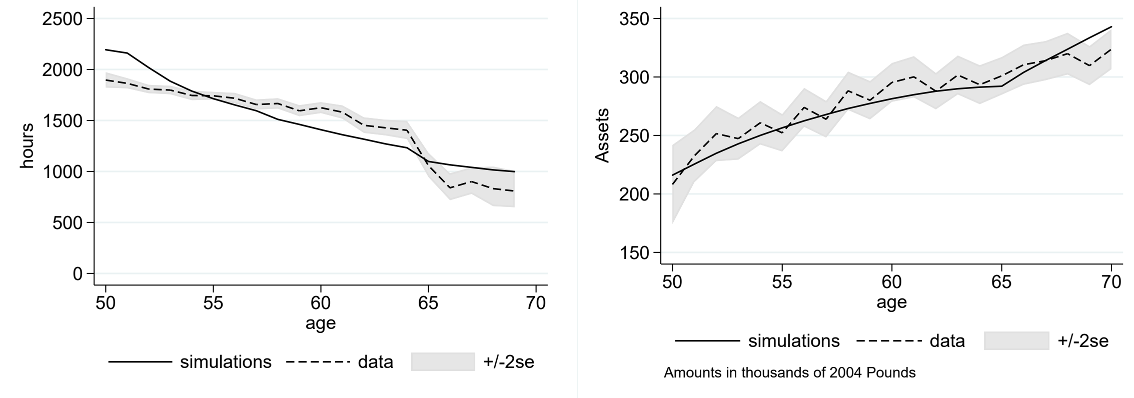


Figure 7.1: Life-cycle profiles of participation by health level. Simulations (dark gray lines) versus data (light gray lines).



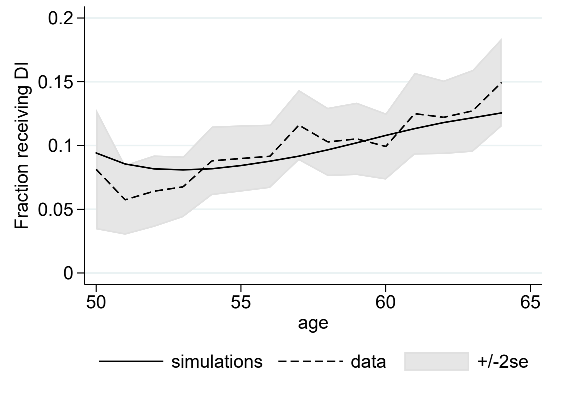
(a) Hours

(b) Assets

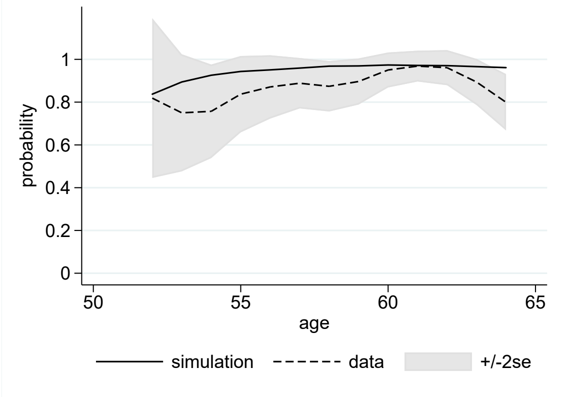
Figure 7.2: Life-cycle profiles of mean hours worked (left panel) and mean assets (right panel).Simulations versus data

intervention that reduced marginally the amount of DI. Table 7.1 reports the elasticities computed for the entire sample and by age group and health level. The elasticity of DI application is equal to 0.90 using all individuals in the sample. This value is in the range of previous empirical estimates for the US (from 0.2–1.3) surveyed by Bound and Burkhauser (1999). The elasticity varies remarkably with age and health level. Relatively younger individuals aged 50–54 show an elasticity 40% larger than those aged 60–64 (0.92 and 0.75, respectively). This is consistent with French and Song (2014) and Maestas *et al.* (2013), who find that younger individuals are more responsive to work disincentives provided by DI. Further, individuals with higher levels of health are elastic to changes in benefit generosity, whereas those in poor health hardly respond to the policy change. These results are in line with the heterogeneity found by Low and Pistaferri (2015) using three levels of

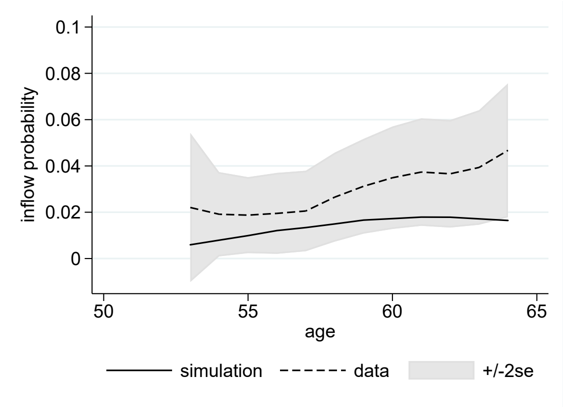
amount. The second elasticity is instead the ratio between labor market non participation in the baseline and reformed scenarios divided by the percentage change in DI benefit amount.



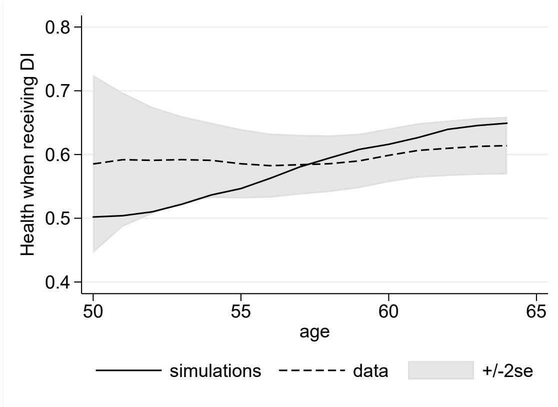
(a) DI rate



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



(c) $Pr(d_t = 1|d_{t-2} = 0)$



(d) Average health if in DI

Figure 7.3: Life-cycle profiles of DI rate, persistence, inflow and average health of DI recipients. Simulations versus data.

health. The elasticity of non-participation is equal to 0.36. The value is again in the range of previous estimates for the US and Canada (0.21–0.93) surveyed by Bound and Burkhauser (1999).[29] Compared to the US, in the UK there is no need to be out of the labor market for several months in order to apply for DI, and there is not a long waiting period for the final examiner decision. Therefore, one might expect greater elasticity than in the US. However, the peculiar aspect of the UK DI program of providing a low flat rate benefit amount results in mainly low-skilled workers with fewer employment opportunities entering the program. This means that given the characteristics of the target population, the demand for DI is likely to be rather inelastic to marginal changes in the benefit amount. This is consistent with the findings of Mullen and Staubli (2016), according to whom the elasticity of non-participation to benefit generosity in Austria is lower for low-skilled and poorer workers.

[29]More recently, Mullen and Staubli (2016) estimate elasticities of 0.7–1.2 on labor force withdrawal for Austria, and Marie and Vall Castello (2012) estimate an elasticity of -0.22 on LFP for the partially disabled in Spain.

Table 7.1: Elasticities of DI application and labor market non-participation to benefit generosity.

	All	Age			Health (percentiles)				
		50–54	55–59	60–64	<10	10–20	20–30	30–50	>50
DI application	0.90	0.92	1.07	0.75	-1.68	0.26	1.00	1.23	2.22
non-participation	0.36	0.66	0.51	0.17	0.01	0.59	0.72	0.36	0.12

8 Policy experiments

The estimated structural model is used to analyze the effects of changing the rules governing the DI program on labor supply, DI benefit participation, and welfare. I study two counterfactual policies. First, I focus on financial incentives to return to work and investigate the behavioral responses to a policy reform that allows DI beneficiaries to keep a portion of their benefit for an additional year if they start working. Second, I consider non-financial return-to-work incentives and simulate a policy scenario in which the DI beneficiaries are reassessed with a certain probability to verify their eligibility.

The behavioral responses to these policies show partial effects of the hypothetical reforms and are not meant to capture general equilibrium effects. In all experiments, I measure the welfare consequences of altering the DI rules as a consumption equivalent variation, that is a proportional reduction in consumption from age 50–90 that makes the individual ex-ante indifferent between the baseline (status quo) and the reformed scenario. All policies are revenue-neutral; whether positive or negative, the costs of policy changes are offset by proportionally adjusting the income tax rates (see Appendix H for details).

8.1 Financial incentives to return to work

In several institutional contexts, recent reforms of the DI program have introduced measures to promote the labor market inclusion of people with disabilities and to incentivize DI recipients to return to work. Some examples are the 2005 reform in Norway that allows DI recipients to keep a portion of their benefit when they return to work and the proposed "$1 for $2 offset" in the US, which consists of a reduction of $1 in benefits for every $2 in earnings above the earnings cap instead of a 100% reduction. Similarly, in the UK, the Pathways to Work program provides both financial and non-financial support to DI claimants to find a job and exit the DI roll. Its pilot implementation started in 2003, and the program became mandatory for all DI claimants found able to perform some kind of work (those placed in the "work-related activities group") with the introduction of the Employment and Support Allowance in 2008.

In this first policy experiment, I simulate the behavioral responses when DI beneficiaries returning to work after having received DI for at least a year are allowed to keep a portion of the benefit for an additional year, provided that they work a sufficient number of hours (750 per year) and do not earn more than £15,000 a year. This policy reform mimics the financial support provided by the Pathways to Work program. According to the program rules, individuals receiving DI for at least 13 weeks who find a job of at least 16 hours a week and do not earn more than £15,000 a year can continue to receive about half of the DI benefit amount for up to 12 months.

Two competing mechanisms are at play in policies providing return-to-work incentives: on the one hand, labor supply might increase among DI beneficiaries and some might be induced to exit DI, resulting in reduced program costs. On the other hand, the option of working while receiving DI increases the program generosity and thus DI enrollment.

The aggregate effects presented in Figure 8.1 shows that the labor supply increase prevails. The figure reports the overall effects of the reform for individuals aged 50–64 when DI beneficiaries are allowed to keep from 20% to 100% of the DI amount in the year following the return to work. The effects are non negligible. When 100% of the amount is received for an additional year, DI participation decreases by 0.56 percentage points (5.7%) and the increase in LFP is 0.45 percentage points (0.5%). The reduction in DI exceeds the increase in LFP, this means that the reduction in DI participation does not entirely translate into an increase in LFP.

Aggregate figures mask heterogeneous responses to the policy change. In what follows, I focus on the reformed scenario in which individuals are allowed to keep the entire DI amount for an additional year if they exit the program. Among individuals below SPA, 2.7% change their behavior. This percentage becomes 16.4% if I focus on those receiving DI at least once in the baseline or in the reformed scenarios. Table 8.1 illustrates the heterogeneity of the effects. Most recipients reduce the number of years in DI (93%) with an average reduction of 3 years. Those who increase DI participation (7%) do it by 3 years on average. In the first group, DI recipients mainly anticipate program exit (66%),[30] the second group instead

[30]The LFP of this group is discontinuous after a DI episode, with some individuals remaining out of DI (14%) and some others re-entering the program only for short periods depending on their health (10% increase the number of DI episodes).

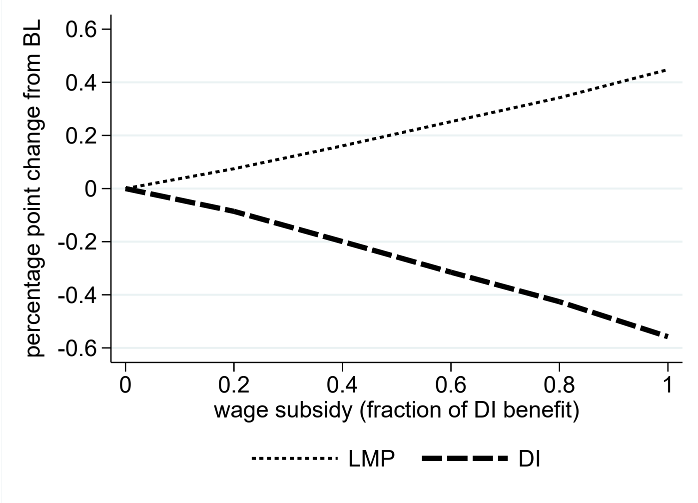


Figure 8.1: Effects on participation and DI rate of a proportional wage subsidy introduction.

is mainly composed by new DI recipients (79%). Figure 8.2 shows that those who change behavior are on average in better health than those who continue to receive DI as in the baseline scenario. The health of individuals spending more years in DI is slightly higher than the health of those reducing DI participation and they also have higher savings, but these differences shrink after age 55 when the change in behavior is mainly observed (see Table 8.1). This suggests that they are at the margin of program entry after age 55 and the increased generosity of the program induces them to apply.

Table 8.1: Statistics for those who increase or decrease DI participation if allowed to keep the DI benefit for an additional year when returning to work.

years in DI	% change DI part.	Δ years	% of DI recipients	age distribution (%)		
				50-54	55-59	60-64
less	93	-3	15.3	36	30	34
more	7	3	1.2	9	29	62

After having characterized those who change behavior, in Table 8.2 I focus on DI rate and LFP for those receiving DI in the baseline or in the reformed scenario (sub-sample 1, which refers to the net effect of the policy change) and those receiving DI in the baseline scenario (sub-sample 2, which refers to the partial effect of the reform for those reducing DI participation). Columns 3 and 6 of Table 8.2 shows that LFP increases by about 5 percentage points and DI rate decreases by 6.1 percentage points when considering only those reducing DI participation. The net effect in columns 2 and 5–when considering those receiving DI in the baseline or in the reformed scenario–is instead a 4.6 points increase in LFP and a

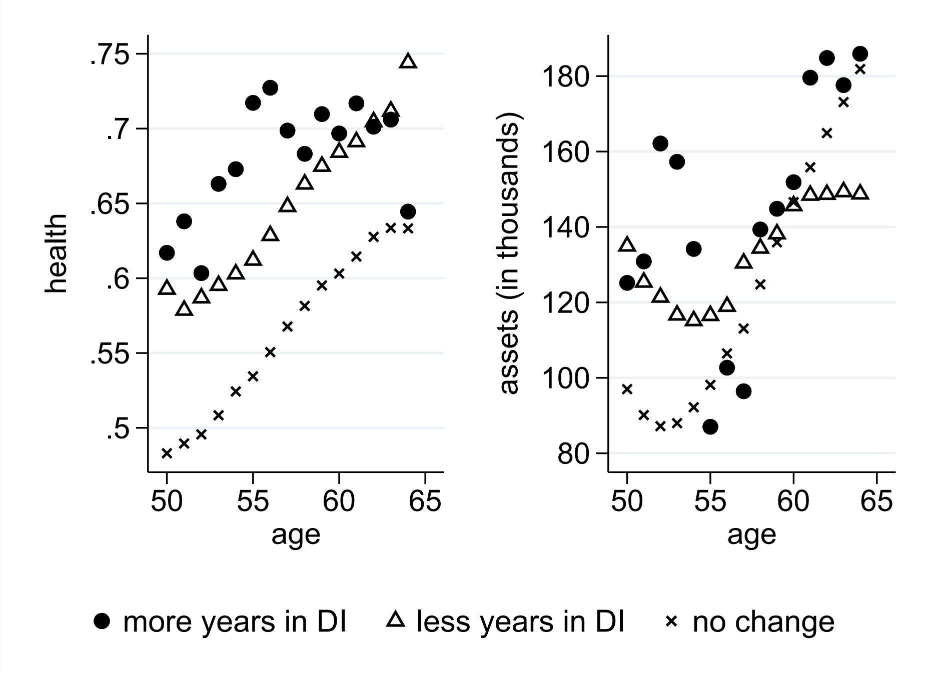


Figure 8.2: Patterns of health (left panel) and assets (right panel) for those who increase, decrease or do not modify DI participation if allowed to keep the DI benefit for an additional year when returning to work.

5.7 points decrease in DI rate. These figures can be compared with those estimated by Adam *et al.* (2010) analyzing data on the pilot implementation of the UK Pathways to Work program. They estimate a 5.8 percentage point increase in employment due to the reform.[31] The samples considered are different: Adam *et al.* (2010) use all new DI applicants aged 18–59 in a given year, whereas I look at the behavior over the life-cycle of a specific cohort of men. Consistently with my findings, the observed effect seems to be mainly driven by applicants aged over 40. Adam *et al.* (2010) do not find any increase in the application to DI as a consequence of the increased generosity of the program, probably due to the short time horizon considered (12-24 months). Instead, they document a negative short-run effect of the program on DI participation, which vanishes after the first 12 months. This is in part consistent with the model prediction according to which among those receiving the wage subsidy in the reformed scenario about 67% exit DI even in the absence of the return-to-work incentive.

The model allows to investigate the heterogeneity of the policy effects by health, wealth and age. Looking at the overall sample, those more responsive have health between the first

[31] Similar increases in LFP are also found by Campolieti and Riddell (2012) for Canada (5.7 percentage points) and Ruh and Staubli (2018) for Austria (6.7 percentage points). Smaller effects are found by Benítez-Silva *et al.* (2011) for the US and Kostøl and Mogstad (2014) for Norway (8.5 percentage points increase in LFP among claimants aged below 50 but no effect among those approaching retirement age).

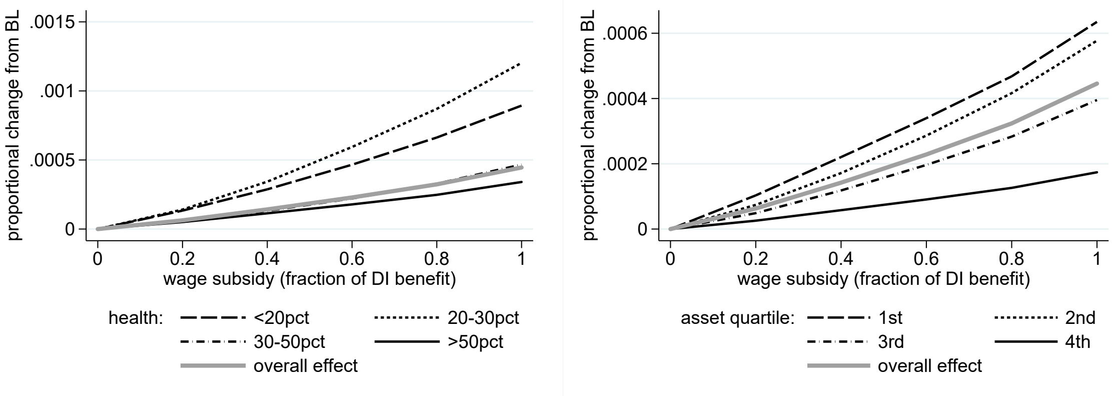
Table 8.2: Proportional changes in DI rate and LFP conditional on health, wealth and age due to a policy reform that allows to keep the DI benefit for an additional year when returning to work.

	Δ DI rate			Δ LFP		
	all	sub-sample 1	sub-sample 2	all	sub-sample 1	sub-sample 2
	(1)	(2)	(3)	(4)	(5)	(6)
overall	-0.56	-5.68	-6.12	0.45	4.63	5.00
health						
<10th pct	-1.24	-1.38	-1.41	0.06	0.08	0.09
10-20th pct	-3.82	-7.85	-8.17	2.97	6.41	6.59
20-30th pct	-1.66	-6.62	-7.92	1.51	6.16	7.43
30-50th pct	-0.56	-9.68	-10.29	0.53	8.97	9.58
>50th pct	-0.08	-11.97	-12.28	0.08	11.69	12.00
assets						
1st quartile	-1.40	-5.54	-5.84	1.13	4.38	4.66
2nd quartile	-0.62	-11.18	-12.28	0.50	9.42	10.34
3rd quartile	-0.16	-4.02	-4.60	0.11	3.33	3.85
4th quartile	-0.05	-1.16	-1.43	0.05	1.23	1.40
age						
50-54	-0.61	-7.21	-7.35	0.53	6.27	6.40
55-59	-0.50	-5.38	-5.78	0.40	4.47	4.84
60-64	-0.56	-4.75	-5.45	0.40	3.48	4.06

Notes: Statistics refer to individuals aged below 65 (all), receiving DI in the baseline or in the reformed scenarios (sub-sample 1) and receiving DI in the baseline scenario only (sub-sample 2).

and the third deciles, have lower wealth and are younger (see columns 1 and 4 of Table 8.2). More can be learned conditioning on those receiving DI in the baseline or reformed scenarios (sub-sample 1 and 2 in Table 8.2). The reduction in DI participation is increasing with health (from -1.4 to -12 percentage points) as it is the increase in LFP (from 0 to 12 percentage points), this suggests that the subsidy incentivizes those who have recovered from a bad health shock, and are still in DI, to exit the program. This is not surprising as individuals in relative good health are better-off if they reduce DI participation as they are compensated by a larger income–the wage subsidy, on average, represents 36% of employment income for individuals who exit DI due to the policy. Those most affected are the youngest, in fact the reduction in DI rate is 7.2 points for the 50-54 age group and 4.8 for the 60-64; LFP increases by 6.3 for the former and by 3.5 for the latter. Finally, the effect is non-monotonic in wealth: the effect picks in the second quartile (11.2 points decrease in DI and 9.4 increase in LFP) and than decreases for higher and lower levels of assets. The comparison between sub-sample 1 and sub-sample 2 in Table 8.2 confirms that the disincentive induced by the subsidy is concentrated on individuals in relative better health, aged above 55, and with assets in the second and third quartile.[32]

Finally, the model allows to compute the welfare consequences of the one-year wage subsidy introduction, and in particular to investigate heterogeneous welfare effects by health. Figure 8.3 shows that the overall effect of the revenue-neutral policy change on welfare is positive (grey lines) even if rather small. A wage subsidy of the same amount of the DI benefit implies a welfare gain of 0.04% of consumption. The gains are remarkably heterogeneous by initial health and wealth levels and they decrease as the level of initial health and wealth increase: they are three times higher among those with health below the 20th percentile (wealth in the first quartile) than those with health above the median (wealth in the fourth quartile). This highlights that the costs of providing an additional year of benefit and of the marginal increase in DI entry produced by the policy reform do not exceed the gain experienced by individuals in terms of decreased dependency from DI and increased attachment to the labor market. Notice that, in all reformed scenarios considered revenue neutrality is reached through a tax reduction (see Table H.1 in Appendix H). Moreover, the welfare increase for those in bad health (health below the 20th percentile) is mainly driven by the increased generosity of the program (12% only is driven by the tax reduction), whereas it is almost entirely driven by the tax reduction (about 90%) for those in good health (health above the median).[33]



(a) by initial health level

(b) by initial asset level

Figure 8.3: Welfare effects by initial health and initial wealth level of a proportional wage subsidy introduction.

[32]The asset quantiles are computed from the simulated distribution of assets for individuals aged 50–64 in the baseline scenario. The results are quantitatively the same when simulations from the reformed scenario are used instead.

[33]Computations are not shown. They refer to a wage subsidy of the same amount of the DI benefit and are obtained comparing welfare gains with and without adjustments of the tax rates to obtain revenue neutrality.

The result that the wage subsidy is welfare improving – as the effect of the policy on the flows off DI are more consequential than the flows into DI – is likely to be driven by specific features of the institutional context. The peculiar aspect of the UK DI program of providing a low flat rate benefit amount results in mainly low-skilled workers with less generous employment opportunities entering the program. This means that given the characteristics of the target population, the benefit inflows are likely to be rather inelastic to marginal changes in the benefit generosity, also due to the rather strict application procedures. The absence of a long waiting period in unemployment to apply for DI reduces also the skills depreciation of those entering DI, making the wage subsidy particularly effective in bringing them back to work.

8.2 Continuous reassessment of DI beneficiaries

Return-to-work policies may also include non-financial work incentives. For example, the Ticket to Work program in the US permits work trial periods without losing benefit entitlements and allows beneficiaries to retain eligibility to public healthcare insurance when they return to work but also requires a reassessment of beneficiaries' disability status. The UK Pathway to Work program, in addition to the financial support already mentioned, requires beneficiaries to participate in regular interviews with an adviser to improve work readiness by helping individuals with job search and with managing health-related problems within a work context. These policies are aimed at reducing the implicit tax on labor supply and often imply a continuous monitoring of the health evolution of beneficiaries.

To assess the effect of continuous health monitoring to increase the incentive to exit DI, I simulate individuals' behaviors under alternative scenarios in which the health of DI recipients is reassessed with a certain probability on a yearly basis. In the most extreme reform scenario, when the reassessment probability is set at 1, individuals have to pass an annual health test to remain in the program.

Figure 8.4 reports the aggregate effects on LFP and DI rate for individuals aged 50–64 when the reassessment probability is set at values ranging from 10 to 100%. The effects are large. DI rate is reduced by 3 percentage points (30.4%) when the reassessment probability is 10% and by 8 points (82%) when the DI recipients are assessed on a yearly basis. The reduction in DI does not entirely translate into an increase in LFP: when the reassessment has a probability of 10% LFP increases by 1.5 points (1.8%), the increase is of 4.6 points

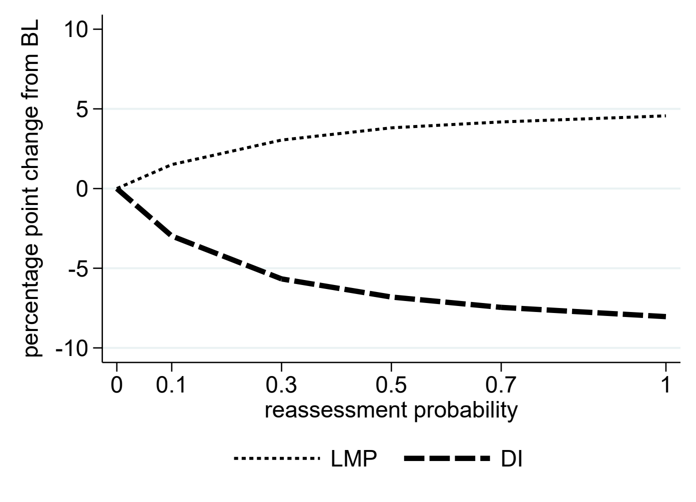


Figure 8.4: Effects on participation and DI rate of introducing an eligibility reassessment.

(5.5%) for the yearly reassessment.

In what follows, I focus on the reformed scenario in which the probability of being reassessed is 10% as it is likely to be more relevant due to its easier applicability with respect to a yearly reassessment of all DI recipients. The change in behavior concerns 52% of DI recipients in the baseline scenario and the reduction in DI participation is on overage of 5 years. The reformed scenario is characterized by an increasing number of DI episodes of smaller duration, for example the number of individuals with 2 or more episodes is more than 3 times bigger than in the baseline scenario. Figure 8.5 reports the life-cycle distribution of health (left panel) and assets (right panel) of those who change behavior in the reformed scenario compared to those who do not. As expected, those who exit DI due to the reform are on average in better health. They also have lower wealth in particular if aged 60–64, the age range in which reducing DI is more likely (45% are in this age group).[34]

Focusing on those receiving DI in the baseline scenario, I document a 30.6 (82.3) percentage points reduction in DI and a 16.4 (42.5) points increase in LFP. This suggests that only half of those who exit DI participate in the labor market. Among those remaining inactive, I find a modest increase in the receipt of means tested benefits, in particular a 3% (6%) increase in the receipt of Guarantee credit concentrated on those with lower levels of health and in the first asset quartile. This is consistent with observing lower levels of wealth among those reducing DI (see Figure 8.5). Those not in DI, not working and not receiving other benefits mainly rely on savings.

[34]Very similar patterns emerge when the reassessment probability is set at 100%, with lower levels of health for those who do not change behavior.

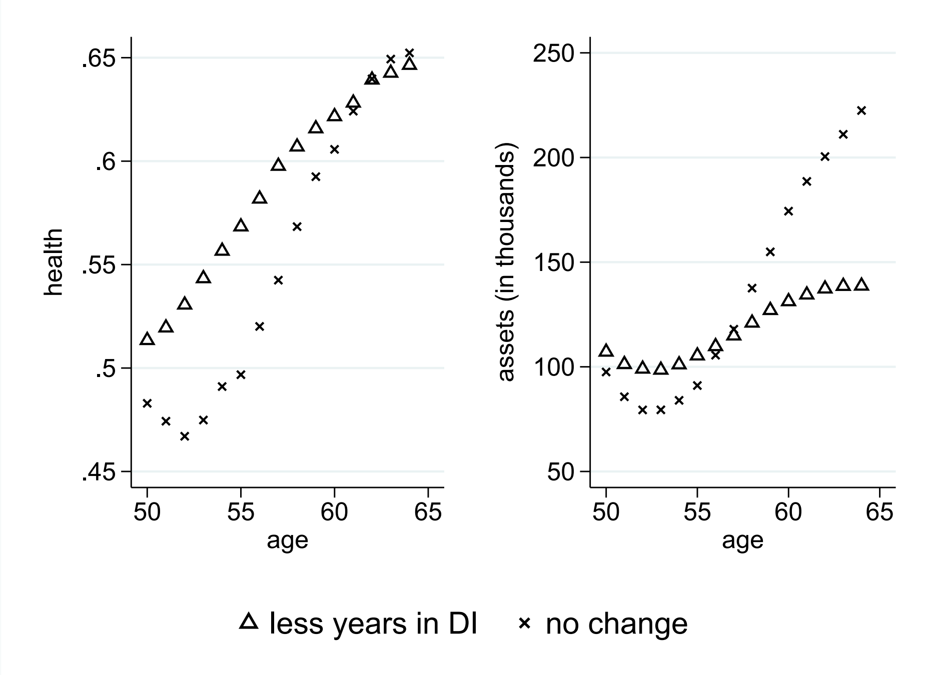


Figure 8.5: Patterns of health (left panel) and assets (right panel) for those who decrease or do not modify DI participation when the reassessment probability is set at 10%.

As already pointed out in the return-to-work policy experiment, the aggregate effects hide large variability by health, wealth and age that the model allows to exploit. In the overall sample, those most affected by the reform have health in the first three deciles, are in the first quartile of assets and are older. These categories are those who reduce DI rate the most, however the increase in LFP does not change much with age and it is larger for those at the margin of program entry, i.e. with health between the first and the second deciles. Additional insights can be gathered looking at individuals receiving DI in the baseline scenario. Table 8.3 shows that, consistently with Figure 8.5, the reduction in DI participation is increasing with health, it is 14 percentage points among those with health in the first decile and 49 points for those with health above the median, and the same is true for the increase in LFP (from 0 to 38 points). The decline in DI rate is increasing with age (from 22 to 36 points), whereas the increase in LFP picks among those aged 55–59 (22 points). The effect on DI rate is rather constant at 33 percentage points for wealth above the first quartile, and those in the second quartile respond the most in terms of LFP (24 points increase).[35]

Differently from the wage subsidy experiment, the aggregate effect on welfare is negative (see grey lines of Figure 8.6) and it amounts to -0.01% of consumption when the

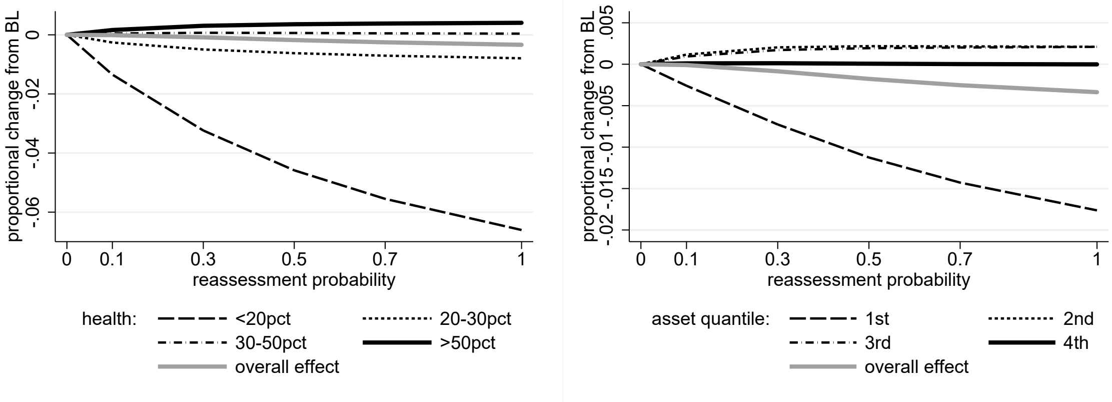
[35] Similar patterns of heterogeneity emerge when the reassessment probability is set at 100%.

Table 8.3: Proportional changes in DI rate and LFP conditional on health, wealth and age due to a policy reform that reassesses the health of DI beneficiary with a 10% probability each year.

	Δ DI rate		Δ LFP	
	all	DI in BL	all	DI in BL
	(1)	(2)	(3)	(4)
overall	-2.97	-30.63	1.50	16.42
health				
<10th pct	-12.66	-14.03	0.27	0.30
10-20th pct	-18.46	-38.19	9.38	19.47
20-30th pct	-8.83	-36.15	6.27	26.21
30-50th pct	-2.69	-47.05	1.77	34.01
>50th pct	-0.33	-49.37	0.19	37.94
assets				
1st quartile	-7.33	-29.39	2.95	14.29
2nd quartile	-1.87	-33.96	1.37	24.57
3rd quartile	-1.38	-34.76	0.86	21.45
4th quartile	-1.33	-29.84	0.84	13.92
age				
50-54	-1.86	-21.93	1.29	15.17
55-59	-3.00	-32.68	2.03	21.75
60-64	-4.15	-35.72	1.18	12.98

Notes: Statistics refer to individuals aged below 65 (all) and receiving DI in the baseline scenario (DI in BL).

reassessment involves 10% of the beneficiaries each year. However, panel (a) of Figure 8.6 reveals that the welfare effect is negative for levels of initial health below the third decile and particularly large for those with initial health below the second decile (-1.3% of consumption) and third decile (-0.3% of consumption), and positive for higher levels of health. The heterogeneity by initial wealth reveals negative effects for the first decile (-0.3%) and positive for higher quantiles (panel (b) of Figure 8.6). This highlights that the savings from a large reduction in DI participation induced by the policy reform (and the resulting redistribution as a reduction in the tax rates, see Table H.1 in Appendix H) do not compensate the reduced insurance experienced by those having lower levels of health (and wealth) at age 50 who are more exposed to the risk of DI episodes in the following years.



(a) by initial health level

(b) by intial asset level

Figure 8.6: Welfare effects by initial health and initial wealth level of introducing an eligibility reassessment.

8.3 Sensitivity

This section presents a number of robustness checks to assess results sensitivity to specific model assumptions. For each alternative model specification I report the estimated parameters in Table G.1 and the main results from the policy experiments in Tables G.2 and G.3 of online Appendix G.

The results are robust to setting the relative risk aversion coefficient to 3, a value closer to estimates in the literature (French, 2005; French and Jones, 2011). Estimation results are reported in column 2 of Table G.1 (column 1 reports the estimated parameters of the baseline specification, already shown in Table 6.2, for ease of comparability). Column 2 of table G.2 shows that the response to the financial return-to-work incentive is slightly smaller with about a half percentage point (14%) contraction in the reduction of DI rate for those receiving DI in the baseline or in the reformed scenarios. Among those who exit DI due to the health reassessment a larger fraction is active in the labor market with respect to the baseline model specification (from 54 to 58%). However, the main conclusions are preserved also in terms of welfare implications (see column 2 of Table G.3)

The exogenous health process is specified as the sum of a persistent and a transitory component. However, transitory shocks might capture also measurement error, at least partly. If the variance of the transitory component is overestimated this might affect the individual responses to policy changes, in particular when introducing the health reassessment probability. To check whether this is the case, I re-estimate the model reducing by half the variance of the transitory component (σ_η^2). The model fits the data equally well and estimation results are reported in column 3 of Table G.1. Also in this case the effect of the first policy experiment is smaller: the decrease in DI reduces by about 8%. The effect of the health reassessment is instead rather stable as it is the welfare effect. An increase in health persistence, given by the reduction in transitory fluctuations, might reduce the convenience to exit DI to go back to work. At the same time, given the symmetric nature of the transitory component the reduction in negative shocks is compensated by an equivalent reduction in the positive ones and this might explain why no changes are observed in the second policy scenario.

In the model specification in Section 4, I introduced a non-linearity in the fixed cost of work through the parameter ϕ_{P2} which shifts (downward) the intercept of the fixed cost of work for those working part-time from age 65 onward. This parameter allows to reproduce, at least partly, the drop in hours worked observed in the data at age 65 (SPA), which is due to a considerable fraction of individuals not exiting the labor market but reducing hours worked. This change in behavior is not driven by institutional incentives but it seems to be driven by preferences (ONS, 2015). In column 4 of Table G.2 I report parameter estimates when the fixed cost of work is not allowed to vary at age 65 ($\phi_{P2} = \phi_{P0}$). The results are broadly confirmed, both in terms of DI and LFP participation, and in terms of welfare consequences.

Finally, the model predicts a very large rejection probability for individuals in bad health, this is likely to be partly driven by the assumption of no direct or indirect costs of applying for DI in the model resulting in a large fraction of individuals applying for DI. This model assumption is justified by the institutional setting in which both employed and unemployed individuals can apply for DI and in a few weeks after applications accepted applicants start to receive the benefit. However, it might be reasonable to assume that individuals applying for DI have to pay the cost of filling in the forms and participating in the different steps of the application process, or to experience the public response to their disability condition in the form of stigma and discrimination. To capture these potential costs, I introduce in the equation for leisure (see Equation 4.3) a time cost of applying for DI, ϕ_A. The calibration of this parameter is to some extent arbitrary, as I do not have information on the decision to apply for DI in ELSA nor from other external (administrative or survey) data sources. For this reason I experiment with three different values of ϕ_A, 50, 200 and 500 hours respectively. The introduction of application cost affects the rejection probability for the unhealthy only when the cost is set at high values: with ϕ_A set at 500 the probability shrinks from 68 to 63% for those with health in the first decile, and from 87 to 74% for those with health in the second and third deciles. The DI application elasticity to benefit generosity is substantially higher, suggesting that in presence of costs in DI application individuals are more responsive to changes in the benefit amounts.

The results from the policy experiments are broadly confirmed, in particular for what concerns the welfare consequences. The major differences are observed when considering the model specification with the highest application cost. I find that the reduction in DI participation induced by the first policy experiment shrinks by 36% among those receiving DI in the baseline or in the reformed scenarios. In the second policy experiment instead the fraction observed working after being moved out from the benefit due to the health reassessment increases by 5 percentage points. This suggests that having paid the cost of applying makes less attractive to leave the benefit and return to work (as a result of the wage subsidy), increases the benefit persistence and therefore, the likelihood of finding individuals in relative good health and with substantial working capacity among those in DI. See columns 5–7 of Tables G.1–G.3.

9 Conclusions

In this paper, I investigate the consequences of two alternative return-to-work policies—a wage subsidy and a continuous eligibility reassessment—on welfare, LFP, and DI enrollment using English data. To this end, I develop and estimate a life-cycle model of labor supply, DI application, and saving behavior. I model the decisions of men living with a partner and approaching retirement age who face uncertainty regarding wage realization, health, and life expectancy. The model is able to replicate quite well asset profiles, LFP and its heterogeneity by health level, 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 process.

Comparing the two alternative policies, I find that a wage subsidy provided to DI beneficiaries for the first year of work is promising, as it increases LFP by 4.6 percentage points and DI outflows by 5.7 percentage points without decreasing welfare. The continuous eligibility reassessment has a greater effect on LFP and DI enrollment but is welfare-decreasing, in particular among low-health individuals. I document that individuals in relatively better health are more responsive to work incentives, supporting the adequacy of the 2008 UK reform that introduced return-to-work policies only for DI claimants with less severe disabilities. Finally, I show that results are robust to numerous alternative model specifications.

My main conclusions are likely to be robust to specific changes in modeling assumptions. In the baseline model specification, there are no application costs, therefore the flows in are likely to be overestimated by the model. However, this effect is compensated by the absence in the model of a direct cost of receiving DI in terms of reduced labor market opportunities (the wage offer only depends on age and health and not on DI status in the previous period) which might produce an overestimation of the benefit outflows. The sensitivity analysis in Section 8.3, which introduced the application costs, shows that indeed the effect of the wage subsidy decreases with the increase in the size of the application costs, but remains sizable.

The results in this paper are limited to cohabiting men aged above 50. Couples tend to be wealthier, less at risk of under-saving, in better health, and have lower mortality rates than singles. Therefore, these findings should be extended with caution to single individuals. Another caveat is that 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 labor market and to claim DI (Blau, 1998; Blau and Gilleskie, 2006; Borella *et al.*, 2017). Finally, the model assumes that health is exogenous, and it does not account for the possibility that changes in the DI benefit and DI benefit receipt might induce individuals to modify their behaviors, thus affecting their health level and dynamic (Cole *et al.*, 2019) as well as their life expectancy (Gelber *et al.*, 2018; Black *et al.*, 2018). For example, forcing DI beneficiaries to work by increasing the work incentives or reducing the DI generosity might have direct negative effects on individuals' health. Moreover, if return-to-work policies substantially increase labor supply among DI claimants, this might affect wages and thus attenuate the LFP response. The investigation of heterogeneous effects by household type, the role of labor demand, 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.

Andrews, I., Gentzkow, M. and Shapiro, J. M. (2017) Measuring the Sensitivity of Parameter Estimates to Estimation Moments*, *The Quarterly Journal of Economics*, **132**, 1553–1592.

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.

Autor, D. H. and Duggan, M. G. (2003) The rise in the disability rolls and the decline in unemployment, *The Quarterly Journal of Economics*, **118**, 157–206.

Autor, D. H. and Duggan, M. G. (2006) The growth in the social security disability rolls: A fiscal crisis unfolding, *Journal of Economic Perspectives*, **20**, 71–96.

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. (2015a) 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. (2015b) Effect of pensions and disability benefits on retirement in the United Kingdom, in *Social Security Programs and Retirement Around the World: Disability Insurance Programs and Retirement* (Ed.) D. A. Wise, University of Chicago Press, pp. 81–136.

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

Black, B., French, E., McCauley, J. and Song, J. (2018) The effect of disability insurance receipt on mortality, 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, 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.

Borghans, L., Gielen, A. C. and Luttmer, E. F. P. (2014) Social support substitution and the earnings rebound: Evidence from a regression discontinuity in disability insurance reform, *American Economic Journal: Economic Policy*, **6**, 34–70.

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.

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.

Capatina, E. (2015) Life-cycle effects of health risk, *Journal of Monetary Economics*, **74**, 67–88.

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

Cole, H. L., Kim, S. and Krueger, D. (2019) Analyzing the effects of insuring health risks: On the trade-off between short-run insurance benefits vs. long-run incentive costs, *Review of Economic Studies*, **86**, 1123–1169.

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

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.

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.

Gelber, A., Moore, T. and Strand, A. (2018) Disability insurance income saves lives, Working Paper WP 18-005, Stanford institute for economic policy research.

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

Gruber, J. and Kubik, J. D. (1997) Disability insurance rejection rates and the labor supply of older workers, *Journal of Public Economics*, **64**, 1–23.

Hosseini, R., Kopecky, K. A. and Zhao, K. (2021) The evolution of health over the life cycle, *Review of Economic Dynamics*.

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., Palme, 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. 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.

Marie, O. and Vall Castello, J. (2012) Measuring the (income) effect of disability insurance generosity on labour market participation, *Journal of Public Economics*, **96**, 198 – 210.

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

Mullen, K. J. and Staubli, S. (2016) Disability benefit generosity and labor force withdrawal, *Journal of Public Economics*, **143**, 49 – 63.

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

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

ONS (2015) *Office for National Statistics: Participation Rates in the UK Labour Market*, available at `https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemployeetypes/compendium/participationratesintheuklabourmarket/2015-03-19`.

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.

Ruh, P. and Staubli, S. (2018) Financial incentives and earnings of disability insurance recipients: Evidence from a notch design, *American Economic Journal: Economic Policy*.

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, special Issue: The Role of Firms in Tax Systems.

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.

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

Zaresani, A. (2018) Return-to-work policies and labor supply in disability insurance programs, *AEA Papers and Proceedings*, **108**, 272–76.

Zaresani, A., Olivo-Villabrille, M. *et al.* (2021) Return-to-work policies’ clawback regime and labor supply in disability insurance programs, Tech. rep., Institute of Labor Economics (IZA).

Online Appendix:
Disability Insurance and the Effects of Return-to-Work policies

Chiara Dal Bianco

A Data

The English Longitudinal Study of Ageing (ELSA) is the English version of the Health and Retirement Study (HRS) for the US. It began in 2002 and it targets people aged 50 and over and their partners, living in private households in England. The original sample was drawn from households that had previously responded to the Health Survey for England (HSE) between 1998 and 2001 and then refreshed at several waves (waves 3, 4, 6, 7 and 9). The survey follows the same group of respondents over time and takes place every two years. The questionnaires are designed to measure the ageing process of individuals by looking at changes in health, economic and social circumstances. The survey provides data about household and individual demographics, health - physical and psychosocial, work and pensions, income and assets, housing, cognitive function, social participation, expectations, objective health measures such as walking speed and weight (collected every two waves).

A.1 The health measure

Table A.1 reports the list of conditions included in the health index and their averages for respondents age 50–64 and 65–90. Limitations with ADL include: dressing, including putting on shoes and socks; walking across a room; bathing and showering; eating, such as cutting up food; getting in and out of bed; using the toilet, including getting up and down. Limitation with Instrumental ADL include: using map to figure out how to get around strange place; preparing a hot meal; shopping for groceries; making telephone calls; taking medications; doing work around house and garden; managing money; e.g. paying bills, keeping track of expenses. The CES-D scale is based on questions about feelings experienced in the past weeK: (i) felt depressed, (ii) felt that everything was an effort, (iii) sleep was restless, (iv) were happy, (v) felt lonely, (vi) enjoyed life, (vii) felt sad, (viii) could not get going.

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

ELSA variables	**age<65**	**age≥65**
physical functions		
difficulty sitting 2 hours	0.11	0.11
difficulty getting up from chair	0.16	0.25
difficulty walking 100 yards	0.07	0.14
difficulty climbing several flights stairs	0.18	0.35
difficulty climbing one flight stairs	0.07	0.14
difficulty stooping, kneeling or crouching	0.23	0.37
difficulty reaching or extending arms	0.07	0.10
difficulty pulling or pushing large objects	0.08	0.14
difficulty lifting or carrying weights	0.09	0.17
difficulty picking up 5p coin from table	0.03	0.05
limitations with Activities of Daily Living (ADL) (from 0 to 5)	0.23	0.41
eyesight (from 1=excellent to 6=blind)	2.36	2.56
glaucoma	0.02	0.07
cataracts	0.04	0.21
other eye problems	0.02	0.05
hearing (from 1=excellent to 5=poor)	2.60	3.00
problem of incontinence	0.04	0.10
mental, cognitive and intellectual functions		
Any emotional, nervous or psychiatric problems	0.10	0.06
limitations with Instrumental ADL (from 0 to 5)	0.18	0.37
Depression (CESD scale from 1 to 8)	2.88	2.94
Diagnosed psychiatric problems	0.07	0.04
Diagnosed dementia, Alzheimer's disease or Parkinson's disease	0.01	0.03
diagnosed conditions		
High blood pressure or hypertension	0.31	0.42
A stroke (cerebral vascular disease)	0.02	0.07
Diabetes or high blood sugar	0.09	0.14
Chronic lung disease	0.03	0.07
Asthma	0.09	0.09
Arthritis	0.21	0.32
Osteoporosis	0.01	0.02
Cancer	0.02	0.07
Angina	0.05	0.12
Myocardial infarction	0.04	0.11
Heart failure	<0.01	0.01
Heart murmur	0.02	0.04
Abnormal heart rhythm	0.05	0.10

Note: Unless differently specified the variables are binary. The table reports averages computed on male respondents living with a partner interviewed in wave 1 to 6 of ELSA (11,317 aged 50-64 and 10,316 aged 65-90).

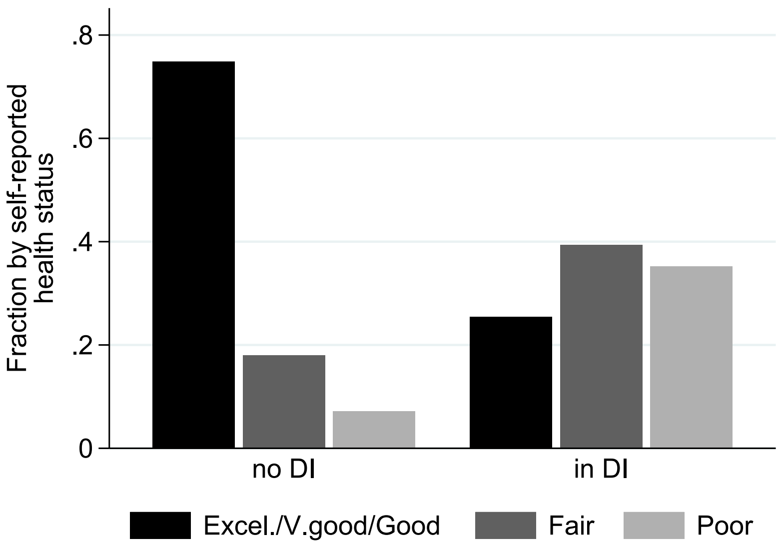


Figure A.1: Self-reported health distribution by DI status.

A.2 Additional tables and figures

Figure A.2: Fraction receiving DI by cohort.

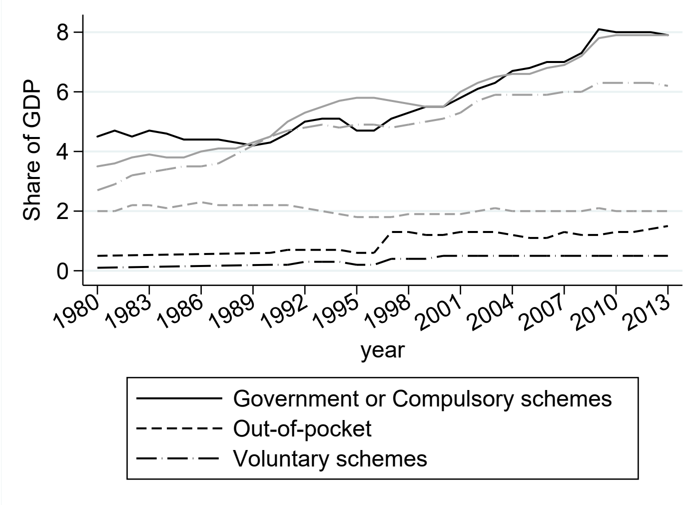


Figure A.3: Health care expenditure as share of GDP divided by government/compulsory schemes, household out-of-pocket payments, voluntary healthcare payment schemes. Black lines = UK. Grey lines = US. Source: OECD statistics.

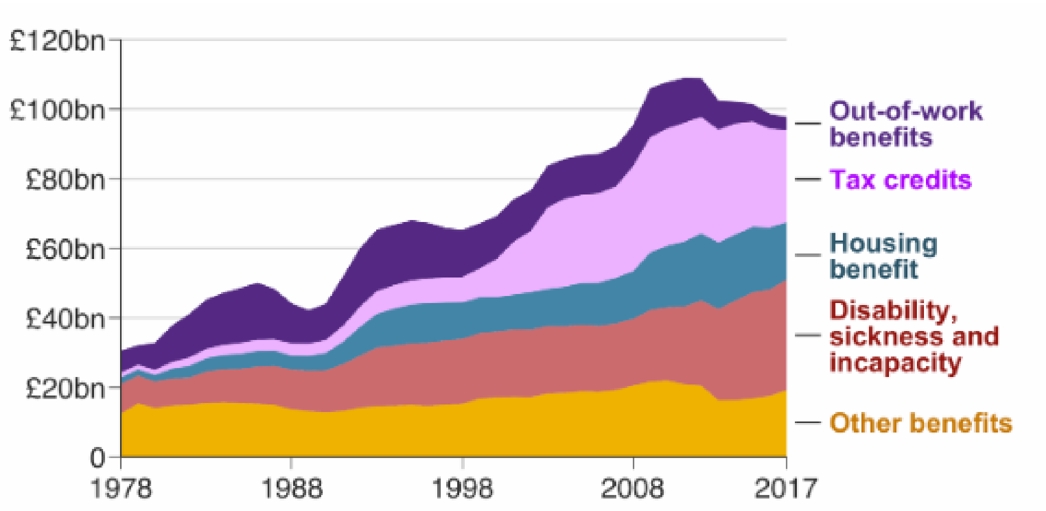


Figure A.4: Breakdown of working-age benefits spending over time

Notes: 'Other benefits' includes council tax benefit and minor housing-related benefits. Source: Institute for Fiscal Studies (`https://ifs.org.uk/publications/14011`).

B Exogenous processes

B.1 Health process

The three parameters of the random component of the health process ($\sigma^2_{\nu_H}$, σ^2_η and ρ_H) plus the initial variance at age 50 (σ^2_0) are identified by the variances and covariances of the adjusted residuals, $\tilde{\zeta}^H_{it}$. The initial period variance ($t = 0$), the following periods variances

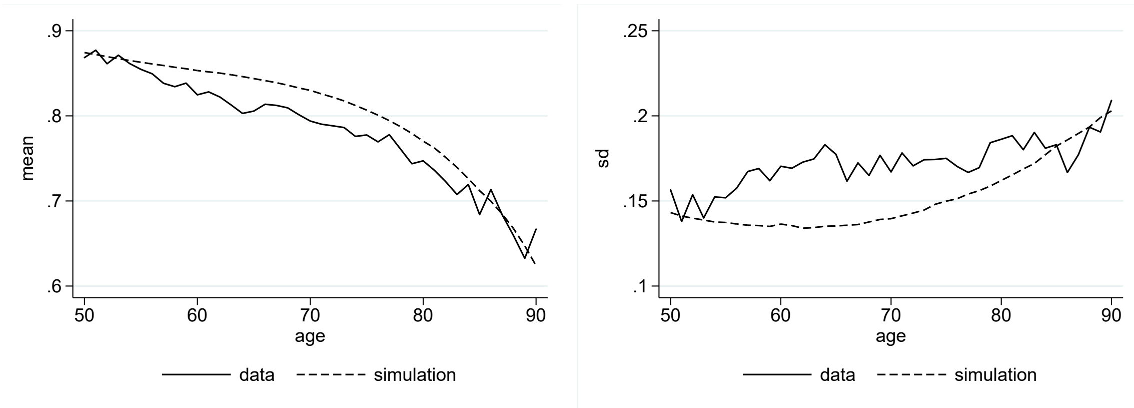
($t = 1, \ldots, T$) and the lag ℓ covariances are equal to

$$\begin{aligned}
Var(\tilde{\zeta}^{H}_{i0}) &= \sigma_0^2 + \sigma_\eta^2 \\
Var(\tilde{\zeta^{H}}_{it}) &= \rho^{2t}\sigma_0^2 + \frac{1-\rho^{2t}}{1+\rho^2}\sigma_{\nu_H}^2 + \sigma_\eta^2 \\
E(\tilde{\zeta^{H}}_{it}\tilde{\zeta^{H}}_{it-\ell}) &= \rho^{\ell}\left(\rho^{2(t-\ell)}\sigma_0^2 + \frac{1-\rho^{2(t-\ell)}}{1+\rho^2}\sigma_{\nu_H}^2\right)
\end{aligned}$$

Where ρ is ρ_H according to the main text notation. Identification requires to have at least three periods of data. Note that, given the biennial nature of ELSA data I consider lags ℓ that are multiple of 2 up to lag 8.

The mean and standard deviation of health over age in the data and in the simulations are shown in Figure B.1. Figure B.2 reports the fit of specific moments of the health distribution by age. Simulations are obtained setting the fixed effect equal to the average fixed effect for the cohort of interest (1946-55), family size to two and the unemployment rate at 4.9%, which is the 2004 annual unemployment rate for men in England.

Table B.1 shows the variances and covariances of the adjusted error term in the data and in the simulations.



(a) Mean

(b) Standard deviation

Figure B.1: Model fit for the mean and standard deviation of health by age.

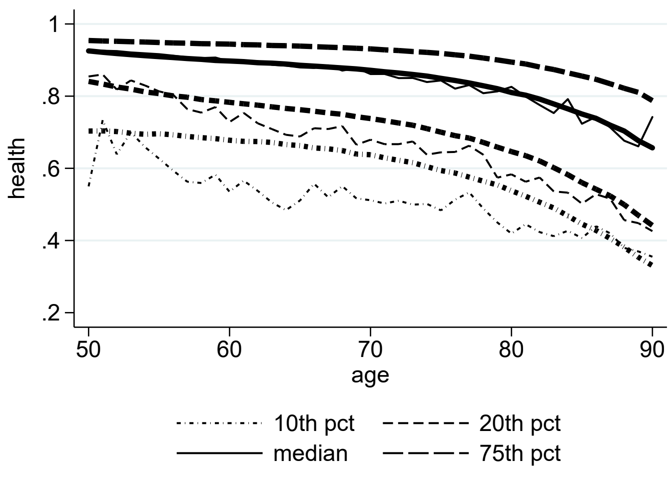


Figure B.2: Model fit for selected percentiles of the health distribution by age.

Table B.1: Variance-covariance matrix of the adjusted residuals. Data vs simulations.

						Age					
Age	50-51	52-53	54-55	56-57	58-59	60-61	62-63	64-65	66-67	68-69	70-71
						Data					
50-51	0.987										
52-53	0.761	1.089									
54-55	0.695	0.790	1.053								
56-57	0.775	0.822	0.934	1.158							
58-59	0.802	0.733	0.894	0.930	1.144						
60-61	0.610	0.730	0.794	0.906	0.955	1.172					
62-63	-	0.762	0.705	0.741	0.860	0.933	1.183				
64-65	-	-	0.313	0.752	0.852	0.948	0.947	1.2290			
66-67	-	-	-	0.800	0.800	0.843	0.879	0.926	1.119		
68-69	-	-	-	-	0.727	0.752	0.791	0.856	0.884	1.117	
70-71	-	-	-	-	-	0.646	0.855	0.870	0.894	0.939	1.117
						Simulations					
50-51	1.02										
52-53	0.809	1.036									
54-55	0.759	0.824	1.055								
56-57	0.710	0.771	0.840	1.064							
58-59	0.665	0.722	0.783	0.848	1.072						
60-61	0.624	0.677	0.735	0.795	0.856	1.082					
62-63	-	0.632	0.688	0.745	0.803	0.866	1.090				
64-65	-	-	0.644	0.696	0.750	0.810	0.871	1.104			
66-67	-	-	-	0.654	0.704	0.759	0.819	0.884	1.107		
68-69	-	-	-	-	0.658	0.708	0.764	0.826	0.887	1.108	
70-71	-	-	-	-	-	0.665	0.716	0.780	0.836	0.896	1.108

Assuming an AR(1) process for health and estimating the process at the mean implies that the persistence of health is constant. This means that individuals in very bad health experience the same persistence of individuals in good health. Figure B.3 plots the transition matrix implied by the data when the health index is discretized in ten deciles. The level of persistence (the values on the diagonal) does not seem to vary remarkably with health; moreover, the process is not far from being symmetric as the probability of a health improvement is almost equal to the probability of a health drop.

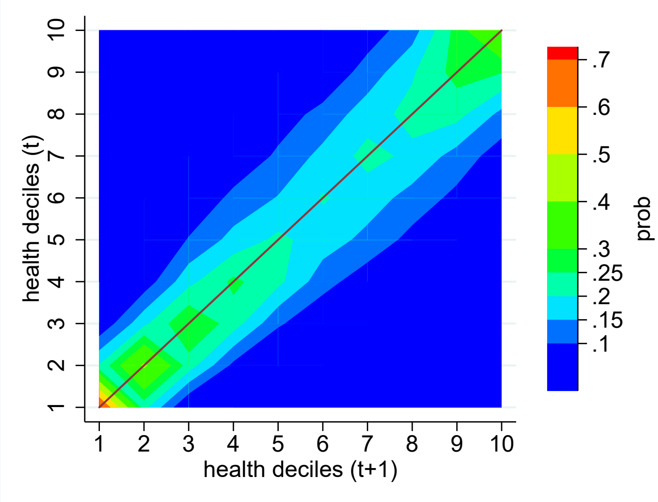


Figure B.3: Transition probability from t to $t + 1$ for the health deciles.

I checked for the robustness of the health measure to the inclusion of self-reported health among the health indicators used in the principal component. Table B.2 reports the estimation of the deterministic and stochastic components of the health process when self-reported general health is included. To ease the comparison, the table also reports the estimates of the continuous index used in the analyses and that do not include self-reported general health (the values are those already shown in Table 6.1 in the main text). The parameters of both the deterministic and the stochastic component are quantitatively the same in the two specifications of the index.

Table B.2: Parameters of the health process when self-reported health (SRH) is used to generate the health index and when it is excluded.

Deterministic component			Random component		
	no SRH	SRH		no SRH	SRH
age	-0.503***	-.463***	ρ	0.967***	0.967***
	(0.126)	(0.125)		(0.011)	(0.011)
$age^2/10$	0.081***	0.0749***	σ_ν^2	0.063**	0.063**
	(0.019)	(0.019)		(0.021)	(0.022)
$age^3/100$	-0.005***	-0.004	σ_η^2	0.158***	0.158***
	(0.001)	(0.001)		(0.028)	(0.028)
			σ_0^2	0.865***	0.886***
				(0.056)	(0.058)
Observations	19,034	19,029			

Even if the interest is on individual decisions up to age 70, I include individuals aged above 70 to estimate the health process because the model is solved and simulated up to age 90. Table B.3 shows that the estimated parameters of the stochastic component are quantitatively the same if I use only individuals aged 50 to 70.

Table B.3: Parameter estimates of the health process stochastic component. In the first column individuals aged 50 to 90 are included in the estimation sample (baseline specification), in the second column only those up to age 70 are included.

Random component		
	50-90	50-70
ρ	0.967	0.976
	(0.011)	(0.013)
σ_ν^2	0.063	0.055
	(0.021)	(0.023)
σ_η^2	0.158	0.172
	(0.028)	(0.029)
σ_0^2	0.865	0.835
	(0.056)	(0.058)

B.2 Wage process

To estimate the wage process and control for selection into participation, I closely follow Low and Pistaferri (2015). I report below the main steps of the procedure. I write the wage equation and the LFP equation as

$$\begin{aligned}\log w_{it} &= f_i + \omega_w(H_{it}, age_{it}) + \epsilon_{it} + \xi_{it}\\ \epsilon_{it} &= \epsilon_{it-1} + \nu^w_{it}, \qquad \xi_{it} \sim N(0, \sigma^2_\xi)\\ P^*_{it} &= \omega_P(H_{it}, age_{it}) + \psi G_{it} + \phi_{it}\\ &= p_{it} + \phi_{it}\end{aligned}$$

where G_{it} is the vector of the exclusion restrictions and $P_{it} = 1$ if $P^*_{it} > 0$. To get rid of the unobserved heterogeneity f_i, I rewrite the wage equation using data in differences, with s denoting a generic lag with $s \geq 1$.

$$\begin{aligned}\Delta^s \log w_{it} &= \omega_w(\Delta^s H_{it}, \Delta^s age_{it}) + \Delta^s \epsilon_{it} + \Delta^s \xi_{it}\\ &= \omega_w(\Delta^s H_{it}, \Delta^s age_{it}) + \sum_{j=0}^{s-1} \nu^w_{it-j} + \Delta^s \xi_{it}\end{aligned}$$

I observe wage growth only for individuals working in both t and $t - s$, therefore the conditional expectation takes the following form:

$$\begin{aligned}E(\Delta^s \log w_{it}|P_{it} = P_{it-s} = 1) &= \omega_w(\Delta^s H_{it}, \Delta^s age_{it}) + E\left(\sum_{j=0}^{s-1} \nu^w_{it-j}|P_{it} = P_{it-s} = 1\right)\\ &= \omega_w(\Delta^s H_{it}, \Delta^s age_{it}) + E\left(\sum_{j=0}^{s-1} \nu^w_{it-j}|\phi_{it} > -p_{it}, \phi_{it-s} > -p_{it-s}\right)\end{aligned}$$

where $P_{it} = 1$ if $h_{it} > 0$ and zero otherwise. Assuming $(\phi_{it}\phi_{it-s})' \sim N(0, I)$, the conditional expectation can be written as:

$$\begin{aligned}E(\Delta^s \log w_{it}|P_{it} = P_{it-s} = 1) = \omega_w(\Delta^s H_{it}, \Delta^s age_{it}) +&\\ \left(\sigma_{\nu^w} \sum_{j=0}^{s-1} \rho_{\nu^w_{t-j}\phi_t}\right) \lambda_{it} + \left(\sigma_{\nu^w} \sum_{j=0}^{s-1} \rho_{\nu_{t-j}\phi_{t-s}}\right) \lambda_{it-s}&\end{aligned}$$

where λ_{it} is the inverse Mills' ratio, $\sigma^2_{\nu^w}$ is the variance of ν^w_{it}, $\rho_{\nu k \phi_\ell}$ is the correlation between νik and $\phi_{i\ell}$. The regression of the wage growth on the controls in differences and the inverse Mills' ratios for each lag s allows to consistently estimate the parameters of the wage process.

In particular, I assume that financial incentives to participate in the labor market and the family structure serve as exclusion restrictions (G). To measure financial incentives, I use household financial constraints, that influence the labor market attachment, and institutional characteristics, that affect the labor supply decision. The former include having a mortgage loan; the latter comprise 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 it. The family structure variables capture family needs that affect participation decision and the couples' preferences for joint retirement, for example to spend time together. The variables considered are having a partner, having children, partner's health, whether the partner is above state pension age and whether the partner is above minimum age to withdraw from the private pension plans.

Table B.4: Estimation of the deterministic component's parameters of the wage process.

Parameter	Employment (1)	Wage growth (2)
Deterministic component		
age	0.437***	0.050
	(0.100)	(0.054)
$age^2/10$	-0.045***	-0.065
	(0.009)	(0.047)
$health$	10.054***	0.303
	(1.308)	(1.080)
$health^2$	-4.158***	-0.140
	(0.888)	(0.633)
p-value exclusion restrictions	0.000	
p-value selection correction terms		0.121
Random component		
$\sigma^2_{\nu^w}$		0.015***
		(0.004)
σ^2_{ξ}		0.041***
	(0.005)	
Observations	12,110	3,079

The first column of Table B.4 reports probit parameter estimates for the selection equation. Additional controls included are having a partner and time fixed effects. The exclusion restrictions are jointly significant (p-value 0.000). The probit selection equation allows to construct the inverse Mills's ratio to be included in the wage growth equation to account for selection bias. Estimates of the wage equation are shown in the second column of Table

B.4. Given the low number of observations, estimates are not very precise. The wage offer decreases with age and it increases with the level of health.

Finally, the following set of moment conditions on the adjusted error term are used to identify the parameters of the random component of the wage process.

$$E(\Delta^s(\epsilon_{it}+\xi_{it})|\phi_{it}>-p_{it},\phi_{it-s}>-p_{it-s}) = \sigma_{\nu^w}\lambda_{it}\sum_{j=0}^{s-1}\rho_{\nu^w_{t-j}\phi_t} + \sigma_{\nu^w}\lambda_{it-s}\sum_{j=0}^{s-1}\rho_{\nu_{t-j}\phi_{t-s}}$$

$$E(\Delta^s(\epsilon_{it}+\xi_{it})^2|\phi_{it}>-p_{it},\phi_{it-s}>-p_{it-s}) = \sigma^2_{\nu^w}\left(s - p_{it}\lambda_{it}\sum_{j=0}^{s-1}\rho_{\nu^w_{t-j}\phi_t} - p_{it-s}\lambda_{it-s}\sum_{j=0}^{s-1}\rho_{\nu_{t-j}\phi_{t-s}}\right) + 2\sigma^2_\xi$$

$$E(\Delta^s(\epsilon_{it}+\xi_{it})\Delta^\ell(\epsilon_{it-s}+\xi_{it-s})) = -\sigma^2_\xi$$

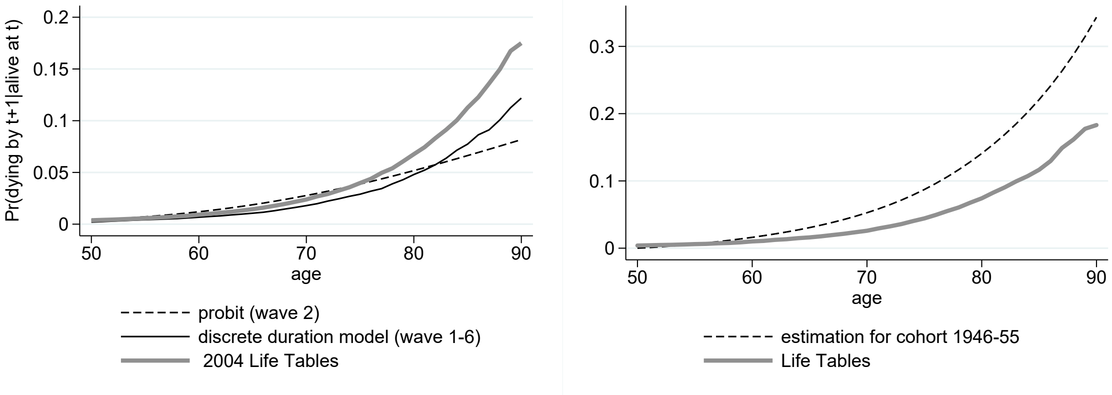
Estimates are reported in the bottom panel of Table B.4. Note that, given the biennial nature of ELSA data, the typical lag (difference in the interview years) between consecutive waves is 2 (70%). However, due to several reasons such as fieldwork length, interview schedule, possible work interruptions and no participation of the respondent in a particular wave, s can take value 1 (13%) or values greater than 2 (18%), most commonly 3 (15%).

Given that the fixed effect cancelled out when computing the adjusted residuals g_{it}, I recover the fixed effect $\hat{f}_i$ for those observed working and set the fixed effect for each simulated individual equal to the average fixed effect for those born between 1946 and 1955, i.e. the reference cohort for the model estimation.

B.3 Survival probability

In the model, I use mortality rates conditional on health for men born between 1946 and 1955. To derive them, I use ELSA data linked to administrative death records which allow to know the exact year of death of any individual (including attriters) up until February 2013 (this information is available in the public data release of ELSA).

Figure B.4a shows cross-sectional mortality rates from the life tables (for year 2004) and compares it with mortality rates computed on ELSA data. The solid black curve is obtained by estimating a discrete time duration model (complementary log-log hazard function) on a yearly panel of male individuals interviewed between wave 1 and wave 6. The model controls for age and uses log(time) as baseline hazard function. The dashed black curve shows predictions from a probit regression of the annual death probability on a second order polynomial in age for individuals observed in the second wave of ELSA, collected in 2004, to be consistent with the choice of using 2004 life tables. This comparison reveals that mortality rates computed with ELSA data are lower than comparable mortality rates from the life tables (for year 2004). In the derivation of survival probabilities conditional on health level I assume that the mortality risk perceived by individuals is consistent with the life tables and re-scale mortality in each health-age group in order to match the life tables. I make this assumption to be sure that mortality rates are corrected for selection into participation and attrition problems and I take from the data the heterogeneity in mortality rates by health.



(a) **(b)**

Figure B.4: Mortality rates in the life tables and in the data. Panel (a) reports mortality rates for ELSA sample. Panel (b) reports predicted mortality rates for cohort 1946-55.

Data are biennial but, given the linkage with administrative death records, I construct a dummy variable taking value one if the individual dies by next year and zero otherwise. This allows me to estimate the probability of dying by $t+1$ conditional on being alive in t, for each age and health level in t. I do it estimating a fixed effect regression and controlling for a four-grade polynomial in age interacted with four discrete health categories.[36]

Details on the procedure used to derive mortality rates are provided below.

1. 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 observations for male respondents. To control for cohort and family size effects, I estimate these probabilities using fixed-effect regressions. When I predict from the estimated regressions, I set family size to two and the fixed effect equal to the average fixed effect for those born between 1946 and 1955 (the cohort of interest);

[36] I discretize health in four categories, below the 20th percentile, between the 20th and the 30th percentiles, between the 30th percentile and the median and above the median.

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

$$\hat{Pr}(death^{D}_{t+1}) = \sum_{i=1}^{4} \hat{Pr}(H_t = i) * \hat{Pr}(death^{D}_{t+1}|H_t = i); \tag{B.1}$$

   Figure B.4b shows the probability computed according to equation B.1 compared to the life tables. Note that the in-sample predictions for the cohort of interest (1946-55) lay below the life tables (not shown). However, using the individual specific fixed effect I can only compute mortality rates up to age 67. To predict out of sample and account for cohort effects, I use the average fixed effect for the 1946-55 cohort, which affects the intercept of the curve, and the parameters of the age polynomials which affect the shape. The resulting curve, shown in Figure B.4b, over predicts the mortality rates compared to the cross-sectional life tables starting from age 56.

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

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

4. 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 $\hat{Pr}(death^{C}_{t+1}|H_t = i) = \alpha_t * \hat{Pr}(death^{D}_{t+1}|H_t = i)$.

Figure B.5 shows the generated mortality curves by health and the corresponding simulated curves. Note that the remarkably good fit suggests that also health fits well the data as shown in Subsection B.1.

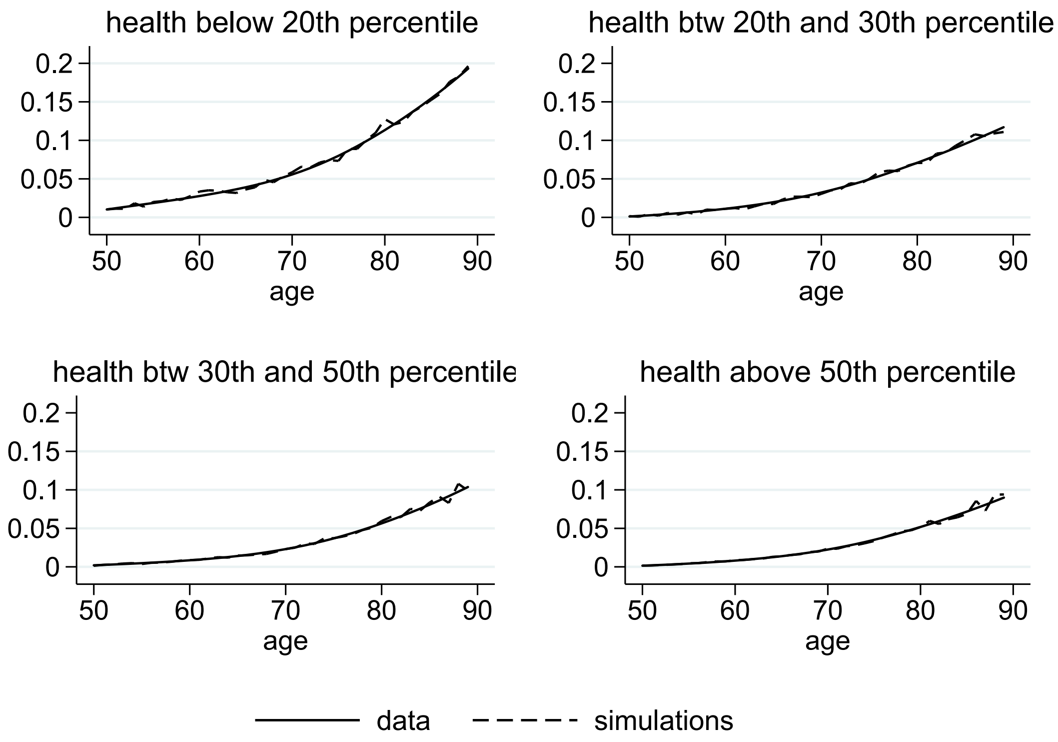


Figure B.5: Mortality rates by health level. Data vs simulations.

C Value function and model solution

In the model when individuals apply for DI, at the same time, they decide whether they would work or not if they were rejected. If the individual applies for DI, we have two different possible choices: he applies and work if rejected or he applies and remains out of the labor market if rejected. In the first case the value function becomes $V^{app} = (p_{allowed})V^{DI} + (1-p_{allowed})V^{work}$, in the second case $V^{app} = (p_{allowed})V^{DI} + (1-p_{allowed})V^{inactive}$, where $p_{allowed}$ denotes the probability of being allowed the benefit. As an example, when $h_t > 0$, $dapp_t = 0$ and age is below SPA the value function is the following:

$$V^{work}(X_t) = \max_{a_{t+1},h_t} \left\{ U(c_t, l_t) + \beta\pi^s_{t+1} \iint_{\substack{H_{t+1},\\ w_{t+1}}} V(X_{t+1}|X_t)dF(X_{t+1}|X_t) + \beta(1-\pi^s_{t+1})b(a_{t+1}) \right\} \quad \text{with} \quad l_t = L - h_t - \phi_H(\hat{H} - H_t) - \phi_{P_t}, \tag{C.1}$$

such that $a_{t+1} = a_t + y((1-c_w)w_t h_t, ra_t); \tau) + ys_t + tr_t - c_t$. In solving the model, the maximization of life-time utility is with respect to savings a_{t+1}, which is equivalent to maximize with respect to consumption c_t. Individuals are assumed to contribute a fixed fraction of their salary, c_w, to the pension fund up to SPA when they start to receive the pension. From age 70 onward there is no uncertainty on future wages but only on future health, because individuals are assumed to exit the labor market by that age.

I denote the vector of preference parameters with ϑ:

$$\vartheta = \{\beta, \nu, \gamma, L, \phi_H, \phi_{P0}, \phi_{P1}, \phi_{P2}, \alpha_1, \alpha_2, \alpha_3, \phi_B, K\}$$

and the parameters that determine the data generating process for the state variables with χ:

$$\chi = \left\{ r, \omega_H(age_t), \sigma_{\nu_H}, \sigma_\eta, \rho_H, \omega_w(H_t, age_t), \sigma_{\nu_w}, \{ys_t, \pi_t^s\}_{t=1}^T \right\}.$$

The state variables are discretized into a finite number of points on a grid, and the value function is evaluated at each point of the state space. I take 50 grid points for assets, 5 for pension wealth, 5 for the wage shock, 6 for the persistent health component and 5 for the transitory health component, additionally I have 2 states for DI claiming status for a total of 75,000 grid points. The grid for assets is more finely discretized for lower values of assets.

I take expectations with respect to shocks in health and wages and with respect to mortality risk. To capture uncertainty over the persistent components of health and wages, I convert θ_t and ϵ_t into discrete Markov chains.

D Calibration of K and ϕ_B

To calibrate the bequest parameters (K and ϕ_B) I follow De Nardi *et al.* (2010) (see Appendix D of the Supplementary Material of their paper). I compute the marginal propensity to bequeath (mpb) out of an extra pound and the consumption value of wealth at which the bequest motive becomes operative. I do it for a person who starts period t with cash on hand x, has level of health $\dot{H}$ and dies the next period with probability one.

Then, I calibrate the curvature of the bequest function K and the bequest weight ϕ_B in such a way that the *mpb* is equal to 0.98 and the bequest motive becomes operative when $x > \pounds 8000$. These values are in line with those implied by the parameter estimates in French (2005). French (2005) estimates imply a *mpb* ranging from 0.97 to 0.98, and a value of x between \$8000 and \$16000 in 2002 values, corresponding to $\pounds 7300$–$\pounds 14700$ in 2004 values.

The maximization problem is

$$\max_e \frac{1}{1-\nu}(c^\gamma \ell^{1-\gamma})^{1-\nu} + \beta\phi_B \frac{(e+K)^{(1-\nu)\gamma}}{1-\nu} \tag{D.1}$$

$$\text{subject to } e = (1+r)(x-c)$$

where $\ell = L - \phi_H(\hat{H} - \dot{H})$ and e is the bequest left. I fix $\dot{H}$ equal to the worst level of health, i.e. $\hat{H} - \dot{H} = 1$. Solving the maximization problem we find the optimal e:

$$\begin{aligned} e &= \frac{x - Kf}{q+f} \\ \text{with } f &= \left((1+r)\frac{\beta\phi_B}{\ell^{(1-\gamma)(1-\nu)}}\right)^{\frac{1}{\gamma(1-\nu)-1}} \\ \text{and } q &= -\frac{1}{1+r} \end{aligned} \tag{D.2}$$

A person leaves a bequest e only if $x > Kf$. If x is large enough, the *mpb* out of extra pound today is

$$\frac{\delta}{\delta x}\left(\frac{e}{1+r}\right) = \frac{1}{(1+r)(q+f)}. \tag{D.3}$$

Equations D.2 and D.3 give the conditions to calibrate the curvature of the bequest function K and the bequest weight ϕ_B, respectively. The resulting values in the baseline specification are $\phi_B = 42.8$ and $K = 403,368$.

E Moment conditions and asymptotic distribution of parameter estimates

E.1 Moment conditions

The set of data moments used in the estimation of the structural parameters are obtained by estimating life-cycle profiles for assets, participation, hours worked, DI claiming rate, DI persistence, DI inflow and average health if in DI; the profiles are representative of the cohort of interest, those born in 1946-55. The procedure is similar to the one implemented in French (2005). To increase sample size, I use information on both singles and couples. Taking as an example hours profile, I regress log hours, $log(h_{it})$, on an individual specific effect f_i, a polynomial in age, 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 LFP as the outcome, age polynomials are interacted with health categories.

$$\log(h_{it}) = f_i + \sum_{s=1}^{S} \pi_s age_{it}^s + \sum_{k=1}^{K} \pi_k 1\left\{size_{it} = k\right\} + \pi_C couple_{it} + \pi_U U_t + u_{it} \tag{E.1}$$

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 effects, time effects and family size effects.

The estimated fixed effects $\hat{f}_i$ are regressed on a set of ten-year cohort dummies and the $couple$ dummy, this allows to compute the conditional expectation of $\hat{f}_i$ for a specific cohort of individuals cohabiting with a partner, $E\left[\hat{f}_i|cohort = c, couple = 1\right]$. I then simulate from the estimated model fixing family size at two, unemployment rate at 4.9% and the individual fixed effect with the average fixed effect for the cohort of interest. Specifically, I replace f_i with $\tilde{f}_i = \hat{f}_i - E\left[\hat{f}_i|cohort_i, couple_i\right] + E\left[\hat{f}_i|cohort = c, couple = 1\right]$. 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.[37]

The life-cycle profile of assets, including both housing and non-housing wealth, estimated using data between 2002 and 2008 might be largely influenced by the rapid house price increase in that period. 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%. If the assets profile used to estimate the model's parameters 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. To account for this, I assume that the house price increase and the resulting wealth increase for homeowners do not affect individual decisions in terms of consumption, retirement and LFP. Therefore, I strip out

[37]The estimated data profiles are shown in Section 7 together with the profiles implied by the estimated model.

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 profile.

I use waves 1–6 for the assets profile, whereas I restrict the data to the first four waves of ELSA, a period in which DI policies and parameters were relatively stable to estimate the remaining moment conditions.

E.2 Parameter estimation

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

$$\sqrt{I}(\hat{\vartheta} - \vartheta_0) \sim N(0, \mathbf{V}) \tag{E.2}$$

with variance-covariance matrix $\mathbf{V} = (1 + \tau)(\mathbf{D'WD})^{-1}\mathbf{D'WSWD}(\mathbf{D'WD})^{-1}$, where $\mathbf{S}$ is the variance-covariance matrix of the data, $\mathbf{D}$ is the Jacobian matrix of the population moment vector (Equation E.3) and $\mathbf{W}$ the plim of the weighting matrix $\hat{W}$.

$$D = \left.\frac{\partial \theta(\vartheta, \chi_0)}{\partial \vartheta'}\right|_{\vartheta=\vartheta_0} \tag{E.3}$$

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 $\mathbf{S}$ and the weighting matrix $\mathbf{W}$ are estimated with their sample analogue. The partial derivatives in the Jacobian matrix $\mathbf{D}$ are straightforward to compute by taking the numerical derivatives of $\hat{\theta}_I(.)$.

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

$$\frac{I}{1+\tau}\hat{\theta}(\vartheta, \chi_0)'\mathbf{R}^{-1}\hat{\theta}(\vartheta, \chi_0) \sim \chi^2_{N-8} \tag{E.4}$$

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

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

I report below the measure of sensitivity proposed by Andrews *et al.* (2017) that allows, among other things, to formally show how perturbations of moments conditions affect parameter estimates. The sensitivity measure called Λ, is defined as $\Lambda = (\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}$. To ease the interpretation, I report the matrix content graphically.[38] Figures E.1–E.8 report the product between each element Λ_{pj} of Λ and the standard deviation of the j^{th} moment, such that the values reported represent the effect of one-standard-deviation change in the moment on each parameter.

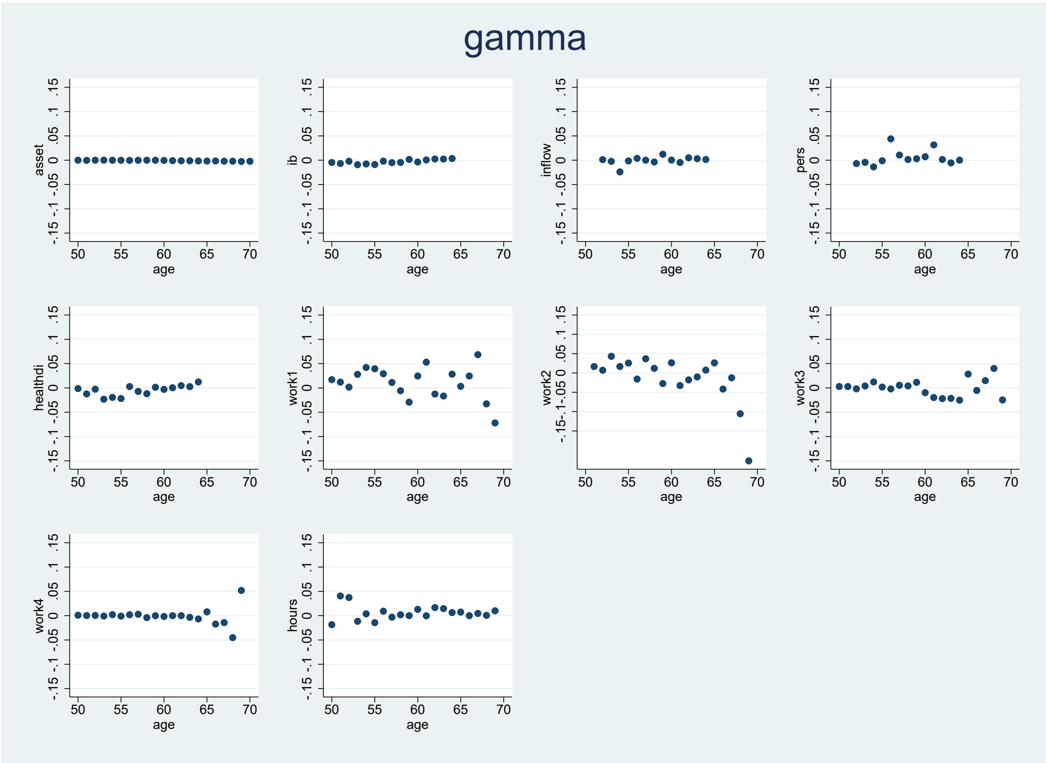


Figure E.1: Effect of one-standard-deviation change in the targeted moments on relative risk aversion parameter, γ.

[38] The full matrix is available upon request.

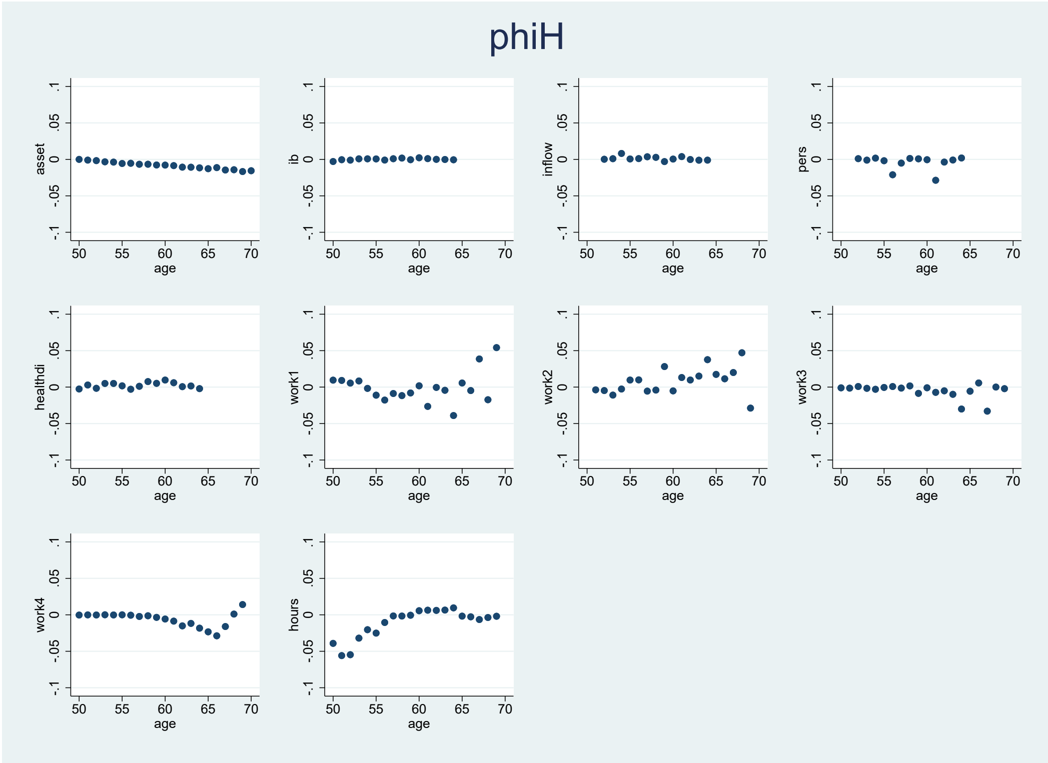


Figure E.2: Effect of one-standard-deviation change in the targeted moments on the cost of being in bad health, ϕ_H.

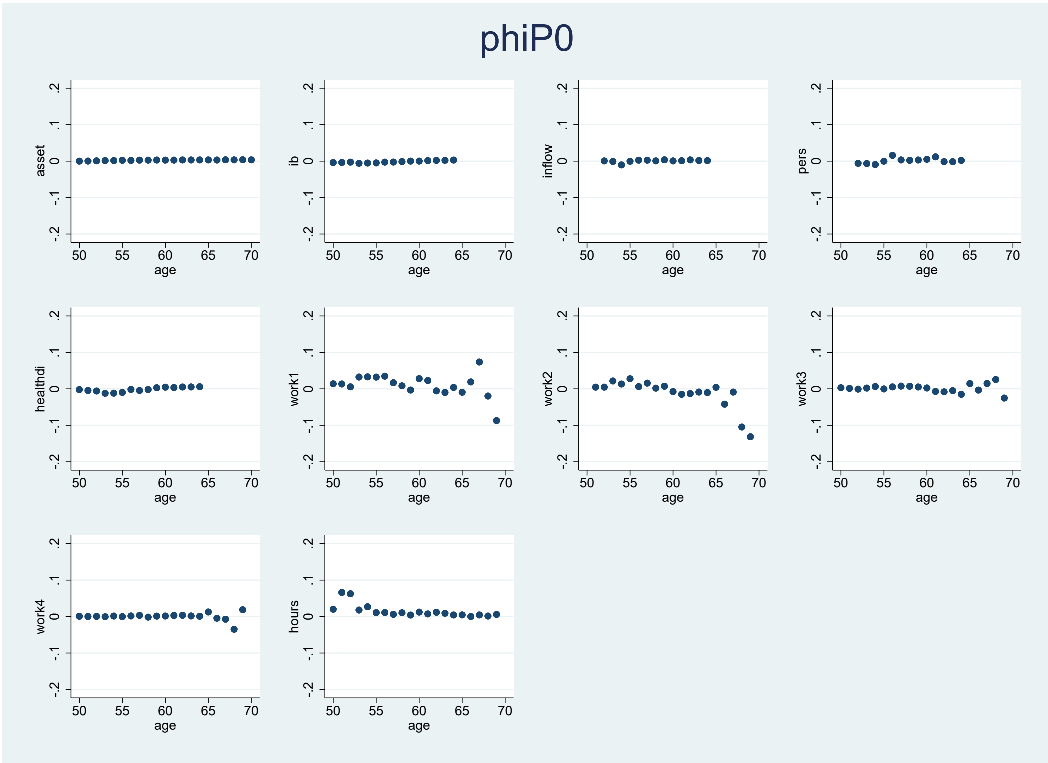


Figure E.3: Effect of one-standard-deviation change in the targeted moments on the intercept of the cost of work, ϕ_{P0}.

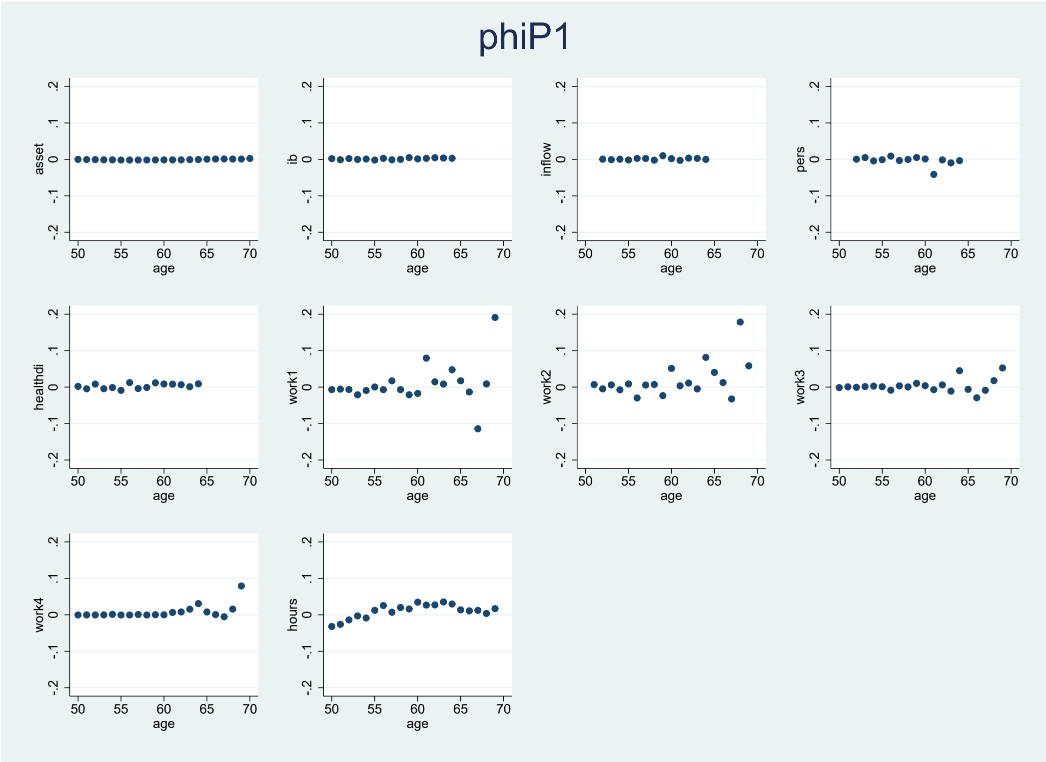


Figure E.4: Effect of one-standard-deviation change in the targeted moments on the slope of the cost of work, ϕ_{P1}.

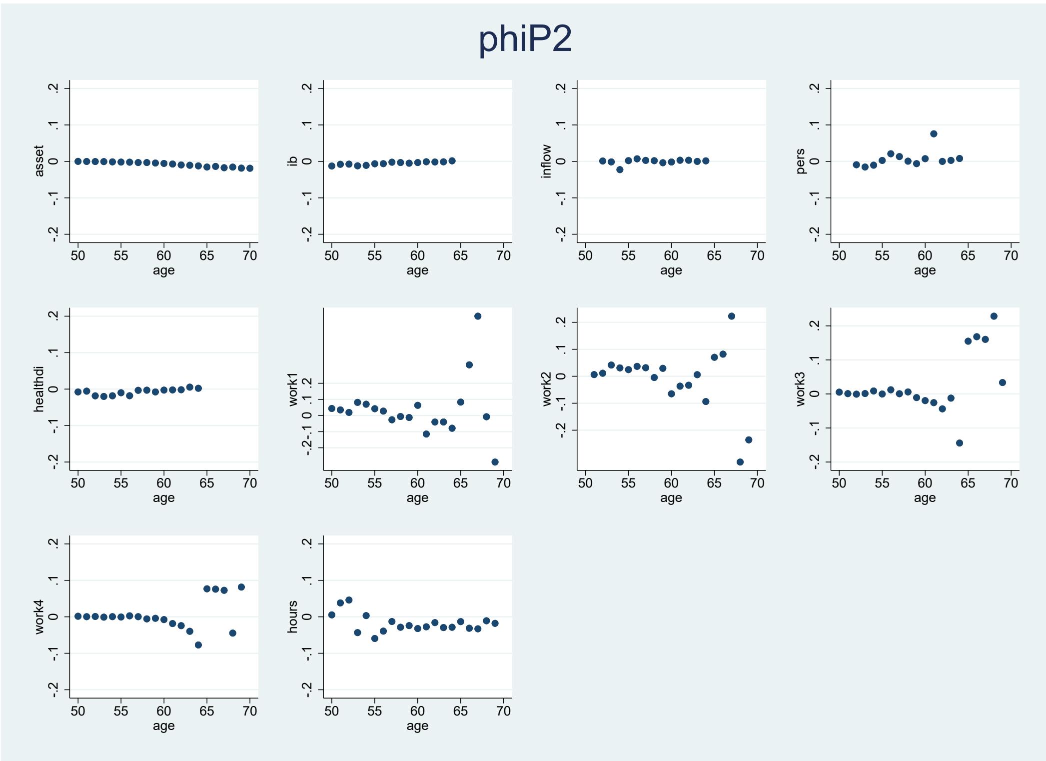


Figure E.5: Effect of one-standard-deviation change in the targeted moments on the intercept of the cost of work form age 65 onward, ϕ_{P0}.

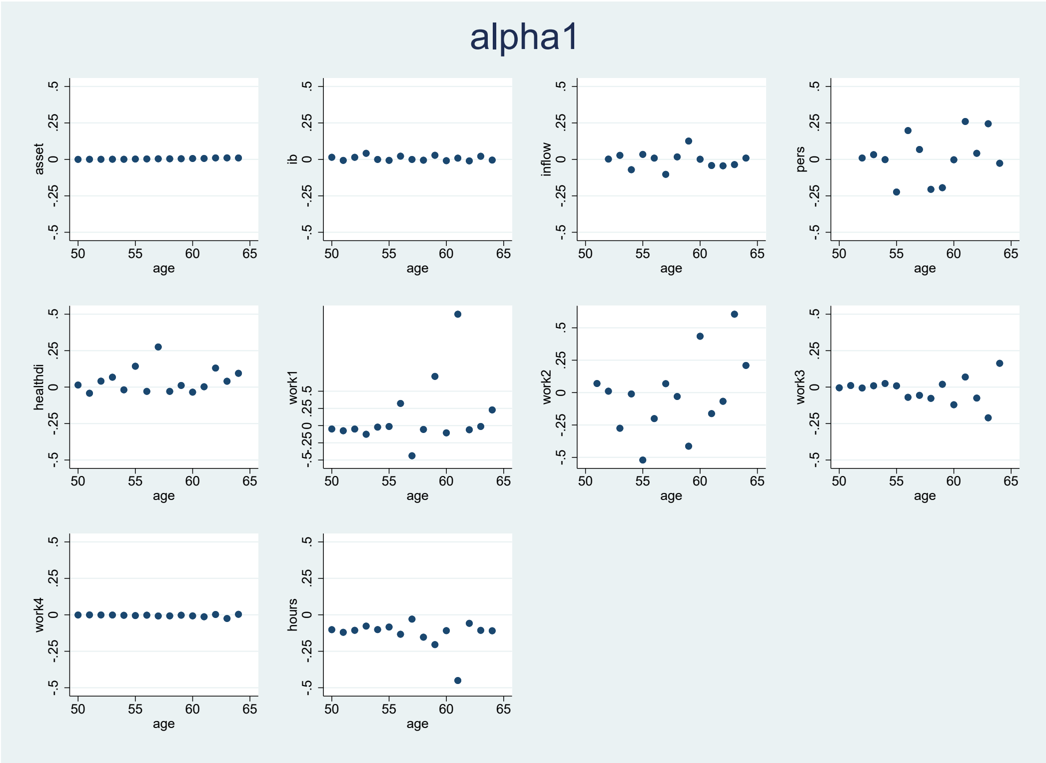


Figure E.6: Effect of one-standard-deviation change in the targeted moments on the parameter of the acceptance probability, α_1.

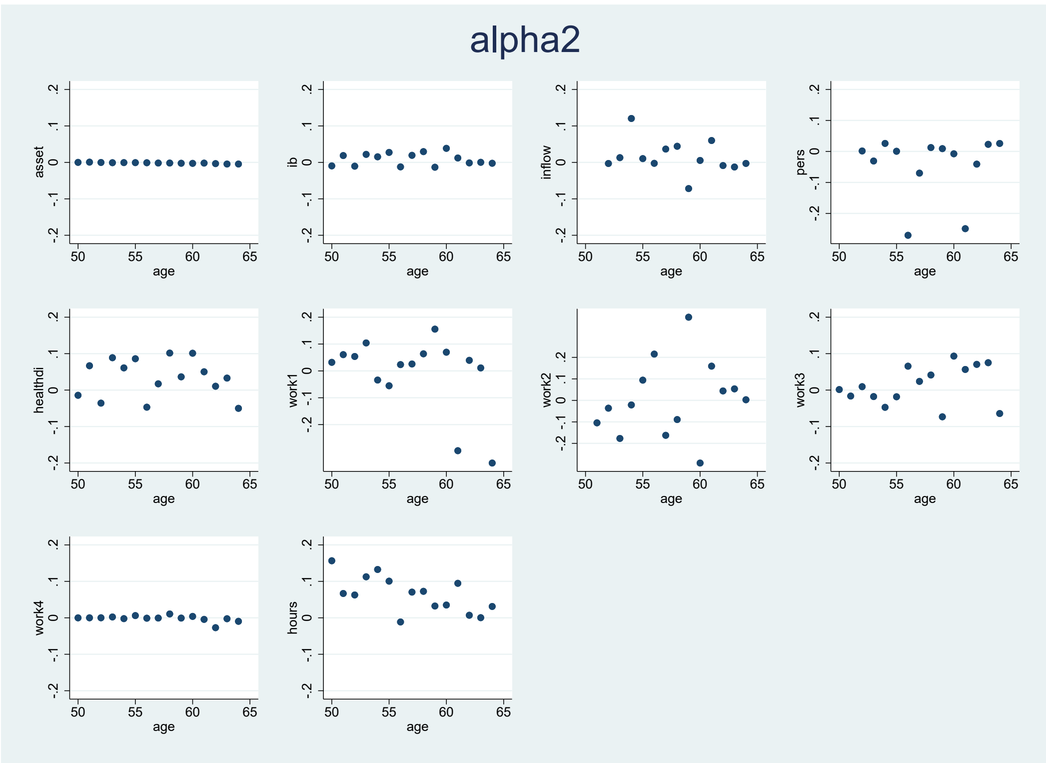


Figure E.7: Effect of one-standard-deviation change in the targeted moments on the parameter of the acceptance probability, α_2.

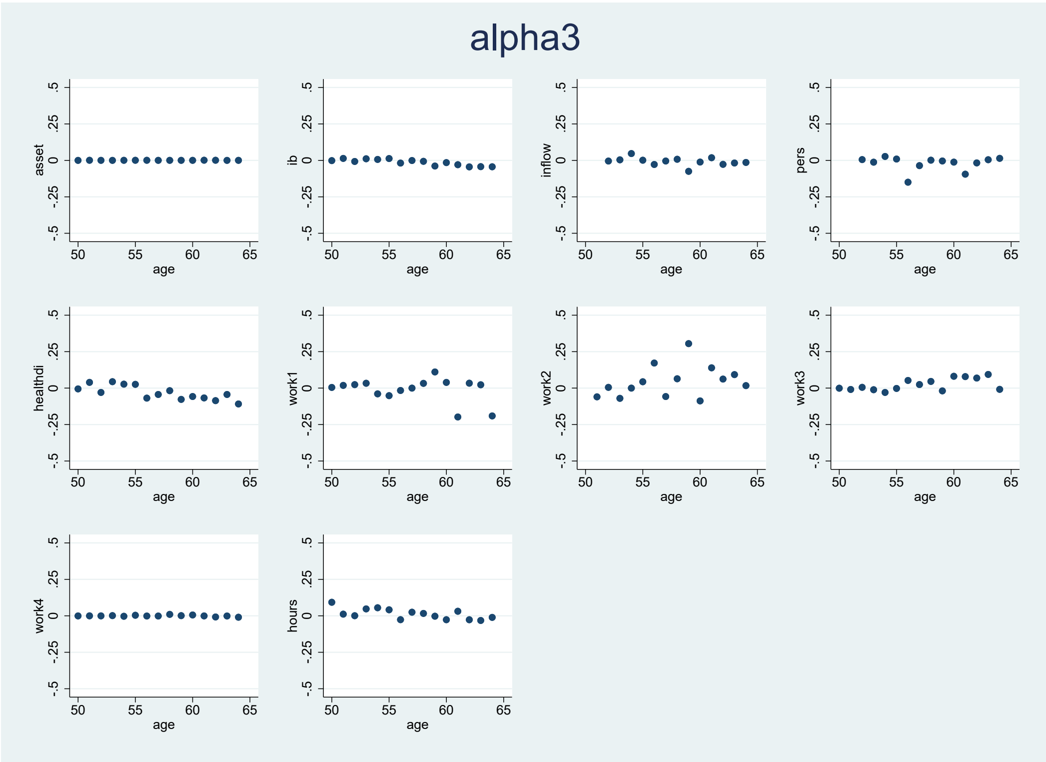


Figure E.8: Effect of one-standard-deviation change in the targeted moments on the parameter of the acceptance probability, α_3.

F Model fit

Moments not directly targeted in the estimation procedure.

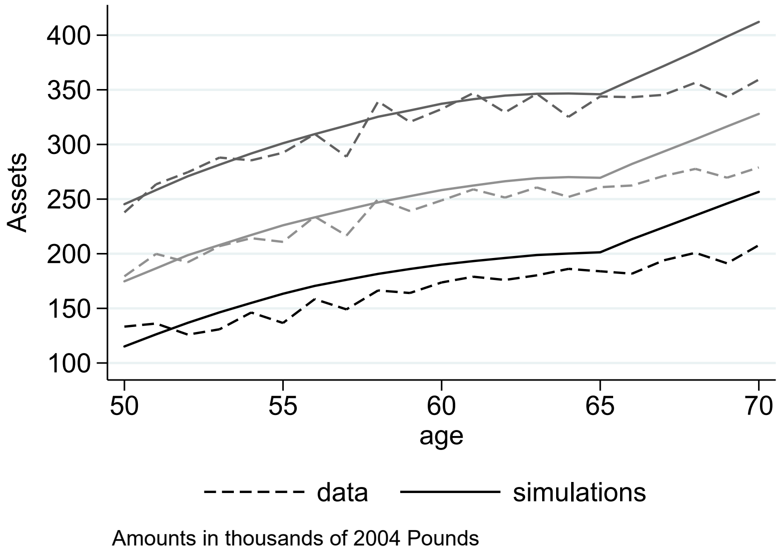


Figure F.1: Life-cycle profiles for median, first and second tertiles of assets. Simulations versus data.

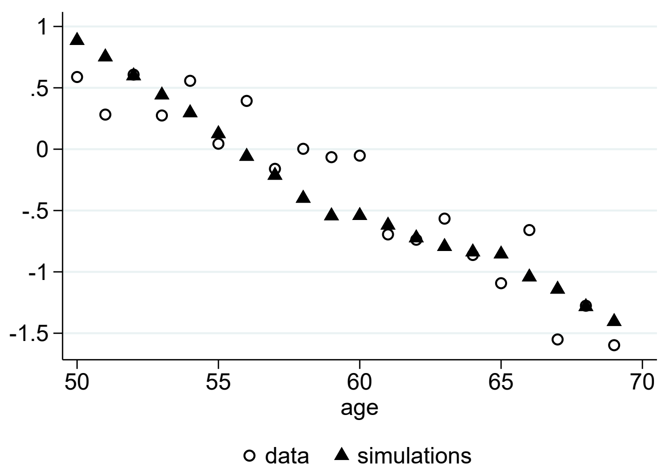


Figure F.2: Earnings life-cycle profile (deviation from life-cycle means). Simulations versus data.

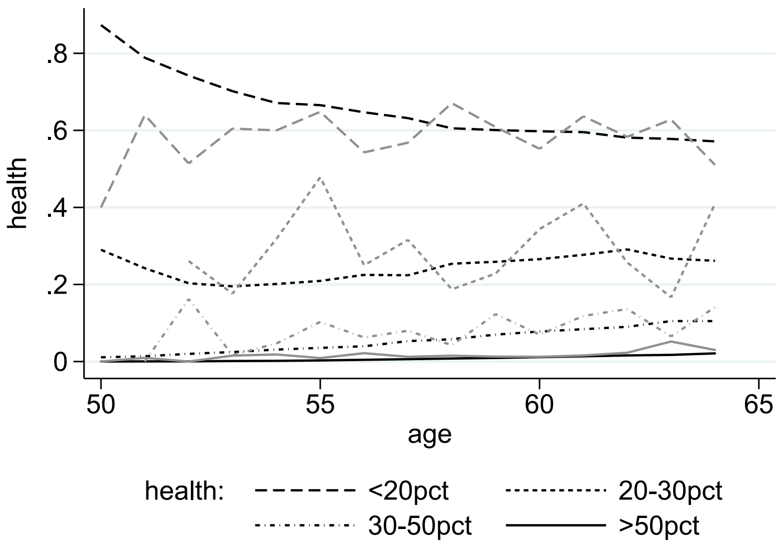


Figure F.3: Fraction receiving DI by health level. Simulations versus data.

Table F.1: Percentage receiving Income Support and Pension Credit by age. Data and simulations.

	Income Support		Pension Credit	
age	data	simulations	data	simulations
50–55	3.2	3.9	–	–
56–59	3.9	4.6	–	–
60–64	–	–	6.0	8.8
65–70	–	–	6.8	0.4

G Policy experiments and robustness analyses

Table G.1: Structural parameter estimates. Baseline and alternative model specifications.

Parameter	BL	$\nu = 3$	$\sigma_\eta^2/2$	$\phi_{P2} = \phi_{P0}$	$\phi_A = 50$	$\phi_A = 200$	$\phi_A = 500$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
γ	0.616	0.618	0.663	0.590	0.629	0.616	0.607
ϕ_H	4,201	4,014	4,329	3,975	4,306	4,256	4,149
ϕ_{P0}	909	892	1,075	707	969	854	777
ϕ_{P1}	46	47	47	60	42	45	46
ϕ_{P2}	467	497	666	–	612	492	416
α_1	0.200	0.183	0.273	0.073	0.267	0.210	0.264
α_2	0.124	0.136	0.174	0.064	0.169	0.176	0.261
α_3	0.122	0.135	0.173	0.062	0.169	0.175	0.254

Table G.2: Effects of policy experiments and elasticities to benefit generosity for different model specifications.

	BL	$\nu = 3$	$\sigma_\eta^2/2$	$\phi_{P2} = \phi_{P0}$	$\phi_A = 50$	$\phi_A = 200$	$\phi_A = 500$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experiment 1: DI benefit for an additional year if back to work.							
all							
Δ DI rate	-0.56	-0.51	-0.52	-0.52	-0.51	-0.45	-0.40
Δ LFP	0.45	0.42	0.43	0.43	0.42	0.36	0.32
sub-sample 1							
Δ DI rate	-5.68	-4.88	-5.24	-5.66	-4.97	-4.32	-3.66
Δ LFP	4.63	4.04	4.25	4.54	-4.03	3.60	3.11
Experiment 2: 10% reassessment probability.							
all							
Δ DI rate	-2.97	-3.14	-2.89	-3.18	-3.07	-3.17	-3.41
Δ LFP	1.50	1.84	1.49	1.58	1.72	1.98	2.39
sub-sample 2							
Δ DI rate	-30.63	-30.50	-29.23	-34.55	-30.02	-30.85	-31.80
Δ LFP	16.42	17.67	15.50	18.83	16.14	17.42	18.73
Elasticities							
DI app.	0.90	1.06	1.10	0.74	0.84	1.15	1.16
non-part.	0.36	0.45	0.41	0.33	0.45	0.49	0.63

Notes: Experiment 1: individuals are allowed to keep the DI benefit for an additional year if they go back to work. Experiment 2: DI recipients face a 10% probability of having their health reassessed on a yearly basis. All: individuals aged below 65; sub-sample 1 (receiving DI in the baseline or in the reformed scenarios); sub-sample 2: receiving DI in the baseline scenario only.

Table G.3: Welfare effects of policy experiments for different model specifications.

	BL	$\nu = 3$	$\sigma^2_\eta/2$	$\phi_{P2} = \phi_{P0}$	$\phi_A = 50$	$\phi_A = 200$	$\phi_A = 500$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experiment 1: DI benefit for an additional year if back to work.							
welfare effect	0.0004	0.0004	0.0004	0.0005	0.0004	0.0003	0.0003
welfare effect by health							
<20th pct	0.0009	0.0008	0.0011	0.0008	0.0009	0.0008	0.0008
20-30th pct	0.0012	0.0011	0.0016	0.0011	0.0012	0.0010	0.0010
30-50th pct	0.0005	0.0004	0.0004	0.0004	0.0004	0.0003	0.0003
>50th pct	0.0003	0.0003	0.0003	0.0004	0.0003	0.0002	0.0002
Experiment 2: 10% reassessment probability.							
welfare effect	-0.0001	<0.0001	>-0.0001	>-0.0001	0.0003	0.0002	0.0002
welfare effect by health							
<20th pct	-0.0135	-0.0152	-0.0121	-0.0170	-0.0120	-0.01490	-0.0172
20-30th pct	-0.0026	-0.0025	-0.0027	-0.0023	-0.0022	-0.0022	-0.0022
30-50th pct	0.0005	0.0007	0.0004	0.0011	0.0006	0.0010	0.0011
>50th pct	0.0016	0.0019	0.0015	0.0020	0.0019	0.0021	0.0023

Note: Experiment 1: individuals are allowed to keep the DI benefit for an additional year if they go back to work. Experiment 2: DI recipients face a 10% probability of having their health reassessed on a yearly basis.

H Welfare effect and revenue neutrality

The lifetime expected utility of an individual under the scenario s can be written as:

$$E_0U(s) = E_0\left\{\sum_{t=0}^{T}\beta^t\left(\pi_t\frac{1}{1-\nu}(c_{it}(s)^\gamma l_{it}(s)^{1-\gamma})^{1-\nu} + (1-\pi_t)\frac{1}{1-\nu}\phi_B(a_t(s)+k)^{(1-\nu)\gamma}\right)\right\}.$$

Let $s = b$ for the baseline scenario and $s = b'$ for the reformed scenario, in which some features of the benefit program have been modified. I define ρ as the proportion of consumption an individual is willing to pay from age 50 to age 90 to be indifferent between the baseline and the reformed scenario. We can obtain ρ by equating $E_0U(b')|_\rho = E_0U(b)$, where

$$\begin{aligned}
E_0U(b')|_\rho \equiv & E_0\left\{\sum_{t=0}^{T}\beta^t\left(\pi_t\frac{1}{1-\nu}(((1-\rho)c_{it}(b'))^\gamma l_{it}(b')^{1-\gamma})^{1-\nu} + \dots\right.\right.\\
& \left.\left.\dots + (1-\pi_t)\frac{1}{1-\nu}\phi_B((1-\rho)(a_t(b')+k))^{(1-\nu)\gamma}\right)\right\}\\
E_0U(b) \equiv & E_0\left\{\sum_{t=0}^{T}\beta^t\left(\pi_t\frac{1}{1-\nu}((c_{it}(b))^\gamma l_{it}(b)^{1-\gamma})^{1-\nu} + \dots\right.\right.\\
& \left.\left.\dots + (1-\pi_t)\frac{1}{1-\nu}\phi_B(a_t(b)+k)^{(1-\nu)\gamma}\right)\right\}
\end{aligned}$$

such that ρ satisfies the following expression $1-\rho = (\frac{E_0U(b)}{E_0U(b')|_{\rho=0}})^{\frac{1}{(1-\nu)\gamma}}$.

The policy experiments presented in section 8 are revenue neutral. I assume that revenue costs or savings resulting from policy changes are offset by iterative proportional adjustments of the progressive earned income tax rates (see the first column of Table I.4 in Appendix I for the tax schedule implemented in the model). Let the present discounted value of government revenues (GR) and costs (GC) be

$$\begin{aligned}
GR = & \sum_{i=1}^{N}\sum_{t=1}^{T} r_t\left(\tau(w_{it}h_{it}, di_{it}, pb_{it}) + \tilde{\tau}(ra_{it}) + nic_{it})\right)\\
GC = & \sum_{i=1}^{N}\sum_{t=1}^{T} r_t\left(tr_{it} + di_{it} + appcost * \mathbb{1}(dapp_{it}=1) + sb_{it}\right)
\end{aligned}$$

GR are the sum of (labor, pension and DI) income taxes ($\tau(\dots)$), taxes on investment income ($\tilde{\tau}(\dots)$) and national insurance contributions (nic). GC are the sum of means tested benefits (tr), DI benefit (di), the cost for the government of assessing DI eligibility $appcost$ (Source:`www.nao.org.uk/wp-content/uploads/2016/01/Contracted-out-health-and-disability-assessments.pdf`) and the basic amount of the state pension (sb).[39]

[39] See Appendix I for details on how pension wealth enters the model.

I assume that the present discounted value of net revenue flows is fixed and equal to $\bar{D} = GR - GC$. When I change the benefit structure I proportionally adjust τ to keep $\bar{D}$ fixed. Table H.1 reports the percentage change applied to the tax rates to reach revenue neutrality. For example, in experiment 1 when the subsidy is equal to 20% of the benefit amount the tax rates reported in Table I.4 are adjusted as follows: $0.1 \times (1 - 0.0011)$, $0.22 \times (1 - 0.0011)$ and $0.4 \times (1 - 0.0011)$.

Table H.1: Percentage changes in the tax rates to reach revenue neutrality of policy changes.

	Experiment 1				
subsidy	20%	40%	60%	80%	100%
	-0.11	-0.24	-0.37	-0.51	-0.69
	Experiment 2				
reassessment probability	10%	30%	50%	70%	100%
	-4.51	-8.74	-10.46	-11.34	-12.18
	Model fit				
benefit amount	-10%	+10%			
	-1.81	2.33			

I 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. Details on IB are privided in Section 3. Retirement Pension can be received starting from state pension age (65 for men). 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.

In the model, I assume that everyone receives the basic state pension (sb) amount at age 65 (£3,640). Additionally, I introduce heterogeneity in pension wealth with the state variable q_t. As a measure of pension wealth I use ELSA generated variables on pension wealth for both State pension and occupational/private pensions. These measures are computed as the present discounted value of imputed future pension incomes assuming that individuals receive their state pension and annuitize private pensions at age 65 (see Figure I.1 which

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

shows that a large fraction of individuals start to receive the (state and/or private) pension at age 65).

In the model, individuals that work are assumed to contribute 6% of their earnings to the pension pot (this is the average contribution rate to Defined Contribution pension funds). At age 65, the pension wealth is annuitized using as annuitization rate the average annuitization rate at age 65 implied by the procedure used to generate the ELSA wealth variables, that results to be 5.4%. More specifically, I use equation I.1 taken from Crawford (2012) and set the real discount rate (δ) and the real indexation of pension income (i) in retirement at 2.9% (the real interest rate in the model); I additionally set the number of years needed to draw the pension at current age minus SPA (r) and I take the life expectancy (LE) from the life tables. This allows me to invert the formula and derive the pension amount at SPA (P) from the generated wealth variable ($Wealth$).

$$Wealth = \sum_{s=r}^{r+LE} \left(\frac{1}{1+\delta}\right)^{s} (1+i)^{s-r} P \tag{I.1}$$

The implied pension amounts results too high with respect to what observed in the data. Therefore, the pension wealth amount is proportionally adjusted in such a way that the simulated distribution of pension income matches the data one (see Table I.1). In particular, I reduce the pension wealth by about 20%.

Table I.1: Distribution of pension income (public and private) between age 66 and 70 in the data and at age 65 in the simulations.

	mean	10th pct	25th pct	50th pct	75th pct	90th pct
data	10,644	2,460	4,904	8,180	13,309	20,913
simulations	10,264	3,409	4,721	8,215	14,519	20,485

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. The threshold changes with age and is set to reproduce the fraction receiving the benefit. The amount of the benefit is set at £3,000 (using 2004 average amounts from the Department of Work and Pension tabulation tool.).

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 $\pounds$8,000 in cap-

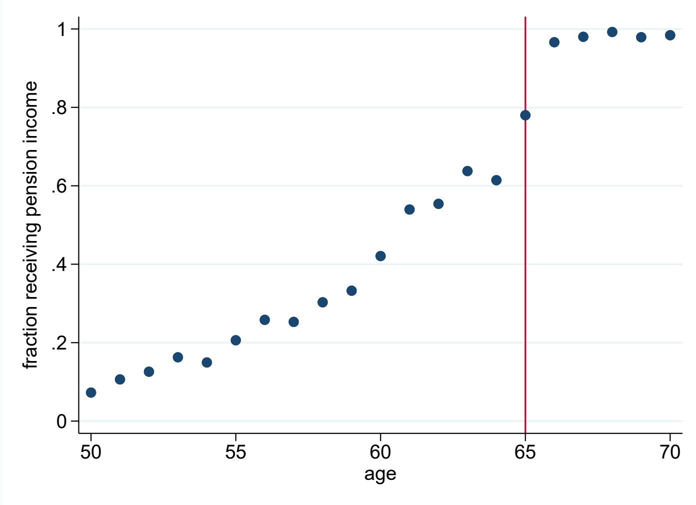


Figure I.1: Fraction receiving any type of pension by age. ELSA data waves 1-6.

ital. 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 I.2. 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 I.2 (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 I.2. The income taken into account is the same as for GC except WTC that are deducted.

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

Table I.2: 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

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 I.3). 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.

Table I.3: Working tax credit

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

The income tax schedule is based on three bands.

Table I.4: 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, state and private pensions, incapacity benefit and interest income (ra) net of personal tax-free allowances and other exemptions. The main tax allowances are listed in Table I.5.

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 I.5: 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)