Università
Ca'Foscari
Venezia

Scuola Dottorale di Ateneo
Graduate School

Dottorato di ricerca
in Economia
Ciclo XXIX
Anno di discussione 2017

Health, Disability Insurance and Labour Supply: Evidence from a dynamic structural model

SETTORE SCIENTIFICO DISCIPLINARE DI AFFERENZA: SECSP/01

Tesi di dottorato di **Chiara Dal Bianco**, matricola: 956071

Coordinatore del Dottorato
Prof. Giacomo Pasini

Supervisore del Dottorando
Prof. Andrea Moro

Co-supervisore del Dottorando
Prof. Giacomo Pasini

Chiara Dal Bianco

Health, Disability Insurance and Labour Supply: Evidence from a dynamic structural model

December 2016
Dissertation for the Ph.D. degree
Department of Economics
Ca' Foscari University of Venice

Acknowledgments

First and foremost I am grateful to my supervisor Andrea Moro for his continued support and guidance, his constant encouragement, his precious advice and his extremely valuable contribution of time.

My sincere thanks for helpful comments and constructive suggestions to my co-supervisor Giacomo Pasini, the internal referee Francesca Zantomio and the two external referees James Banks and Stephen Pudney.

I would like to thank Mariacristina De Nardi, who gave me the opportunity to visit UCL, and Guglielmo Weber. They have followed and supported the development of my Ph.D. research project from the first steps and provided me with extremely helpful suggestions.

Special thanks go to my family. To my parents for all their love and encouragement. And most of all to my loving, supportive, and patient husband Simone: we made it!

Data from The English Longitudinal Study of Ageing, obtained through the UK data Archive, have been used with permission.
All responsibility for the analysis and interpretation of the results lies with the author.

Contents

Introduction

This thesis presents a quantitative assessment of the labour market effects and the distributional consequences of disability insurance (DI) programme in the UK and its insurance value for potential beneficiaries.

In an ageing society, the concern with the sustainability of countries' pension systems and disability related programmes has widely inspired researchers in the last decades. To design effective welfare policies it is fundamental to understand the economic consequences of the changes in the population demographic structure and the trajectories to retirement of older workers, that often go through disability programmes (Wise, 2012, 2016).

Focusing on the UK context, recent studies have shown that even if retirement pathways have changed substantially in the last thirty years, unemployment spells and in particular disability and sickness benefits continue to represent important routes out of the labour market, together with the most widespread direct transition from employment to retirement (Banks *et al.*, 2012). The UK context is characterized by a long season of reforms, started in the mid-nineties, aimed at reducing public spending on disability benefits targeted to individuals whose health limits their current ability to carry out paid work. Whereas the introduction of Incapacity Benefit in 1995 caused a sharp decline in the number of claimants, in particular among workers approaching state pension age, the effect of the most recent Employment and Support Allowance is less clear cut. On the one hand the lack of suitable data and on the other hand the short period after the reform, have allowed only descriptive analyses and first evidence of heterogeneous effects among age, sex, education, and health groups (Banks *et al.*, 2015a).

To contribute to the discussion on the effectiveness of the recent reform and to shed light on the role of state provided DI in the relationship between health and labour supply in the UK, I develop and calibrate a model of labour supply and saving behavior of males living with a partner at the end of their working life. In the model, individuals make a discrete choice about whether to participate in the labour market and in case of participation they choose hours of work. Moreover, if their health level is below a calibrated threshold they can apply for state provided DI. The model accounts for several crucial institutional features affecting such decisions: the financial incentives provided by the pension systems are considered by carefully modelling accrual in both public (earnings related) and private pensions; along with the contributory health-related benefit (Incapacity Benefit), financial incentives provided by non-contributory and means-tested benefits are included in the model by accurately reproducing eligibility to Disability Living Allowance and Attendance Allowance as well as Income Support (including the premium for low-income households containing at least one disabled individual) and Working Tax Credits (which has a supplement for disabled workers).

The model allows for uncertainty on wage realization, health development and life expectancy. In developing the model, particular attention has been devoted to the measure of health and the evolution of health over time. I construct a continuous health index using a set of objective health indicators collected in the English Longitudinal Study of Ageing (ELSA) and replicating the same health domains covered by the health assessment, used to determine eligibility to DI benefit. The specified process for health has a deterministic component that depends on age and a stochastic component allowing both persistent and transitory shocks. Health status enters the deterministic component of the exogenous wage process (productivity channel) and the probability of surviving to the next period, moreover there is a time cost of being in bad health which affects utility through leisure. I do not consider the medical expenditures channel 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 costs of institutionalization arise (they are not covered by the National Health Care system).

The parameters of the health process, the wage process and survival probabilities conditional on health are estimated outside the structure of the model due to the exogeneity assumption. The remaining parameters are calibrated to match profiles generated by the dynamic model with data life-cycle profiles of assets, participation by health level, hours worked and the fraction receiving disability benefit. The calibrated model is able to replicate the main facts observed in the data. I also show the ability of the model in replicating other features of the data not considered in the calibration phase.

To assess the effect of alternative policies I simulate several scenarios by altering some of the parameters governing the DI programme. I document a number of relevant findings. First, a simulated scenario in which Incapacity Benefit (IB) is removed reveals the presence of unused working capacity among IB recipients and the important redistributive role of the benefit. After benefit removal, among those receiving IB in the baseline scenario 22.4% become active and average participation increases by 4% for individuals aged between 50 and 64. The effect is mainly driven by those with health close to the eligibility threshold, i.e. those at the margin of programme entry. In terms of distributional consequences, about a quarter of the individuals are not affected by benefit removal, whereas the others experience a change in the present discounted value of disposable income: an increase for those participating more in the labour market as a result of IB removal and a decrease for those in worse health and without any working capacity left. The effect of removing the benefit can be seen in the thicker left tail of the income distribution and the consequent increase in inequality.

Second, the insurance value of IB shows large heterogeneity by health and wealth levels. Those with lower level of health at age 50 and with assets in the first quartile are willing to pay between 10% and 44% of their assets at age 50 on average to have the DI insurance in the benefit system. Third, a stricter health assessment to receive the benefit might help target efficiency and increase labour market participation but at the cost of increased inequality, in particular among those with lower health levels. Fourth, an alternative policy lever, such as policy interventions aimed at improving non-discrimination and accessibility in the work place to foster labour supply inclusion of individuals with disability, increases participation, reduces benefit duration and especially reduces inequality in income distribution. Finally, a more generous benefit does not have a large effect on benefit take-up and, at the same time, it reduces the dispersion in the income distribution by increasing income in the lower quantiles.

The vast literature on the economic effects of disability insurance, which I survey in Chapter 1, has mainly exploited exogenous variations provided by the details of application process, and nonlinearities or changes over time in the programme structure. My approach, in the spirit of structural retirement literature (French, 2005; French and Jones, 2011; Bound *et al.*, 2010; Low and Pistaferri, 2015), allows the investigation of behavioral responses to alternative hypothetical policy reforms accounting for the full dynamic effects of the policies on the agents' choices. Moreover, it allows the evaluation of the welfare effects and the distributional consequences of the alternative DI structures.

This work contributes to the existing literature in several ways. First, I propose a richer model than in previous works: labour supply decision (both the extensive and the intensive margins), DI application, private and public pension accrual and pension claiming decision are modelled in a unified framework to shed light on the role of DI as a potential retirement pathway[1]. Second, the model assumes a continuous, more comprehensive measure of health than previous studies, which have extremely simplified the treatment of this variable for computational reasons or data constraints failing to reproduce outcomes from the data in some relevant dimensions[2]. This allows my simulations to better evaluate the policy effects by health level. Finally, this is to my knowledge the first study analyzing the UK disability benefits system using a structural approach.

The rest of the thesis is structured as follows. In Chapter 1, I first provide a brief overview of theoretical works and empirical evidence on the relationship between health and labour supply. Then, I focus on the role of public provided disability insurance in this relationship by reviewing the literature on economic analysis of programmes targeted on people with disabilities. Chapter 2 describes the model of lifetime decision-making, its calibration to the data and its ability to replicate main patterns observed in the data. Chapter 3 presents results from several policy experiments aimed at exploiting the effects of alternative DI structures on DI rate, labour supply and income distribution, and measuring the insurance value of the programme.

[1]For example, French (2005) does not model DI application decision and private pension accrual and claiming decision within his model of retirement and health; Bound *et al.* (2010) specify a discrete choice dynamic programming model and abstract from savings and consumption decision assuming that in each period a person consumes all of his income; Low and Pistaferri (2015) focus on DI application and consider only the extensive margin of labour supply decision.

[2]As conjectured in French (2005).

Chapter 1

Health, disability insurance and labour supply: a survey

Labour force participation of older workers and their retirement pathways have been an important research field in the last decades. A persistent increase in life expectancy, the resulting change in the demographic structure of the population and the risk of an increasing number of individuals having health conditions and disabilities causing work limitations, have risen the concern about the sustainability of countries' pension systems and disability insurance programmes. As a result, a large literature has tried to identify the role of financial incentives and health patterns in explaining trends in labour market exit.

Great attention has been devoted to retirement incentives arising from public old age pension provision and, for some countries, occupational schemes (Gruber and Wise, 1999, 2004, 2007). The effort in extending working lives to improve the sustainability of social security, highlights the importance of understanding the working capacity of individuals approaching retirement, in particular for what concerns their health status and the extent to which it limits their ability to work (Wise, forthcoming). Trends in mortality and health have been widely investigated to understand whether increased life expectancy brings forward longer periods of morbidity and higher disability risk (see for example Morciano *et al.* (2015) for the UK). In this context, the insurance and the financial incentives provided by sickness and disability policies are of particular importance (Wise, 2012, 2016), especially in the case of adverse economic conditions. A recent report on sickness and disability programmes in OECD countries (OECD, 2010) addresses the importance of dealing with the widespread use of disability benefits and promoting labour market participation of people with disability in order to avoid a 'medicalisation' of labour market problems experienced in previous downturns.

The purpose of this survey is to provide first a brief overview of theoretical works and empirical evidence on the relationship between health and labour supply (Section 1.1) and then to focus on the role of public provided disability insurance in this relationship (Section 1.2). Literature on DI effects on labour market participation and welfare consequences of different DI programmes are reviewed in Sections 1.3 and 1.4 respectively. The main features of DI programme in the UK, recent reforms and literature evaluating their effects are discussed in Section 1.5. A summary of the results and a discussion on unaddressed research needs is provided in Section 1.6.

1.1 Health and labour supply

The role of health in the labour supply decision of individuals has been a central topic in the literature. The theoretical foundation of this empirical question can be found in Grossman health investment model. The Grossman model (Grossman, 1972) assumes that health affects utility directly, and indirectly by reducing the time available for leisure and work activities. Two relevant implications of Grossman model are that health should be treated as endogenous and that endogenous health affects labour supply. In most of the literature, however, health is included as an exogenous determinant of labour market outcomes under the assumption that most of the variation observed in health is due to exogenous shocks rather than health investments or unhealthy behaviours that deteriorate the stock of health (Currie and Madrian, 1999).[1]

In the empirical literature, there is consensus regarding the direction of the effect of health on several outcomes, such as participation, hours worked and wages, however there is little agreement on the magnitude of such effects as highlighted in Currie and Madrian (1999). Great attention has been devoted to older workers approaching retirement age (Lumsdaine and Mitchell, 1999), for whom the risk of the onset of health conditions is higher and they have more valuable labour market exit options relative to younger workers.

The individual response to health shocks in terms of participation is influenced by

[1] The literature that treats health as endogenous is growing (see for example Scholz and Seshadri (2013); Hai and Heckman (2015)) but a survey of these recent works is out of the scope of this chapter.

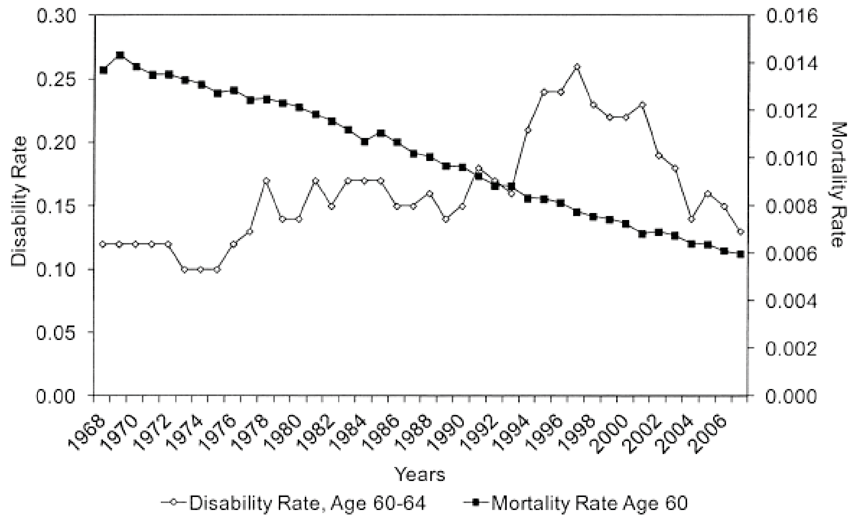
the institutional setting, both in terms of the generosity of the social security system and the labour market structure. In the US, health insurance plays an important role: workers covered by an employment-based programme might increase labour supply after an health deterioration to insure themselves against future increased healthcare expenditures (Madrian, 1994; Kapur, 1998; Bradley *et al.*, 2013). In most European countries out of pocket expenditures represent only a marginal part of healthcare costs due to the availability of national health care services. Several papers provide estimates of the causal effect of health drops on labour supply in the European context, for example García-Gómez (2011) and Trevisan and Zantomio (2016) adopt a unified framework for a large set of European countries.

The uncertainty on health developments suggests that insurance should play an important role in explaining the effect of health deterioration on labour market participation, in particular when old. In addition to private and publicly provided health care systems, another important source of insurance against the onset of health limitations affecting working capacity are state provided disability insurance programmes (Bound and Burkhauser, 1999a). These programmes have been shown to play a large role in explaining the transition from work to retirement, as many workers pass through DI in their departure from the labour force (Wise, 2012, 2016). In the remaining of this chapter I will focus on the literature investigating the labour supply effect of DI programmes and the potentially large effects that changes in these programmes could have on the labour force participation.

1.2 Publicly provided disability insurance

Publicly provided disability insurance is part of a set of rules developed in some countries that guarantee a minimum level of protection against health shocks for all workers in private and public sector. In particular, disability related benefits are designed to insure workers against reduced working capacity, and consequent loss of earnings capacity, resulting from health deterioration (Bound and Burkhauser, 1999a). On top of a primary insurance role, disability schemes have a redistributive role: together with earnings replacement benefits linked to earnings history, social security systems often provide means-tested sickness and disability benefits targeted to low income individuals.

Figure 1.1: Trends in mortality and DI participation in the United Kingdom.



In 2011, in OECD countries the average public spending on disability amounts to 2.2% of GDP. Figure 1.1, taken from Wise (2012), reports for the UK the time series of the fraction in disability benefit receipt among individuals aged 60 to 64 and mortality rate for those aged 60. Similar patterns emerge when looking at other OECD countries. Disability rate does not decrease with increased life expectancy. On the contrary, it seems to increase or, at the end of the observation period, to stabilise at about 15%. Trying to understand the determinants of this pattern is not an easy task.

Undoubtedly, DI has social value in particular for those at the bottom of the wealth distribution, for which savings and private insurance do not represent alternative insurance mechanisms against permanent health deterioration. Several aspects need to be considered in order to design an efficient insurance scheme, among the others the minimisation of public spending, optimal targeting of those in need, and optimal benefit duration. Substantial attention has been devoted in the literature to understand to what extent disability generosity (both in terms of screening process and benefit amount) provides work disincentives. Less attention has been instead devoted to the insurance value of DI for beneficiaries and to what extent DI programmes are able to insure individuals against work-limiting health risk.

1.3 The effects of screening process and benefit amount on labour market participation

Most of the literature investigating the work disincentive provided by DI focuses on the estimation of the elasticity of non-participation in the labour market and the elasticity of participation into DI programmes with respect to programme generosity. To get a reliable estimate of these quantities one needs exogenous variations in DI benefit structures. However, DI is generally a within country uniformly administered insurance programme, limiting the possibility of treatment-control comparisons. Reforms of the programme over time, even if generally rare, might provide interesting exogenous variations within country. Further concerns to the challenge of identification is given by the state of the economy: confounding factors, such as trends in employment, earnings and health add difficulty in identifying the causal effect of DI on labour supply participation.

Among the first to investigate the labour supply effect of DI is Parsons (1980a,b). He uses cross-sectional variation in potential DI benefit in the US. He regresses participation decision on DI replacement rate to estimate the elasticity of participation with respect to DI generosity. He estimates an elasticity between 0.49 and 0.93 (result confirmed by Slade (1984) who uses administrative data to derive DI replacement rate) which implies that the decline in participation among prime-age males during the post-World War II can be virtually explained by the growth in DI programme and in general welfare. However, this approach overstates the effect of DI generosity on labour supply: DI replacement rate decreases as past earnings increase, thus it is difficult to state whether benefit recipients are those with higher replacement rates or those with low labour market attachment and discontinuous careers (Bound, 1989; Haveman and Wolfe, 1984). Bound (1989) estimates a 'placebo' regression using only non-applicants and finds similar results with respect to Parsons' estimates, confirming the weakness of his approach.

The seminal paper in this literature is Bound (1989). Bound (1989) has been the first to propose a 'natural experiment' framework in this context. He proposes to use rejected applicants as control group for work attachment in the absence of DI. He argues that participation rate among rejected applicants represents an upper bound of participation rate among those receiving the benefit in the Social Security Disability Insurance (SSDI), in the US. Among rejected applicants, less than half did any paid work in the previous year and less than a quarter of them worked the full year. Assuming that poor health is the main reason why rejected applicants did not come back to work, Bound estimates that employment rate among males aged 45 to 64, beneficiaries of SSDI, would have been 35 percentage points higher during the 1970s in the absence of the programme. This translates into half the decline in labour force participation for males aged 45 to 54 observed between the 1950s and the 1970s explained by SSDI. The percentage decreases to less than 25% when considering males aged 55 to 64. The following literature on DI tries to give a more transparent causal interpretation to the associations emerged in Bound's work.

Changes in the programme structure before and after the observational period do not allow to extend the results. To assess the robustness of Bound's results over time, Bound and Waidmann (1992) use time series data to gauge the effect of the DI programme's expansion between the 1950s and 1980s in the US on labour force attachment. Under the assumption of no time trend in the proportion of truly disabled working-aged males, any change in the proportion of males reporting themselves as disabled can be attributed to factors other than health. They compare trends in self-reported disability and labour force participation and they find that since the 1970s the increase in the fraction of men aged 45 to 54 reporting themselves as disabled corresponds to a comparable decrease in labour force participation. In the same period in the US there has been an expansion of the DI programme. This result confirms the negative association between DI generosity and labour force attachment, even though no causal interpretation can be inferred.

Another potential problem in Bound (1989) is the presence of other transfers targeted to the disabled that may act as substitutes for SSDI that are not accounted for. This might results in an overestimation of the programme's effect on disability (Bound, 1989, 1991; Bound and Burkhauser, 1999b).

Parsons (1991) instead claims that Bound's estimate of participation among recipients in the absence of the disability benefit is not an upper bound, rather it understates the true effect of DI. The main concern is with the assumption that accepted applicants are as likely to work as denied applicants: Bound considers only the first application to DI, but denied applicants can decide to pursue appeal or to reapply. Moreover, the application process implies a period out of the labour force that can deteriorate applicants' human capital and their chances of finding a job after benefit denial. Heterogeneity in health levels and in tastes for work among accepted and denied applicants can cause a bias in Bound's estimate, however the direction of the OLS bias is not obvious.

The literature has expanded and tested Bound's findings using several methodological approaches. In Section 1.3.1 I survey the studies that use the comparison between accepted and denied applicants, exploiting a unique feature of the US DI programme. Moreover, I present results from the only paper I am aware of that accounts for the long period out of the labour market waiting for the final decision in estimating the effect of DI on late employment in the US. To overcome the adequacy of denied applicants as control group, many studies exploit the exogenous variation provided by reforms of the DI structure across countries and, in particular, within country and over time. A survey of these works is in Section 1.3.2. Evidence on the role of economic conditions in affecting the impact of DI programme on labour market participation are surveyed in Section 1.3.3. Finally, Section 1.3.4 presents works using dynamic models of individuals' behaviour to investigate behavioral responses to different DI structures and to analyse economic and welfare consequences of several reformed scenarios.

1.3.1 Comparison of accepted and denied applicants

The use of rejected applicants as comparison group for accepted applicants (Bound, 1989) has been widely questioned in the literature. The availability of richer datasets and the use of exogenous variations in the application process have allowed to test and extend Bound's results which have been proved remarkably robust.

Lahiri *et al.* (2008) develop a model of DI application behavior using data from the Survey of Income and Programme Participation for the years 1990, 1991 and 1992 matched with Social Security Administrative data for 1989-1995. The richness of the data allow them to account for pre-application differences in labour market attachment between accepted and denied applicants. They estimate that 37% of DI beneficiaries would return to work in the absence of the benefit with respect to the 50% estimated by Bound (1989). Moreover, they find heterogeneous effects on the elasticity of application with respect to benefit amount: males and in particular low earners with discontinuous careers (such as blue collar workers) are more responsive to changes in the benefit amount.

von Wachter *et al.* (2011) use longitudinal administrative data sources to extend Bound (1989) analysis in several directions. They consider not only older male applicants aged 45 to 64 but also younger applicants aged 30 to 44, that have become increasingly relevant starting from the early 1990s. They provide results disaggregated by different impairment, industry, and earnings groups. Results of Bound (1989) for older workers are confirmed using their detailed dataset and are robust to different specifications. Focusing on younger workers, they find a higher employment potential with respect to older ones, in particular among those applying for mental health and musculoskeletal conditions. Distinguishing between applicants allowed at different stages of the application process, they find that those allowed at later stages are not clearly classifiable as unable to work. Moreover the earnings pattern of applicants has been declining over time with respect to nonapplicants and rejected applicants show discontinuous careers; allowed applicants instead experience earnings decline only in the year before application. The effects on earnings after application is negative and permanent for both allowed and rejected applicants relative to nonapplicants. von Wachter *et al.* (2011) confirm previous results for older workers but provides evidence of an increasing employment potential of DI recipients since the late 1970s, mainly driven by younger applicants. Possible explanations are increased income inequality and the worsening of the economic conditions among individuals in poor health, but also increased work disincentive provided by DI to low-earners in a context of an overall deterioration of economic conditions.

Three papers, Chen and van der Klaauw (2008); Maestas *et al.* (2013); French and Song (2014), use the comparison between those who are denied and allowed benefits as source of identification by exploiting discontinuity in eligibility criteria and random assignments of examiners/judges to applications to causally estimate the effect of DI on labour supply in the US.

Chen and van der Klaauw (2008) apply a regression-discontinuity approach to estimate the effect of DI receipt for 'marginal' applicants (40% of applicants), for which the final eligibility decision requires to consider vocational factors. The discontinuity they use come from the rules applied to determine awards among marginal applicants. In particular, the procedure generates discontinuities in the award rate as a function of age. They estimate that DI causes a reduction in participation by 6 to 12 percentage points for this group of applicants. Applying Bound's approach they also estimate participation in the absence of DI during the 1990s and find that it would have been 20 percentage points higher.

Maestas *et al.* (2013) exploit exogenous variation in application outcomes coming from randomly assignment of applications to examiners with different allowance propensities at the fist step of the disability assessment process. They use data on the universe of DI applications in a given year (2005-2006) merged with administrative data on DI receipt up to six years after the initial decision and administrative earnings data to look at labour market outcomes two to four years from the initial award. Allowance propensity of the examiners at the beginning of the evaluation process is used as instrument for the final decision. Maestas *et al.* (2013) estimate an employment rate of 28 percentage points higher for beneficiaries on the margin of DI entry had they not received the benefit. The effect differs with impairment severity ranging from zero to 50 percentage points reduction in employment. The fraction of beneficiaries on the margin of DI entry is estimated to be 23 percent of total applicants and they tend to be younger, with low earnings history and with mental impairments.

A complementary study, that confirms and reinforces previous results is French and Song (2014). They focus on the second stage of the application process that is the random assignment of applications to administrative law judges. As in Maestas *et al.* (2013), they use judge specific allowance rate to instrument the final decision and thus to causally estimate the effect of DI receipt on labour supply. Having a longer panel, they also account for possible appeals and reapplications: 60% of rejected applicants are allowed benefits within ten years. They find that benefit receipt causes on average a 26 percentage points decrease in employment. The effect is lower for those over 55, with higher education and mental impairments.

As pointed out by French and Song (2014), a large fraction of denied applicants continue to appeal and reapply in the attempt of getting the benefit. The application or reapplication process requires the US workers to remain out of the labour force or to work a limited amount of hours. This might have dramatic consequences on applicants' human capital. Autor *et al.* (2015) investigate the role of time out of the labour force on late employment by exploiting exogenous variation in the decision time induced by examiners having different processing speed. Identification relies on random assignment of applicants to examiners. Focusing on initially allowed applicants, a one-month increase in processing time has a persistent effect on employment and earnings: annual employment is reduced by 4 to 5% and earnings by 8 to 13% after the decision. Extending the analysis to both awarded and denied applicants, the effect of waiting time is smaller and not persistent. This additional negative effect of DI, independent on DI receipt, causes previous estimates to be downward biased. Autor *et al.* (2015) estimate the disincentive effect of receiving DI equal to a 48 percentage points decrease in employment three years after the application when accounting for this additional channel, instead of 27 percentage points when waiting time is not considered. The long run effect (six years post-application) is a reduction of 25 percentage points instead of 17 in employment rate for those at the margin of programme entry.

1.3.2 Exploiting exogenous variations in DI programme

Another approach to measure the effects of DI is to exploit quasi-experimental variations in DI benefit structure. The advantage of such an approach, when applicable, is the presence of a 'natural' control group that easily allows to disentangle the effect of DI changes from the effect of other confounding factors.

Gruber and Kubik (1997) estimate the effect of screening stringency on labour supply using dramatic rise in state denial rates in the late 1970s in the US. The sharp increase in denial rates was not homogeneous across states and thus they use this heterogeneity as source of identification. They find that a 30% increase in denial rates in that period led to a fall in labour force non-participation among 45 to 64 year old of 8.3%. Using BMI as health measure (values on the tails of the BMI distribution are signals of bad health and of higher risk of disability), they also find that the effect is mainly driven by those in better health: the non-participation among the 'able' fell by 11.1% whereas no effect is found among the 'truly' disabled.

Gruber (2000) exploits Canadian reform on DI programme that took place in 1987. The benefit amount was raised by 36% but it remains unchanged in Quebec that has a distinct DI programme with respect to the rest of Canada. He applies a difference-in-difference approach and estimate an elasticity of labour supply for older workers with respect to DI benefit of 0.28-0.36, suggesting a large effect of DI benefit on employment.

The Canadian peculiar DI programme offers other policy changes in 2001 and 2005 that have been explored by Campolieti and Riddell (2012). In 2001 the main Canadian DI programme introduced an annual earnings disregard and in 2005 the Quebec DI programme introduced 'an automatic reinstatement provision whereby former recipients could remain eligible for up to 24 months'. The focus is on the probability that disability beneficiaries participate in the labour market and on the probability that they exit the benefit. The first policy intervention does not have effect on DI entry or exit probabilities, but has sizable effect on the probability of doing any paid work during benefit receipt:

employment among male beneficiaries increases by 5.7 percentage point at most and the increase is up to 9.5 percentage points for females. No effect is found for the second reform.

In line with Campolieti and Riddell (2012), in a recent paper Kostøl and Mogstad (2014) investigate the effectiveness of incentives aimed at inducing DI beneficiaries to return to work. In 2005, a programme was introduced in Norway that allows DI recipients to keep a portion of their benefit if they return to work. Only a subgroup of beneficiaries were affected by the programme, in particular those awarded DI before 2004. Kostøl and Mogstad use a regression discontinuity design comparing DI recipients awarded just before and after January 1 of 2004. They find that there is substantial working capacity left among DI recipients: the labour force participation rate increased by 8.5 percentage points three years after the programme introduction for those aged 18 to 49, however no effect is found for DI recipients approaching retirement age. Moreover, earnings and to some extent disposable income increased for DI recipients and, at the same time, the costs of the programme decreased, mainly as a result of a reduction in the benefits paid.

Ruh and Staubli (2015) investigate the role of earnings cap in discouraging work among DI beneficiaries using Austrian data. The cap creates an incentive to keep earnings below the threshold to retain full benefit eligibility. The lower the cap the lower the incentive to participate in the labour market, the higher the cap the higher the generosity of DI programme and thus the incentive to enter the programme. They exploit the discontinuity generated by the earnings cap on the implicit tax rate to investigate the earnings elasticity with respect to financial incentives. They find a large earnings response to the cap but modest elasticities driving the behavioural response. They estimate an elasticity of earnings with respect to the implicit net-of-tax rate of 0.17.

Exogenous variations in DI programme eligibility rules of Sweden, the Netherlands and Austria have been investigated by Karlström *et al.* (2008), de Jong *et al.* (2011) and Staubli (2011). Karlström *et al.* (2008) investigate the effect of 1997 abolition of DI eligibility rules specific for individuals aged 60 to 64 in Sweden. Before the abolition, older workers were allowed to receive the benefit for pure labour market reasons or had to pass a less stringent medical assessment. Applying a difference-in-difference approach using as control group those aged 55 to 59, they find no effect on labour supply of the 1997 reforms at least up to three years after the policy change. However they find supportive evidence that other insurance mechanisms act as substitutes to DI for those not eligible for the benefit under the post-reform rules.

Increasing the level of screening can be an instrument to reduce moral hazard problem in insurance programmes. de Jong *et al.* (2011) investigate to what extent this instrument can be used to control the inflow into DI and in particular the duration of benefit receipt, i.e. the outflow from DI programme. They use data from a controlled experiment in the Netherlands. Examiners were instructed to screen more strictly the reintegration reports in 2 out of 26 regions. The results are twofold: stricter evaluations cause a decrease in the duration of sickness absenteeism and reduce DI applications. According to the authors, stricter screening procedures reduce the attractiveness of DI programme and trigger both self-screening of potential DI and resumption during sickness absence.

In Austria, high rate of disability receipt for individual approaching retirement is mainly due to a relaxation in DI eligibility criteria starting at age 55. On September 1, 1996 a reform increased the age at which conditions for DI benefits are relaxed, from 55 to 57. Staubli (2011) exploits this policy change and finds that the share of disability beneficiaries aged 55-56 decreases by 6 to 7.4 percentage points and employment for the same group of individuals increases by 1.6 to 3.4 percentage points. However, the share in unemployment increases by 3.5 to 3.9 percentage points and the share in sickness insurance by 0.7 percentage points. As in previous works, the effects are particularly high among low-skilled, unhealthy workers.

A comprehensive picture of the effect of DI benefit generosity on labour supply is given in Mullen and Staubli (2015). They use a unique administrative database covering labour market histories of Austrian individuals starting from 1972 and they exploit exogenous variations generated by a series of reforms on DI and old age pension system during 1990s and 2000s in Austria. They find that a 1 percent increase in DI benefit generates an overall 1.2 percent increase in the DI claiming rate and a 0.7 percent increase in the period from 2004 to 2010 in which replacement rates were particularly low. As in previous research, those more responsive to benefit amount are low-skilled workers at the beginning or at the end of their working careers. To help comparability with the extensive literature using US data, it is important to notice that in Austrian DI programme it is not necessary to be out of the labour market to apply for the benefit, the applicant is required to withdraw from his work only if he is awarded the benefit.

1.3.3 Effect of economic conditions

As pointed out in the OECD report (OECD, 2010) and widely recognised in the literature, in periods of economic downturns there is a tendency of a 'medicalisation' of labour market problems. This tendency together with the expansion of DI programmes, make it difficult to distinguish the impacts of increasing demand from the impacts of increasing supply of benefits. Autor and Duggan (2003) try to address this issue. They look at the interactions among a liberalization of the disability determination process occurring in 1984, the increased replacement rate and the declining demand for low-skilled workers from the late 1970s to the late 1990s. They instrument the supply of disability benefits using the progressivity of the DI benefit formula: the formula did not change in the period considered, however replacement rate changed a lot due to increased inequality causing a much higher replacement rate for low-earners. The demand for benefit is instrumented using state level labour demand shifts. Under these assumptions, they estimate that high school dropouts became almost twice as likely to exit the labour market in response of an adverse shock following 1984 liberalization. In the aggregate, the increase in DI generosity caused a reduction of half a percentage point in the unemployment rate among adults aged 25-64 due to changes in labour force behaviour of low-skilled workers.

Duggan and Imberman (2009) use the same framework as Autor and Duggan (2003) and conclude that the increased replacement rate, resulting from increased income dispersion between 1984 and 2002, can account for 28% and 24% of the increase in DI receipt among females and males respectively.

To investigate the relationship between economic conditions and programme participation, Black *et al.* (2002) use exogenous shocks to the value of labour market participation caused by variations in the price of coal during the 1970s. While most regions experienced a decline in economic activity, in regions with coal employment and earnings grew as a result of price increase. The scenario was then reversed in the 1980s. They estimate an elasticity of DI payments with respect to local earnings of -0.4. They conclude that permanent variations in labour demand has larger effect on DI than transitory changes.

1.3.4 Estimates based on dynamic models

The papers surveyed in previous sections use different reduced form approaches to quantify the labour supply effect of DI generosity. Some limitations of the results presented are worth noting. Estimates that compare denied and allowed applicants in the US identify causal effects only for those at the margin of programme entry and cannot be extended to the entire population of DI beneficiaries (Chen and van der Klaauw, 2008; Maestas *et al.*, 2013; French and Song, 2014). Exogenous variations in the DI structure over time offer a more comprehensive view of the average effects of the specific reform being analysed, but only in the case in which the reform affects the entire population. Some of the papers presented, however, exploit reforms that affect subgroups of the population (Karlström *et al.*, 2008; Staubli, 2011; Kostøl and Mogstad, 2014). A more general issue is that the policy maker might want to evaluate alternative reformed scenarios, different from those experienced in the past or by other countries with similar institutional characteristics. A structural approach requires to formulate and estimate a dynamic model of individual behaviour under a set of assumptions and allows to perform policy experiments to evaluate alternative welfare programme structures. Moreover, a reasonably rich model allows to account for the interaction with other welfare programmes. In this section I briefly review papers that apply a dynamic structural framework to estimate, among the other things, the effect of DI benefit structure on labour supply.

Halpern and Hausman (1986) were the first to estimate a dynamic two-period model of the US application process in which eligibility is uncertain. They find that changes in benefit level affect the application probability more than changes in the acceptance rate.

Kreider (1999) uses a lifetime framework to extent previous models including uncertainly on future income flows in addition to uncertainty about award probability. He finds that one-third of the decline in labour force participation of males between late 1960's and early 1970's in the US can be explained by the increase in DI benefit level and that the waiting period induces a substantial self-screening among applicants. Using a similar framework, Kreider and Riphahn (2000) show that men and women differ in their behavioural response to DI changes.

Burkhauser *et al.* (2004) use the retirement model framework, developed to study the financial incentives provided by the pension systems, to propose a model of the decision to apply for DI in the US. They explicitly model the timing of the application arguing that it is an important factor to understand the programme costs and caseload given that DI applicants, when awarded, tend to stay in the programme.

Retirement life-cycle models consider as main pathway to retirement public old age pension (Gustman and Steinmeier, 1986; Stock and Wise, 1990; Rust and Phelan, 1997; Gruber and Wise, 2004; French, 2005) with an increased attention on the role of health and wealth in explaining labour market transitions. A natural extension of these models is the introduction of DI as a possible choice. This allows to account properly for the health risk, partially insured by state provided DI. Disability insurance is also a relevant alternative pathway to retirement in many countries, and it is particularly important in the US institutional context studied in most of the structural retirement literature.

A first attempt in this direction is the work of Bound *et al.* (2010). They propose a model for the behaviour of older workers approaching retirement age accounting for the joint role of health, wealth and labour market participation. DI application is among the set of discrete choices that the individual can take in each period. The main contribution of their work is to account for the possible endogeneity of health status to labour market behaviour and to model health as a latent variable having self-reported disability as an indicator. They conclude that health has an important role in explaining earlier exits, however changes in the DI programme structure does not have large effects on the probability to apply for the benefit.

A specific feature of DI in the US is investigated in Benítez-Silva *et al.* (2011). They evaluate a proposed change in the DI programme that consists in a reduction of $1 in benefit every $2 in earnings above the earnings cap (substantial gainful activities) instead of a 100% reduction in benefit when DI beneficiaries' earnings exceed the cap. Using a calibrated life-cycle model they find that the proposed change, aimed at increasing benefit outflow, has negligible effects on both inflow and outflow from DI. However, under the assumption of lowered stigma cost of returning to work in the reformed scenario, the effect would be significantly higher.

More recently Low and Pistaferri (2015) recognise that disability applications in the US are becoming sizable not only among older individuals close to state pension age but also among younger individuals. They develop a life-cycle model in which the screening process of DI is modelled carefully and agents face several sources of risk: health shocks, productivity shocks unrelated to health, and labour market frictions. However, they abstract from the role of health insurance and in particular the health insurance related job lock, particularly important in the US institutional context. They use self-reported disability (whether the individual has a physical or nervous condition that limits the type of work or the amount of work he can do) to measure health, allowing for heterogeneity in disability levels (as in Bound *et al.*, 2010). The focus of their paper is on welfare and behaviours effects of several DI reformed scenarios in the US. As an evidence of external validity of the model, they present model implied elasticities of DI application and of non-participation to benefit generosity and compare them with elasticities estimated in previous reduced form papers. The former elasticity, that is 0.62, is close to the central value of the elasticities surveyed in Bound and Burkhauser (1999b), but shows large heterogeneity among disability levels and productivity types. The latter elasticity is at the bottom of the range of values found in the literature (Bound and Burkhauser, 1999b; Haveman and Wolfe, 2000) and again it shows heterogeneity among disability levels. Among the other results, they investigate the interaction between DI and other social insurance programmes, such as food stamps, showing that an increase in the generosity of food stamps reduces false applications to DI.

1.4 Welfare implications of disability insurance

In their authoritative survey paper, Bound and Burkhauser (1999b) emphasize the prevalent focus of DI empirical literature on the DI effect on labour force attachment, whether less attention has been devoted to the social and welfare values of DI programmes. Looking only at the effect of DI on participation means focusing on the costs of the programme without considering the effectiveness of the programme, i.e. its ability in redistributing resources towards individuals with higher marginal utility of income, given the programme costs.

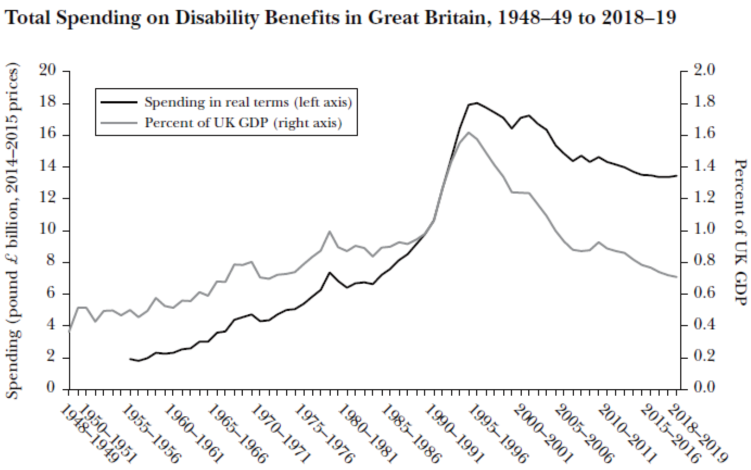
Following Gruber (2000), Bound and Burkhauser (1999b) do a back-of-the-envelope calculation to investigate whether the reduced labour supply caused by an increase in benefit generosity is offset by the increased insurance value. However, this calculation requires a set of assumptions difficult to justify. Other works (Meyer and Mok, 2013; Ball and Low, 2014) quantify the value of DI insurance using longitudinal data and looking at consumption drops that follow a deterioration in the health of individuals.

Few papers have tried to investigate both the effects on behaviour and on welfare of different policy reforms (Bound *et al.*, 2004, 2010; Benítez-Silva *et al.*, 2011; Low and Pistaferri, 2015). The most complete framework to study the incentive insurance trade-off is provided by Low and Pistaferri (2015). They use their estimated dynamic life-cycle model to conduct several counterfactual experiments. They find that a less strict screening process and a more generous benefit increase welfare. This happens because the effect of reduced false rejection offset the effect of increased false applications; moreover, the benefit of an increased level of insurance among the more disabled outweighs the negative effect of too high insurance among the less disabled.

1.5 Publicly provided disability insurance in the UK

Reducing public spending on disability benefit, targeted to individuals whose health limits their current ability to carry out paid work, has been one of the major goal of the UK government in the last decades. Figure 1.2 taken from Banks *et al.* (2015a) shows administrative data on the evolution of public spending on disability benefit. Public spending increases up to early nineties, when it reaches 1.6 percent of national GDP, and then progressively declines as a result of a major reform that came effective in 1995. Both the level of spending and the percent of GDP have remained higher with respect to the 60's and the 70's levels.

Figure 1.2: Public spending on disability benefit from 1948. Source: Banks *et al.* (2015a).



Source: Department for Work and Pensions, Benefit Expenditure and Caseload Tables, March 2014 (https://www.gov.uk/government/publications/benefit-expenditure-and-caseload-tables-2014).
Note: Figure shows spending on Sickness Benefit, Invalidity Benefit, Severe Disablement Allowance, income support on grounds of disability, Incapacity Benefit, and Employment and Support Allowance.

Much of the literature surveyed in previous sections refer to the US institutional context. As it will be clear in the discussion of the UK DI programme, two main differences emerge comparing the UK and the US programmes. The latter is characterized by a waiting period of several months for applications' examination and the DI benefit amount is based upon an average of past earnings: the calculation is based on the average indexed monthly earnings (AIME) as for other Social Security benefits. The former does not require a long assessment period and the benefit amount is not linked to past earnings, it is instead a rather low flat rate amount.

In the first part of this section, I briefly describe the disability benefit structure in the UK and I summarise the main reforms focusing on disability insurance. I will then survey the empirical economic literature on disability benefit provision in the UK and in particular empirical evidence on reforms effectiveness.

1.5.1 The UK disability reforms

Following Banks *et al.* (2012), disability benefits can be classified into work-related injury benefits, disability insurance, non-contributory benefits and means-tested benefits.

The first to be implemented were the work-related injury benefits with the Workmen's Compensation Act in 1897. The aim of this category of disability benefits is to regulate and enact the employer liability to compensate the employee in case of work related accident or disease. It does not prevent to receive other health related benefits and it is not means-tested. It represents the most generous benefit among disability related benefits.

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

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

In line with the 1995 reform, the 1999 Welfare Reform and Pensions Act remarkably tightened eligibility conditions. Eligibility was tested with the Personal Capability Assessment, a health test aimed at fostering return to work. In addition, contribution requirements referred only to contributions paid in the last three years before the start of incapacity. Finally, the reform introduced a benefit cut and a means-testing with regard to private pension income: for private pension income exceeding £85 a week the benefit amount was reduced 'by an amount equal to 50% of that excess'.

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

Finally, in 2008 *Employment and Support Allowance* (ESA) was introduced for new claimants in place of IB. A Work Capacity Assessment, stricter than the previous health test, determines eligibility to the benefit and classifies claimants into able or unable 'to follow work related activities'. If classified as able to follow work related activities, individuals have to attend the Pathways-to-Work programme. From 2011 to 2014, existing IB claimants had been reassessed and those eligible moved to ESA.

For individuals not eligible for contributory benefits, in the 1970s a set of benefits to compensate the extra cost endured by disabled individuals was introduced. The current benefits are the result of the 1992 reform which introduced *Disability Living Allowance* (DLA) for those starting to claim the benefit before age 65. For those aged over 65, *Attendance Allowance* (AA), introduced in 1971, remained available.

Finally, means-tested benefits such as *Income Support* (IS) and *Working Tax Credits* (WTC) have specific premiums for disability. Details on these benefits are reported in the description of 2003/2004 Tax and Benefit system in Apprendix C.

1.5.2 Evidence from the UK disability reforms

The 1995 reform, which replaced IVB with IB, introduced a clear discontinuity in terms of public spending on disability and number of benefit claimants. Disney *et al.* (2003, 2006) argue that the major effects observed for males were due to benefit phased out for those aged above state pension age, whereas the number of claimants among females did not stop to rise. To investigate the effect of tightening the health eligibility condition on employment, they estimate a fixed-effects conditional logit model, accounting for the possible endogeneity of self-reported health to economic status. They reject the hypothesis that the reform raised the severity of the health shock required to enter disability.

In a more recent contribution, Banks *et al.* (2012) use the same data source (British Household Panel Survey) and estimate a probit retirement model controlling for health and a set of year dummies. They find a statistically significant difference between the pre and post reform year dummies, in particular their estimates suggest a higher labour market participation for males after the reform. They do not find significant effect when considering females.

Ball and Low (2014) look at the value of disability benefit for beneficiaries, accounting for other sources of insurance, such as self-insurance through savings, spouse economic status and own labour force attachment. Individuals experiencing work-limiting health conditions and having as only source of insurance the state provided DI suffer a 9% drop in consumption after the onset of disability. This effect is only partially mitigated by other forms of self-insurance. Comparing benefit claim on the onset of disability before and after the 1995 reform they find a higher drop in consumption as a consequence of the reduced benefit generosity but a better targeting of those in need, that are identified as those reporting no work limitations due to health problems while receiving IB.

At the time of Pathways-to-Work announcement by the Department of Work and

Pensions (DWP) in 2002, there were more individuals claiming IB then unemployment benefit. According to Adam *et al.* (2010), the majority of individuals entering IB say they expect to work again. Even if this was true for almost 60 percent of people receiving the benefit in 2004 who left IB within a year, the remaining 40 percent tend to stay longer in the benefit. In particular in 2002, almost 30 percent were expected to be still claiming IB after other eight years (Adam *et al.*, 2010). The scheme was first implemented in some areas and subject to an evaluation process. Adam *et al.* (2010) report results in terms of the economic impact of the programme after the pilot implementation. Using a difference-in-difference approach they compare several economic outcomes for individuals in the piloted areas with claimants in similar areas. They find a significant positive and long lasting effect of the scheme on the probability of returning employed and a significant positive effect of leaving IB which vanishes in 12 months. The interpretation of these findings is that the impact is mostly on individuals who would have left the benefit in the first year in any case, but they would have stayed out of the labour market with a higher probability in the absence of Pathways-to-Work. The major effect seems to have been on women living with a partner who have suffered some physical incapacity.

Banks *et al.* (2015a) use administrative data, Labour Force Survey (LFS) and ELSA to summarise recent trends and heterogeneity in disability benefit receipt, focusing on the effect of the 2008 reform that introduced ESA. They report evidence of heterogeneous patterns across education groups, sex and age with lower education level showing larger drop in benefit participation and tendency of an increase in the number of claimants among the youngest low educated groups. One of the main trend has been the sharp decline in the number of claimants among those aged over 50. Using ELSA data and the detailed health information it provides, they show evidence of an improved targeting of the benefit on those more in need consistent with the decreasing trend, but this is true mainly for females. Comparing disability benefit receipt, employment and unemployment benefit receipt between 2008 and 2012 for different levels of disability (computed as number of declared limitations in a list of 11), they find no strong and conclusive evidence that the decline in disability benefit receipt is associated with increased employment and/or unemployment.

1.6 Summary

The relationship between health and labour supply is of great importance in economics. The design of effective policy interventions requires a deep understanding of this relationship for several reasons. For example, ageing population and pressure on the sustainability of pension systems call for a focus on older workers and their working capacity, given that the probability of adverse health shocks increases with age. In addition, a fundamental role in this relationship is played by DI programmes which are designed to insure workers against reduced working capacity and consequent loss of earnings capacity. In this survey I focus mainly on DI, the incentives/disincentives it provides to labour market participation and its insurance value.

Most of the literature on DI has been developed for the US institutional context which is characterised by remarkable application costs for the individual. The applicant must be out of the labour market, the evaluation takes several months, the rejection rate is high and the entire process is lengthened by the possibility to appeal and reapply. The characteristics of the application process in the US have been exploited in several works in particular to give a causal interpretation to Bound's results (Chen and van der Klaauw, 2008; Maestas *et al.*, 2013; French and Song, 2014). However, these estimated effects of DI on labour supply are for those at the margin of programme entry. The bulk of the literature exploiting exogenous variation in DI programmes over time or across states/regions enlarges the set of estimates of the elasticity of labour supply to DI generosity using data for different institutional contexts. Large heterogeneity emerges between impairment severity, high and low skilled workers, age, sex and institutional contexts.

These analyses mainly rely on the availability of structural changes in the DI programme and good data on benefit take-up and labour supply decisions (von Wachter *et al.*, 2011; Mullen and Staubli, 2015). Accounting for the interaction with other welfare programmes, evaluating the effect of alternative policy reforms and the long run effect on labour supply are difficult to achieve in a reduce form framework. A few papers propose rich life-cycle dynamic models and perform policy experiments to look at the effects on welfare and behaviour of different DI structures (Bound *et al.*, 2010; Benítez-Silva *et al.*, 2011; Low and Pistaferri, 2015).

A first research need that emerges from this survey is that of additional analyses on the social and welfare values of DI programmes, already emphasised by Bound and

Burkhauser (1999b) and still actual. Much of the literature has focused on the impact of DI programmes' generosity on labour force attachment and the implied efficiency costs, however the social benefits of the programmes in terms of equity and insurance values provided need to be considered. The recent paper by Low and Pistaferri (2015) shows the potential of structural analyses to enrich and extend this literature, and provides welfare effects of alternative policies.

A second important aspect in need of investigation is the change in the age distribution of DI beneficiaries. Both von Wachter *et al.* (2011) for the US and Banks *et al.* (2015a) for the UK highlight that the fraction of DI recipients among younger workers is increasing, with a large proportion of them reporting mental health problems. More research is needed to understand these trends among the younger in order to disentangle the effect of adverse economic conditions from trends in health.

Even if an increasing number of younger individuals applies for DI, the fraction receiving the benefit among older workers remains high in several countries. This is because often DI programmes are forms of early retirement. A third aspect in need of further investigation is to consider in a unified framework labour supply, DI application, private and public pension claiming decisions to understand how these welfare programmes interact and how reforms to each of them affect individual behaviour.

The way in which health is measured might play an important role in models that describe labour supply and DI decision, also in the light of a change in the conditions that lead to disability. Bound *et al.* (2010) and Low and Pistaferri (2015) adopt different strategies. The former work uses a latent variable for health, with self-reported disability status as an indicator, and self-reported disability is allowed to be endogenous to labour market behavior. The latter considers self-reported disability as measure of health. The effect of health on labour supply in empirical work is found to be largely influenced by the measure of health adopted (Lumsdaine and Mitchell, 1999). In the context of structural analyses of labour supply and health, existing studies have extremely simplified health processes to keep the model computationally tractable. A fourth undressed research need is to investigate the consequences of more comprehensive health measures.

Finally, as pointed out above, much of the research in this field refers to the US institutional context, in particular for what concerns the few contributions using dynamic life-cycle models. The labour supply effect and the welfare effect of changes in DI programme are affected by the institutional contexts in which they are evaluated that might differ in terms of labour market structures, health institutions and welfare systems. Considering the US institutional context, a crucial aspect not addressed in Low and Pistaferri (2015) is the potential effect of health insurance and related job lock in explaining the relationship between DI generosity and labour market outcomes. In the European institutional context out-of-pocket expenditures are not a crucial part of healthcare costs, as they are in the US (French and Jones, 2011; De Nardi *et al.*, 2010, 2013), due to the presence of national healthcare systems. As a consequence, a direction through which health enters the labour supply behavioural model can be ruled out. Moreover, in many other countries than the US there is no need to be out of the labour market when applying for the benefit (for example in the UK and Austria) resulting in a not straightforward extension of the results for the US.

In the analysis that follows, I address many of the issues outlined above. As opposed to previous research: (*i*) savings, labour supply, private pension and Incapacity Benefit take up decisions are considered jointly in a rich model formulation and interactions among pension, disability and means tested programmes are exploited. (*ii*) The model is used to perform several policy experiments by altering the parameters of the DI programme. The experiments exploit not only labour supply responses but also the distributional consequences of several policy reforms and the willingness to pay for having the programme in the benefit system. (*iii*) The analysis is conducted for the UK institutional context, characterized by an Anglo-Saxon benefit system but a welfare state similar to that of other European countries. I use ELSA data, a survey representative of the population aged over 50 that collects information on several important dimensions such as health, labour supply behaviour and pension and non-pension wealth. (*iv*) The detailed health information contained in ELSA are a crucial feature of this survey and allow to model health and health dynamic in a richer way than in previous studies.

One limitation is in ELSA data only individuals aged 50 or above are surveyed, therefore the analysis will abstract from DI participation when younger and focus on the behaviour of individuals approaching retirement age.

Chapter 2

Model development and calibration

This chapter presents the life-cycle model of savings, labour supply, private pension and Incapacity Benefit take up decisions for males aged 50 or above living with a partner. The agents face uncertainty on health developments, wage offer realization and life expectancy. Assets are defined at the household level and spousal income is modelled as a deterministic function of male's characteristics.

I focus on couples for several reasons. First, in ELSA data 70% of males above 50 live with a partner. Second, the parameters of means tested benefits and tax allowances are different for singles and couples, and pooling together those two groups involves the inclusion of an additional state variable in the model. Third, I select an homogenous sample of individuals, sharing the same insurance channels in facing health and income shocks: couples tend to be wealthier and less at risk of under-saving than singles, moreover married individuals (in particular males) tend to be in better health and to experience lower mortality rates than singles.

The model aims at reproducing individual decisions in terms of IB take up, in order to be used to simulate individual responses to alternative structure of the programme. IB is targeted to individuals in poor health with limited or no working capacity, and more importantly it provides a flat rate amount which represents a valuable replacement income only for low income individuals. Thus, individuals might have an incentive to self-insure against health risk through savings or to take private pension benefits at an early stage. This motivates the importance of modelling savings decision, private pension claiming decision and the possibility to claim Incapacity Benefit as a function of health in a unified framework. Bequest motives are introduced in order to match the pattern of wealth accumulation.

A formalisation of the household head maximization problem is presented in Section 2.1 and model solution in Section 2.2. Section 2.3 presents the data, with a particular focus on the health measure. Section 2.4 presents the three steps of model estimation: life-cycle profiles estimation, exogenous processes estimation and remaining structural parameters calibration. Results from the three estimation steps are reported in Section 2.5.

2.1 The model

I propose a model of lifetime decision-making in which a household head seeks to maximize his expected lifetime utility of the form

$$U(c_t, l_t) + E_t \left[\sum_{j=t+1}^{T+1} \beta^j \Pi^s(j-1, t) \left(\pi_j^s U(c_j, l_j) + (1 - \pi_j^s) b(a_j) \right) \right] \tag{2.1}$$

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 according to a bequest function $b(a_t)$ with a_t assets at time t. Let β be the discount factor, $\Pi^s(j, t)$ the probability of living to age j conditional on being alive at age t and π_t^s the probability of being alive at time t conditional on being alive at time $t-1$. The household head maximizes Equation (2.1) by choosing consumption c_t, hours worked h_t, whether to apply for IB d_t and whether to claim private pension p_t, subject to several constraints.

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

$$U(c_t, l_t) = \frac{1}{1-\upsilon}(c_t^{\gamma} l_t^{1-\gamma})^{1-\upsilon} \tag{2.2}$$

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

$$l_t = L - h_t - \phi_H(\hat{H} - H_t) - \phi_{P_t}\mathbb{1}(h_t > 0) \tag{2.3}$$

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 highest possible level of health $(\hat{H})$ minus the current level of health of the individual (H_t). Following French (2005) and French and Jones (2011), the cost of being in bad health enters the utility as a time cost and the same is true for the cost of participating in the labour market ϕ_{P_t}. The fixed cost of work is the only parameter allowed to vary with age, such that $\phi_{P_t} = \phi_{P_0} + \phi_{P_1} t$[1]. 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.

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

$$b(a_t) = \phi_B \frac{(a_t + K)^{(1-\upsilon)\gamma}}{1 - \upsilon} \tag{2.4}$$

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.

The health process is specified 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):

$$\begin{aligned} \log H_t =& \omega_H(age_t) + \theta_t + \eta_t \\ \theta_t =& \rho_H \theta_{t-1} + \nu_t^H \\ \nu_t^H \sim& N(0, \sigma^2_{\nu_H}) \\ \eta_t \sim& N(0, \sigma^2_\eta) \end{aligned} \tag{2.5}$$

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

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

$$\begin{aligned} \log w_t &= \omega_w(H_t, age_t) + \epsilon_t \\ \epsilon_t &= \rho_w \epsilon_{t-1} + \nu_t^w \\ \nu_t^w &\sim N(0, \sigma^2_{\nu^w}) \end{aligned} \tag{2.6}$$

[1]As in a number of studies among which French and Jones (2011) and Rust and Phelan (1997).

I assume that at time $t-1$ the worker knows ϵ_{t-1} but he only knows the distribution of the innovation ν_t^w.

Following French (2005), I do not directly model joint couple decisions 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 individual's age and after tax labour and pension income.

$$ys_t = ys(income_t, age_t) \tag{2.7}$$

I assume that marital status does not change over the period considered, neither for separation nor for death of the partner.

This set of assumptions has the advantage of keeping the model simple and at the same time accounting for the fact that the head of the household does not rely only on his own income. However, female labour supply has been shown to have an insurance role against permanent labour earnings risk within the household and female participation rates increase as a result of additional uncertainty (Attanasio *et al.*, 2005). Including spousal income as a deterministic function of partners' characteristics does not allow to explore and fully account for this insurance channel.

The probability of survive to period $t+1$ given that the individual is alive in period t, π_{t+1}^s, is a function of age and health in period t.

$$\pi_{t+1}^s = \pi^s(H_t, age_{t+1}) \tag{2.8}$$

Finally, the asset accumulation equation is of the form

$$\begin{aligned} a_{t+1} &= a_t + y(w_t h_t, ra_t, ib_t, pb_t, sb_t; \tau) + ys_t + dla_t \mathbb{1}(H_t \leq \bar{H}) + tr_t - c_t \\ a_t &\geq 0 \qquad \forall t \end{aligned} \tag{2.9}$$

where $y(\ldots, \tau)$ is after tax income; r is the real interest rate; ib_t is the IB amount, received if the individual has health below the threshold $\bar{H}_d$ and claims for it ($d_t = 1$); pb_t is private pension benefit, received starting form the year in which the individual claims for it; sb_t is state pension benefit; dla_t is Disability Living Allowance, received if health is below the threshold $\bar{H}$ (with $\bar{H} < \bar{H}_d$); and tr_t are non taxable transfers (such as Income Support and Pension Credit). The tax function, τ, and the modelled benefits are described in great details in Appendix C. I assume that individuals cannot borrow against future pension income and means-tested benefits ($a_t \geq 0 \quad \forall t$).

Each individual can be endowed with a private retirement plan. Two different plans are modelled: a Defined Contributions private plan ($pen = DC$) and no private plan ($pen = NO$). In principle, the pension plan can be endogenous as individual self-select into a particular type of job offering specific benefits (such as occupational pension funds). Given that I am considering individuals at the end of their working life for whom private pension membership is mainly a predetermined characteristic, the exogeneity assumption seems reasonable.

At each t, in addition to their private plan endowment, individuals observe their age, health status H_t, the amount saved in DC fund (q_t^{DC}), the State Pension accrual (q_t^{SP}) and the wage offer (w_t). They then choose whether to claim for private pension ($p_t = 1$), hours to work ($h_t \in [0, \bar{h}]$), whether to apply for IB ($d_t = 1$) provided that health is below a certain threshold $\bar{H}_d$ and consumption (c_t). IB can be claimed up to State Pension Age (SPA) which is 65 for males in the period considered. I assume that at age 70 everyone is retired.

Individuals face three sources of uncertainty: health uncertainty (see Equation 2.5), wage uncertainty (see Equation 2.6) and survival uncertainty (see Equation 2.8). I assume that the probability of surviving to age T+1 conditional on being alive at age T is zero ($\pi^s_{T+1} = 0$) and I set T equal to 90.

2.1.1 Disability benefits

In what follows I describe which health related benefits have been modelled and under which assumptions, with a particular focus on IB.

At each age between 50 and 64 individuals with health below a certain level can decide to claim for Incapacity Benefit, that is state provided disability insurance in force between 1995 and 2008 (the phase out started in 2008 for new claimants but terminates only in 2014 when the reassessment of all beneficiaries was completed).

According to the rules, IB eligibility is conditional on having paid enough contributions in the three years before the start of incapacity, however if the condition is not met, the applicant can still qualify for a means tested benefit (Income Support) of equal amount. I therefore assume that contributory requirements are always satisfied. Moreover, even if recipients might do some types of work within limits on weekly hours and earned income, I assume work is not allowed while receiving the benefit. This might be a limitation in a context in which there is a gradual phase out from the benefit that allows to continue to receive the benefit while going back to work, but this is not the case for the UK.

Even if IB is a contributory benefit, the amount of the benefit is flat: a lower amount is paid in the first 28 weeks and an higher amount after having passed the 'suitable work test'. In the model the decision period is one year, I therefore assume for simplicity that the annual benefit amount is fixed. In addition, I assume that the decision to receive the benefit is annual and benefit duration is one year, which means that IB receipt in each year is independent of the claiming status in the previous year. The main concern with this assumption is the model ability in replicating benefit persistence[2]. Even if the assumption of yearly benefit is rather strong, in Section 2.5.3 it is shown that the model is able to replicate 80% of the persistence observed in the data only through the persistence in health and wages.

The last set of assumptions concerns the application process. I assume (*i*) that claiming for IB is costless, (*ii*) that when an individual claims for the benefit he will receive it for sure (meaning that rejection rate is assumed to be zero), and (*iii*) that health is measured without error in the examination process. The first two assumptions are a direct consequence of the third one: if the examiner can get an error free measure of the applicant's health, the cost associated with the application is reduced to the cost associated with the medical assessment, which seems negligible, and only agents with health below the eligibility health threshold would apply for the benefit. Moreover, as in the literature the main cost related to DI application has been identified with the long waiting time out of the labour market for the final decision, assuming a costless application seems reasonable in the UK where it is not necessary to wait several months for receiving the benefit. I assume perfectly observed health in the examination process due to data limitations that make it difficult to reasonably identify the error made by the examiner in assessing the applicant's health. One important limitation of this assumption is that it does not allow to assess the effectiveness of the screening process but only changes in individual behaviours under stricter or more generous health eligibility thresholds.

[2] An alternative modelling assumption would be assuming that IB is an absorbing state, but this comes at the cost of preventing individuals to move in and out of the benefit, as it is indeed observed in the data along with patterns of long persistence in the benefit. Another possibility is to assume some sort of persistence, such as a cost of exiting the programme. While the first alternative does not add difficulty to the model, the second alternative requires the IB claiming status to depend on past decisions.

In addition to IB, among non-contributory benefits I include Disability Living Allowance and Assistance Allowance. I assume that the benefit amount is flat and is the same for both benefits (in the model formulation I named these benefits *dla*).

The last category are means-tested benefits and in particular Income Support, Pension Credit and Working Tax Credit. I assume that each entitled individual claims for the benefit. Further details on disability benefits and their model implementation are reported in Appendix C.

2.1.2 State pension

The state pension provision is of two different types: Basic State Pension (BSP) that is received if individuals have payed National Insurance contributions for at least a quarter of their working life and Second Tier State Pension (STSP) which is related to earnings history. For simplicity I assume that everyone is entitled to the full BSP. For what concerns STSP, at age 50 individuals start with an initial level of benefit entitlement[3]. If earnings (*earn*) are above the Lower Earnings Limit (LEL) and below the Upper Earnings Level (UEL), the initial amount is updated according to the following rule:

$$STSP_{t+1} = \begin{cases} STSP_t + \frac{LET*0.4}{(SPA-16)} & \text{if } earn \in (LEL, LET] \\ STSP_t + \frac{LET*0.4+(UET-earn)*0.1}{(SPA-16)} & \text{if } earn \in (LET, UET] \\ STSP_t + \frac{(earn-LEL)*0.2}{(SPA-16)} & \text{if } earn \in (UET, UEL] \end{cases} \tag{2.10}$$

where LET and UET stand for Low and Upper Earnings Threshold[4]. The state pension benefit is simply given by

$$sb_t = \begin{cases} STSP_t & \text{if} \quad age \geq SPA \\ 0 & \text{if} \quad age < SPA \end{cases} \tag{2.11}$$

I assume that each individual starts to receive state pension at SPA, there are in fact few people observed claiming the benefit after SPA even if there are no penalties and modest incentives in terms of benefit amount in postponing benefit receipt (see Appendix A).

[3]In the UK system workers can decide to contract-out the contributions paid to STSP, and to contribute instead to a private pension plan (often an occupational DB plan). In the model there is no distinction between contracted-in and contracted-out earnings related pensions. Individuals are endowed with an initial accrual in STSP and I assume they pay contributions according to their earnings. This might result in an overestimation of pension wealth in the case in which individuals are also endowed with a private pension fund to which they are assumed to contribute 3% of their annual earnings, as explained in Section 2.1.3.

[4]Values for the 2003/2004 tax year are reported in Table C.6 in Appendix F.

2.1.3 Private Pension

In the UK the sources of private pension wealth are mainly two: wealth from defined contribution (DC) pensions and wealth from defined benefit (DB) schemes. As explained more in details in Appendix A, a progressive shift from DB to DC has been observed in the UK as in the US. Therefore I assume that private pension wealth can only be in the form of DC pension funds. The amount in the fund (q^{DC}) in each period t depends on worker's (c_w) and employer's (c_e) contributions and on the rate of return (φ) of the fund. I assume c_w and c_e to be constant fractions of earnings and the rate of return to be deterministic. The DC pension wealth evolves according to the following formula:

$$q_{t+1}^{DC} = (1+\varphi)(q_t^{DC} + (c_w + c_e)w_t h_t) \tag{2.12}$$

Pension amount pb^{DC} depends on the accrued amount in the fund, the lump sum amount the individual decides to receive and on the annuity rates at the time of the annuitisation (r_{DC}). I assume that individuals purchase an annuity fixed in nominal terms, which is the most commonly bought. In addition, when the individual claims for the benefit he gets a fraction ls of the fund as a tax-free lump sum and he annuitises the rest $(1 - ls)$, net of administrative costs l. The tax-free lump sum can reach up to 25% of the pension pot. In the model I set minimum age to claim private pension to 55[5].

The rate of return of the fund is assumed to be deterministic, the annuity rate does not vary with age, the lump sum payment is a constant fraction of the amount in the fund and the amount to save in the retirement account is assumed to be a fixed fraction of annual earned income. The assumption of deterministic rate of return ignores an important source of uncertainty that is likely to affect individual retirement decision and this has to be considered when interpreting the model implications in terms of pension claiming age and interaction of pensions with other benefits.

The benefit at claiming age is given by $pb_t^{DC} = r_{DC}(1 - ls)q_t^{DC}(1 - l)$.

2.1.4 Dynamic programming problem

Having introduced the key ingredients of the model, in what follows I formalise the dynamic programming problem that individuals solve at each time period t. Let $V(X_t)$

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

be the value function at time t, with vector of state variables

$$X_t = (a_t, w_t, H_t, q_t^{SP}) \tag{2.13}$$

for those without a pension plan and

$$X_t = (a_t, w_t, H_t, p_{t-1}, q_t^{DC}, q_t^{SP}) \tag{2.14}$$

for those with a private pension plan, where p_{t-1} is a dummy variable taking value one if in $t-1$ the individual has already annuitized the private pension fund. I can write the value function as

$$V(X_t) = \max\left\{V^i(X_t)\right\} \qquad i = 1, \ldots, 6 \tag{2.15}$$

where the index i denotes the six possible discrete choice combinations: $i = 1$ if the individual is active in the labour market ($h_t > 0$), $i = 2$ if he is active and he claims for private pension ($h_t > 0$ and $p_t = 1$), $i = 3$ if the individual is neither active nor claiming a benefit ($h_t = 0$ and $p_t = 0$), $i = 4$ if he claims for IB ($d_t = 1$), $i = 5$ if he claims for private pension ($p_t = 1$) and $i = 6$ if he claims for both private pension and IB ($p_t = 1$ and $d_t = 1$). Depending on t and on the type of private plan the set of choice variables differs. In particular, for individuals without a private pension plan Equation (2.15) reduces to $V(X_t) = \max\left\{V^1(X_t), V^3(X_t), V^4(X_t)\right\}$. When $i = 1$[6]:

$$\begin{aligned}
V^1(X_t) = \max_{a_{t+1}, h_t} \Bigg\{ & U(c_t, l_t) + \\
& + \beta\pi_{t+1}^s \int_{H_{t+1}} \int_{w_{t+1}} V(X_{t+1}|X_t) dF(X_{t+1}|X_t) + \beta(1-\pi_{t+1}^s) b(a_{t+1}) \Bigg\} \\
\text{with} \quad & l_t = L - h_t - \phi_H(\hat{H} - H_t) - \phi_P \\
\text{s.t.} & \\
a_{t+1} = & a_t + y(w_t h_t, ra_t, sp_t \mathbb{1}(age \geq \text{SPA}); \tau) + ys_t + dla_t \mathbb{1}(H_t \leq \bar{H}) + tr_t - c_t \\
& \text{if } plan = NO \\
a_{t+1} = & a_t + y((1-c_w) w_t h_t, ra_t, sp_t \mathbb{1}(age \geq \text{SPA}); \tau) + ys_t + dla_t \mathbb{1}(H_t \leq \bar{H}) + tr_t - c_t \\
& \text{if } plan = DC
\end{aligned} \tag{2.16}$$

[6]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 cousumption c_t.

The budget constraint differs among pension types because DC plan holders are assumed to contribute a fixed fraction of their salary, c_w, to the fund.

With $i = 2$, there is no accrual in private pension which means that c_w is zero and the amount of the private pension benefit is added to taxable income. Furthermore, when $i = 3$, h_t is set to zero and the maximization is only with respect to savings.

When $i = 4$, the agent claims for IB ($d_t = 1$) and receives the taxable benefit ib_t (in the choice set only if $age <$ SPA):

$$\begin{aligned} V^4(X_t) =& \max_{a_{t+1}} \Big\{ U(c_t, l_t) + \\ &+ \beta\pi^s_{t+1} \int_{H_{t+1}} \int_{w_{t+1}} V(X_{t+1}|X_t) dF(X_{t+1}|X_t) + \beta(1-\pi^s_{t+1}) b(a_{t+1}) \Big\} \\ &\text{with } l_t = L - \phi_H(\hat{H} - H_t) \\ &\quad \text{s.t.} \\ a_{t+1} =& a_t + y(ib_t, ra_t; \tau) + ys_t + dla_t \mathbb{1}(H_t \leq \bar{H}) + tr_t - c_t \end{aligned} \tag{2.17}$$

When $i = 5$:

$$\begin{aligned} V^5(X_t) =& \max_{a_{t+1}} \Big\{ U(c_t, l_t) + \\ &+ \beta\pi^s_{t+1} \int_{H_{t+1}} \int_{w_{t+1}} V(X_{t+1}|X_t) dF(X_{t+1}|X_t) + \beta(1-\pi^s_{t+1}) b(a_{t+1}) \Big\} \\ &\text{with } l_t = L - \phi_H(\hat{H} - H_t) \\ &\quad \text{s.t.} \\ a_{t+1} =& a_t + y(pb_t, sp_t \mathbb{1}(age \geq \text{SPA}), ra_t; \tau) + ys_t + dla_t \mathbb{1}(H_t \leq \bar{H}) + tr_t - c_t \end{aligned} \tag{2.18}$$

Finally, $V^6(X_t)$ enters the maximization in Equation 2.15 only if $age <$ SPA and it results in a obvious combination of the previous two. After age 70 there is no uncertainty on future wages but only on future health, because I assume that individuals exit the labour market by age 70.

For those having a private retirement account the budget constraint at retirement age ($p_t = 1$ and $p_{t-1} = 0$) is slightly different from the one after retirement age. There is in fact the possibility of withdrawing up to 25% (lump sum - ls) of the amount in the account free of taxes, such that

$$a_{t+1} = a_t + y(ra_t, pb_t, sb_t; \tau) + ys_t + ls * q_t^{DC}(1-l) + dla_t \mathbb{1}(H_t \leq \bar{H}) + tr_t - c_t \tag{2.19}$$

2.2 Model solution

Individuals are heterogeneous with respect to state variables, in particular individual i has a vector of state variables X_{it}. Wages and health status will differ across individuals given different realization of wage and health shocks, however given the same age, wage, health status, asset level, retirement decision and pension accrual, different individuals will make the same decisions.

I denote preference parameters with $\vartheta = \left\{\beta, \nu, \gamma, L, \phi_H, \phi_P, \phi_B, K, \bar{H}_d\right\}$ and the parameters that determine the data generating process for the state variables with

$$\begin{aligned} \chi = \{ & r, \omega_H(age_t), \sigma_{\nu_H}, \sigma_\eta, \rho_H, \omega_w(H_t, age_t), \sigma_{\nu_w}, \rho_w, \ \ldots \\ & \ldots \{ys_t\}_{t=1}^{T}, \{\pi_t^s\}_{t=1}^{T}, \left\{pb_t^{DC}\right\}_{t=1}^{T}, \{sb_t\}_{t=1}^{T} \Big\} \end{aligned} \tag{2.20}$$

The model is solved starting from period T and going backward. 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 the expectation with respect to shocks in health and wages and with respect to mortality risk. I integrate the value function with respect to the transitory component of health, η_t , using three-node Gauss-Hermite quadrature (see Judd, 1998). To capture uncertainty over the persistent components of health and wages, I convert θ_t and ϵ_t into discrete Markov chains, following the approach of Tauchen (1986).

2.3 Data

The data used are from the English Longitudinal Study of Ageing (ELSA). ELSA started in 2002 and it is a biennial longitudinal survey representative of English private household population aged 50 and over. Six waves of data are currently released. It is comparable to other notable ageing surveys such as the Health and Retirement Study (HRS) for the US and the Survey of Health, Ageing and Retirement in Europe (SHARE). ELSA contains detailed information on assets, both financial and property wealth, pension fund membership and accrued rights to private pensions, out-of work benefit receipt as well as earnings. It also contains detailed information on health status, both subjective and objective.

The information I need in order to compare model simulations with the data are participation decision, hours worked, pension and non-pension assets, health and Incapacity Benefit rates for males living with a partner. Pension wealth, and in particular accrual in state and private pensions is not directly reported by the respondents. They are asked about the amount in the DC funds they are currently contributing to. To have a comprehensive figure of their private pension wealth, however, it is important to consider all types of private pensions, including personal private pension (non-occupational pensions) and DB pensions. To recover a measure of private and state pension wealth I use the pension wealth derived variables released together with raw data for each wave of ELSA[7]. Pension wealth derived variables include the present discounted value of future or current pensions under some assumptions about earnings growth and participation decision up to state pension age. The procedure used to derive pension wealth variables is standardized starting from the second wave. For this reason I derive initial conditions for moment simulation from individuals interviewed in wave 2 and refresher of wave 3 born between 1946 and 1955 for which information on current economic status, wealth and health are not missing. I select only males living with a partner, i.e. only couples, and assume this status does not change over time. I do not consider self-employed. The final sample counts 776 individuals. Using this sample, I generate the joint initial distribution of assets, accrual, health and wage. Each of the 20,000 simulated individuals receives a draw of assets, accrual, health and wage from the initial distribution.

2.3.1 Measuring health

Health is a multidimensional concept and can be characterized by a large set of indicators, covering different health dimensions. However, researchers are interested in having a synthetic measure to work with. In the literature several health indices have been proposed. The simplest and most used in economic analysis is self-reported general health (SRH). It is collected in most longitudinal surveys and it is typically based on a five-point ordinal scale rating health from 'excellent' to 'poor'. This simple measure is generally strongly correlated with objective measures of health, though it has been criticized because of some limitations: it is subject to measurement error (Crossley and Kennedy, 2002), it suffers from heterogeneity in health perception (Lindeboom and van Doorslaer, 2004) and, finally, it has been proved sensitive to justification bias (Bound *et al.*, 1999). To overcome at least some of these potential problems, other more objective measures such as self-reported ADLs, IADLs and functional limitations have been

[7]The derivation of accruals is described in Appendix E.

used in economic analysis. On top of that, several continuous measures that use a set of (objective and subjective) indicators to recover 'true' latent health have been proposed. Among the others Meijer *et al.* (2011) , Jürges (2007) and Poterba *et al.* (2013), hereafter PVW. 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 aspect of health is of interest.

The continuous measure of health I use is constructed using the rich set of health indicators available in ELSA[8]. In particular I follow PVW approach and I apply principal component analysis to a set of dummy variables covering several dimensions of individuals' health. The weights associated with the first principal component are then used to construct the index.

In selecting which health indicators to use I try to have only objective measures of health in order to limit the three main limitations listed above (measurement error, differences in health perception and justification bias) and therefore I do not include SRH[9]. I select health indicators to replicate the set of health dimensions asked in the *Work Capacity Assessment* (WCA), which is the DWP's method of determining a person's ability to perform any type of work for state IB purposes. The WCA is a measure of the extent to which a person is incapable of performing certain specified everyday activities laid down in legislation. These activities cover physical and sensory functions as well as mental function. The latter consists in performing four activities in presence of mental health problems: daily living, completion of tasks, coping with pressure and interaction with other people. To account for mental health in the health index I include the score obtained by the interviewees in the CESD scale (Center for Epidemiologic Studies Depression).

[8]French (2005), for example, uses a dichotomous health variable. He suggests this might be one of the reasons why his model fails to match labour supply decline by health levels.

[9]However these measures, even if they are arguably more objective than SRH, are not immune to biases.

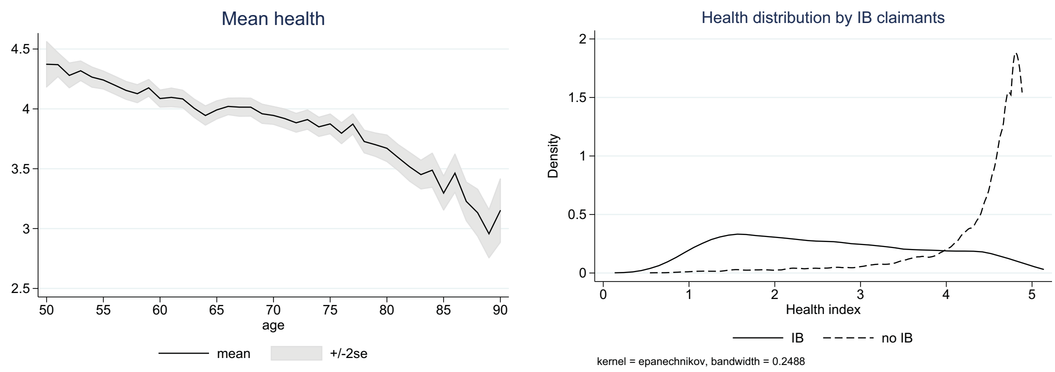
Table 2.1: Variables used for the health index.

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

The complete list of indicators used to construct the index is reported in Table 2.1, together with descriptive statistics for each indicator. In addition to indicators capturing physical, sensory and mental functions (tested in WCA), I add a set of variables included in ELSA questionnaire relevant to better capture the severity of limitations: limitations with ADL and IADL (dummy variables that take value one if at least one limitation is reported), any pain (one if the individual report to suffer from pains) and if the individual is receiving help. Finally, given that in the model I am interested in capturing not only disability but a comprehensive measure of health, in the spirit of PVW index, I add a set of indicators for diagnosed conditions.

Figure 2.1: Health index distribution by age and IB receipt (males younger than 65).

(a) (b)



The health index seems to capture well a comprehensive measure of health. As expected it is decreasing with age (Figure 2.1a) and the distribution is remarkably different among those receiving or not IB (Figure 2.1b).

If I restrict the set of indicators to physical function, sensory function and mental health, i.e. only the dimensions explicitly covered by the WCA, the index does not perform differently in distinguishing between individual receiving or not IB.

2.4 Model Estimation

Model estimation consists of three steps. In the first step I estimate the life-cycle profiles of assets, participation, hours worked and IB rate accounting for cohort and health effects. Second, I estimate exogenous processes of health, wages and survival probability. Third, I fix some parameters to values estimated in the literature and calibrate the remaining structural parameters to moments from the data life-cycle profiles.

2.4.1 First step: estimation of life-cycle profiles

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

Taking as an example hours profile, I regress log hours, $log(h_{it})$, on an individual specific effect f_i, age dummies, a full set of family size dummies and unemployment rate U_t proxing for aggregate time effects. When considering as outcome participation, age dummies are interacted with health.

$$\log(h_{it}) = f_i + \sum_{s=50}^{S} \pi_s 1\left\{age_{it} = s\right\} + \sum_{k=1}^{K} \pi_k 1\left\{size_{it} = k\right\} + \pi_U U_t + u_{it} \tag{2.21}$$

This specification allows to estimate age parameters (and in case of participation age parameters conditional on a certain level of health) accounting for individual fixed effect, time effect and family size effect. I derive an estimate of f_i computed as

$$\hat{f}_i = T_i^{-1} \sum_t \left(\log(h_{it}) - \sum_{s=50}^{S} \hat{\pi}_s 1\left\{age_{it} = s\right\} - \sum_{k=1}^{K} \hat{\pi}_k 1\left\{size_{it} = k\right\} - \hat{\pi}_U U_t \right) \tag{2.22}$$

with T_i number of waves in which individual i participate into the survey. I then regress the $\hat{f}_i$ on a set of ten years cohort dummies[10] to get the conditional expectation of f_i for a specific cohort, $E\left[f_i | cohort = c\right]$. When simulating the data profile I replace the individual effect f_i with $\tilde{f}_i = f_i - E\left[f_i | cohort_i\right] + E\left[f_i | cohort = c\right]$, I fix family size to two, and I set unemployment rate to 4.9%, which is 2004 annual unemployment rate for males in England[11].

The reference cohort c are those born between 1946 and 1955. By doing that my data profiles are representative of the same group of individuals used to set up initial conditions for model simulation.

2.4.2 Second step: estimation of exogenous processes

The exogenous processes for health, wage and mortality, are formalized in Equations 2.5, 2.6 and 2.8 of Section 2.1. In what follows I discuss the estimation of each process.

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

[11] Source: Labour Force Survey. ILO unemployment rate. (`http://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/timeseries/ycol/lms`)

Health process

The parameters of the health process to be estimated are the parameters of the deterministic component ($\omega_H(age)$), the variance of the persistent component ($\sigma^2_{\nu^H}$), the autoregressive coefficient (ρ_H) and the variance of the transitory component (σ^2_η). I first estimate the fixed effect regression in Equation 2.23 to get an estimate of the age parameters ($\hat{\pi_1}$, $\hat{\pi_2}$ and $\hat{\pi_3}$) controlling for time effects and family size effects.

$$\log H_{it} = \pi_1^H age_{it} + \pi_2^H age_{it}^2 + \pi_3^H age_{it}^3 + \sum_{k=1}^{K} \delta_k^H 1\{size_{it} = k\} + \mu^H U_t + \zeta_{it}^H \quad (2.23)$$

$$\begin{aligned}
\zeta_{it}^H &= f_i + \theta_{it} + \eta_{it} \\
\theta_{it} &= \rho_H \theta_{it-1} + \nu_{it}^H \\
\nu_{it}^H &\sim N(0, \sigma^2_{\nu^H}), \qquad \eta_{it} \sim N(0, \sigma^2_\eta)
\end{aligned}$$

Identification of the health shocks parameters comes from the error term ζ_{it}^H. I define the 'adjusted error term' as $g_{it}^H = \Delta\zeta_{it}^H = \Delta\theta_{it} + \Delta\eta_{it}$. The three parameters of interest are identified by the variance, lag one and lag two covariances (see Appendix B for details on moments derivation) and are estimated using standard minimum distance technique[12].

Wage process

The wage process is specified with a deterministic component that depends on age and health status, $\omega_w(age_{it}, H_{it})$, and a stochastic term ζ_{it}^w that is the sum of a persistent (ϵ_{it}) and a transitory component (ξ_{it}).

$$\log w_{it} = \pi_1^w age_{it} + \pi_2^w age_{it}^2 + \alpha_1^w H_{it} + \alpha_2^w H_{it}^2 + \sum_{k=1}^{K} \delta_k^w 1\{size_{it} = k\} + \mu^w U_t + \zeta_{it}^w \quad (2.24)$$

$$\begin{aligned}
\zeta_{it}^w &= f_i + \epsilon_{it} + \xi_{it} \\
\epsilon_{it} &= \rho_w \epsilon_{it-1} + \nu_{it}^w \\
\nu_{it}^w &\sim N(0, \sigma^2_{\nu^w}), \qquad \xi_{it} \sim N(0, \sigma^2_\xi)
\end{aligned}$$

I assume that ξ_{it} reflects measurement error. Fixed effect estimation of Equation 2.24 provides estimates for age (π_1^w, π_2^w) and health (α_1^w, α_2^w) effects on productivity. Even if fixed effect estimation allows to get rid of unobserved heterogeneity and thus of its potential correlation with the regressors, there might be problem of selection bias if

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

wage growths differ between workers and non-workers, given that only accepted wages are observed. To account for selection into participation I estimate Equation 2.24 using both accepted (observed) and offered (unobserved) wages, where offered wages for those not observed working are imputed as described in Appendix B.

As for the health process, identification of the productivity shock parameters comes from the error term ζ^w_{it}.

Mortality risk

I assume that the probability at time t of dying by $t+1$ is a function of age and health status in t. I first compute from the data the probability of dying by $t+1$ conditional on having a certain health level H in t, that is $Pr(death_{t+1}|H_t)$, controlling for cohort effects[13]. To do that I discretize the health measure in four categories, below the 10th percentile ($i = 1$), between 10th and the 20th percentiles ($i = 2$), between the 20th percentile and the median ($i = 3$) and above the median ($i = 4$).

The unconditional probability of dying by $t+1$, $Pr(death_{t+1})$ is then obtained as $Pr(death_{t+1}) = \sum_{i=1}^{4} Pr(death_{t+1}|H_t = i) * Pr(H_t = i)$ where the probability of a certain health level $Pr(H_t = i)$ is computed in the data, as for the conditional probabilities, controlling for cohort effects.

2.4.3 Third step: model calibration

Among preference parameters $\vartheta = \left\{\beta, \nu, \gamma, L, \phi_H, \phi_{P0}, \phi_{P1}, \phi_B, K, \bar{H}_d\right\}$, I fix relative risk aversion ν, discount factor β, time endowment L and curvature of the bequest function K to values estimated in the literature (see Table 2.2).

I set relative risk aversion of the composite good consumption-leisure ν to 2. Holding labour supply fixed and assuming that consumption weight γ varies between 0.4 and 1 (French and Jones, 2011), the coefficient of relative risk aversion for consumption is given by $\gamma(\nu-1)+1$, which ranges from 1.4 to 2, in line with values estimated in the literature (Blundell *et al.*, 1994; Attanasio and Weber, 1995; Banks *et al.*, 2001).

[13]The strategy to control for cohort effects is the same explained in Section 2.4.1 above.

Table 2.2: Calibrated preference parameters.

parameter	description	value
ν	relative risk aversion	2
γ	consumption weight	0.55
L	leisure endowment (hours)	4466
ϕ_H	cost of being in bad health (hours)	800
ϕ_{P0}	fixed cost of work at age 50 (hours)	450
ϕ_{P1}	age trend of the fixed cost of work (hours)	35
ϕ_B	bequest weight	6.2
K	bequest function curvature (£)	300,000
$\bar{H}_d$	eligibility threshold for IB (health level)	3.89
β	discount factor	0.9756
r	real interest rate	0.029

The discount factor β is set to 0.9756 as in Low and Pistaferri (2015) that use the central values of estimates from Gourinchas and Parker (2002) and Cagetti (2003). Time endowment L is fixed to 4,466 hours and the curvature of the bequest function K to £300,000[14] as in French (2005). Consumption weight γ is calibrated to 0.55. Higher values of γ mean stronger preferences for work, thus I calibrate this parameter in order to match participation. The time cost of being in bad health (in hours) is set to 800. I assume that the fixed cost of work has a linear trend in age $\phi_P = \phi_{P0} + t * \phi_{P1}$. I set ϕ_{P0} to 450 hours and ϕ_{P1} to 35, which corresponds to a time cost of 800 hours at age 60 and 1115 at age 69. These two parameters are in line with the estimates in French and Jones (2011)[15] in which the time cost at age 60 is 826 and at age 69 is 1315. The fixed cost of work and the fixed cost of being in bad health are calibrated to match the drop in hours worked and the drop in participation by health.

I calibrate the bequest weight ϕ_B to 6.2 to match the asset profile. Finally, I calibrate the health threshold to be eligible for IB at 3.89, which is the 30th percentile of the

[14] The value is obtained assuming an exchange rate pound to dollar of 1.65 and applying this exchange rate to the amount of $500,000 set in French (2005).

[15] French (2005) estimates the fixed cost of work for males' entire working career, whereas French and Jones estimate the fixed cost of work for older workers aged 60 to 69, which is a population more similar to the one considered in this work.

unconditional health distribution in the data, to match the fraction claiming the benefit by age. The health threshold to receive Disability Leaving Allowance varies with age and it is calibrated in such a way that the fraction of those receiving the benefit by age matches the data.

The rate of return on the safe asset r is set to 0.029, the average real return on UK Government liability between 2002 and 2008 (Capital, 2013). For what concerns benefits amounts, I set the amount for IB and DLA equal to the average amount received by males aged between 50 and 69 in 2004, that is £3,460 for IB and £3,000 for DLA [16]. The parameters of DC pension are calibrated as follows: contribution rates c_w and c_e are set to 6%; the rate of return of the fund (φ) is assumed deterministic and set to 7%[17]; the annuity rate r_{DC} is set to 4% in line with rates reported for males aged 65 in the UK compulsory market by Cannon and Tonks (2011); administrative costs are set to 10% as in Crawford and O'Dea (2016); and the lump sum payment ls is assumed equal to 15% of the pension pot.

2.5 Results

2.5.1 Decision profiles

Decision profiles for assets, participation, hours worked and fraction claiming for IB are estimated following the procedure presented in Section 2.4.1. To estimate the decision profiles I consider only data up to 2008/2009, which means only data up to wave 4. This sample selection is motivated by the fact that in 2008 there has been the introduction of ESA which gradually replaced IB. The health threshold under which individuals can claim for the benefit would be calibrated to match IB recipients up to 2008/2009 and thus would represent the eligibility condition before the introduction of ESA.

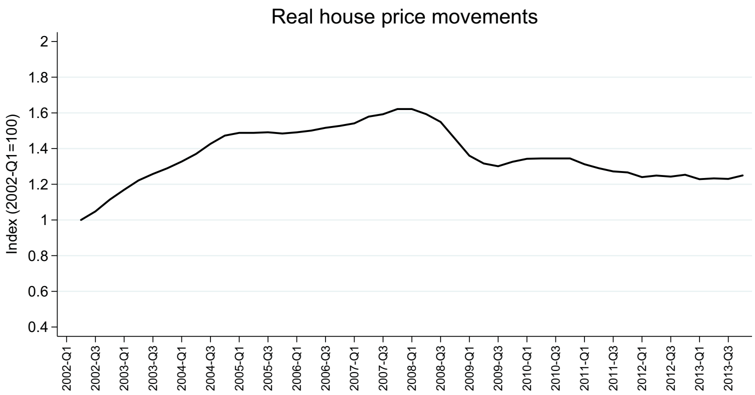
Assets are defined at the couple level and they include net financial and housing wealth. Pension wealth such as state pension accrual and private pension accrual are modelled separately. In the period considered, that is between 2002 and 2008, house prices have grown rapidly. In Figure 2.2, taken from Blundell *et al.* (2016), I report real house price movements in England from 2002 to 2013. Between the first two waves

[16] Source: amount data from DWP tabulation tool.

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

of ELSA, in 2002 and in 2004 respectively, house prices increase by 40%. A life-cycle profile using an asset measure that combines housing and non-housing wealth might be largely influenced by this positive shock in illiquid wealth.

Figure 2.2: Real house price movements in England form 2002 to 2013.



Note: Data from Blundell *et al.* (2016): English house price data is taken from the Land Registry and converted into real terms using the IPD.

In the model the mechanisms through which assets can increase are savings in non-pension wealth and contributions to pension funds. The latter are liquid assets, the former become liquid when the agent decides to annuitise non-pension wealth. The rate of return on non-pension wealth is calibrated to 2.7%, that is the average return of government securities in the period from 2002 to 2008. If the assets profile used to calibrate the model is not corrected for house price changes, than the assets increase observed in the data would be explained by savings in pension and non-pension wealth.

Nakajima and Telyukova (2012, 2016) propose and estimate a model of housing decision in which housing and non-housing wealth are modelled as separate state variables. They provide evidence of the importance of house equity in explaining different wealth decumulation patterns after retirement between homeowners and renters. Blundell *et al.* (2016) compare assets profiles in England and the US and attribute the slower wealth draw down in England mainly to the growth in house prices over the period, confirming the importance of housing as an asset in explaining wealth trajectories in retirement.

The scope of this work is not to investigate wealth trajectories in retirement distinguishing between housing and non-housing wealth, I am focusing on the end of the working life and I treat separately illiquid retirement wealth to capture financial incentives to exit the labour market provided by different wealth sources. The solution I implement to account for the price movements is to strip out house price changes by dividing net primary housing wealth by the house price index, using as reference year 2004, and to assume a price increase equal to the real rate of return used in the model. The underlying assumption is that house price increase, and the resulting wealth increase for homeowners, does not affect individual decisions in terms of consumption, retirement and labour supply. The corrected net primary housing wealth is added up to net non-housing wealth and used to estimate the asset life-cycle profile corrected for cohort effects. Concerns might arise on the consequences of this assumption on model predictions. Following De Nardi *et al.* (2010), a robustness analysis is to include in the model a time trend in the rate of return and allow for uncertainty in the rate of return realization by solving the model with independent and identically distributed (i.i.d.) interest rate shocks. Under these assumptions no corrections are required to housing wealth. Analyzing the robustness of model implications to the introduction of uncertainty in the rate of return is left for future research.

In the estimation of the assets profile I use 9,347 individual-year observations for 4,225 males aged from 50 to 70, interviewed in waves from 1 to 4. The resulting assets profile is reported in Figure 2.3a. Assets are slightly increasing from age 50 to 60 and almost flat after age 60.

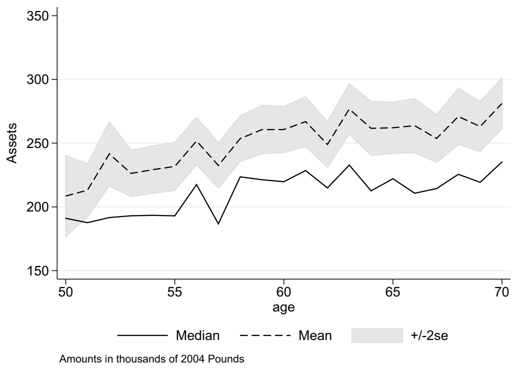
Figure 2.3c shows participation in the labour market conditional on health. The estimated profiles are obtained using 11,860 individual-year observations for 5,124 males observed in the first four waves of ELSA, aged between 50 and 75. In particular for each age the graph shows participation rates among individuals having health below the 20th quantile, between the 20th and the 30th quantile, between the 30th quantile and the median, and above the median. Health quantiles refer to the unconditional health distribution for individuals aged 50 to 90 in the data. As expected participation decreases with age and it is lower for individuals having a lower level of health.

Aggregate hours profile for those active in the labour market is reported in Figure 2.3d. In ELSA information on hours are reported as weekly hours worked in the current job. I infer annual hours using information on the date in which current job started and assuming weekly hours worked do not change over the period. To estimate the profile I use 4,904 individual-year observations for 2,388 males aged between 50 and 75, interviewed in waves form 1 to 4. Mean hours is almost flat about 2000 hours up to age

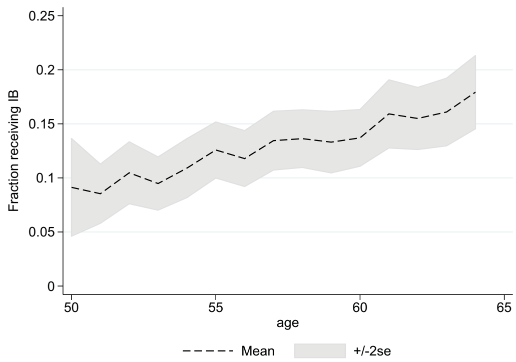
65 when it drops to 1000 hours.

Figure 2.3: Profiles of decision variables corrected for cohort and health effects.

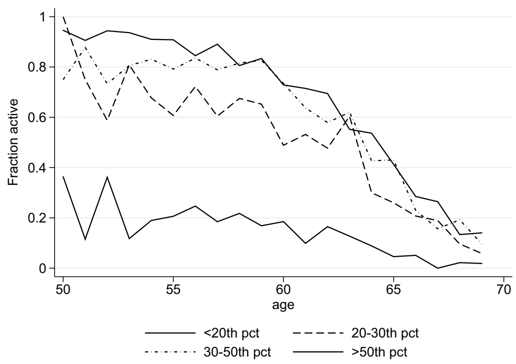
(a) Mean and median asset profiles.



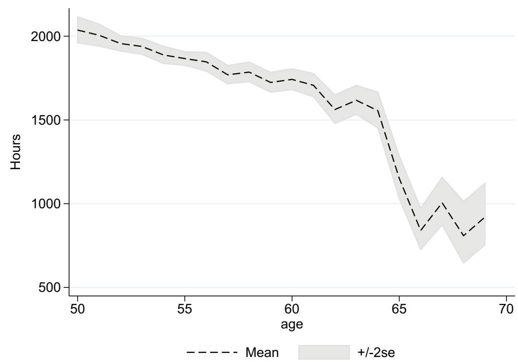
(b) Incapacity Benefit profile.



(c) Participation profiles.



(d) Hours profile.



Finally, Figure 2.3b shows the fraction of those claiming for IB. The estimation is performed using 6,853 individual-year observations for 3,243 males aged between 50 and 64. In what follows I indicate the state provided disability insurance with IB specifically for this reason. The profile is quite noisy up to age 52, it is increasing with age and for the cohort considered it reaches about 18% for those aged 64. After age 64, having reached SPA, it is not possible to claim IB anymore.

2.5.2 Exogenous processes

The estimation of the exogenous processes is carried out using the six waves of data currently available and not only the first four waves as for the decision profiles. The underlying assumption is that the introduction of ESA does not effect health developments, wage offer and mortality risk.

In Tables 2.3 and 2.4 I report parameter estimates for health and wage processes for the deterministic components and the stochastic components respectively.

Health process

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

Table 2.3: Parameters of the deterministic component of health and wage processes. (Fixed effect regression controlling for unemployment rate and household size. Standard errors in parenthesis.)

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

In the first column of Table 2.3 the parameter estimates of the third order polynomial in age are reported. Health is decreasing with age and the declining trend becomes steeper after age 70.

The error component model specified for health has a persistent AR(1) component and a transitory component. The autoregressive parameter estimate is 0.977 suggesting high persistence of the process (see Table 2.4). In a recent paper, 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 as health measure self-reported health and they 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.

Table 2.4: Parameters of the stochastic component of health and wage processes. (Standard errors in parenthesis.)

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

I include in the fixed effect regression family size and unemployment rate to control for time effects. In the simulation I set family size to two and unemployment rate to 4.9%. For what concerns unobserved heterogeneity captured by the fixed effect, I first recover the fixed effects as follows

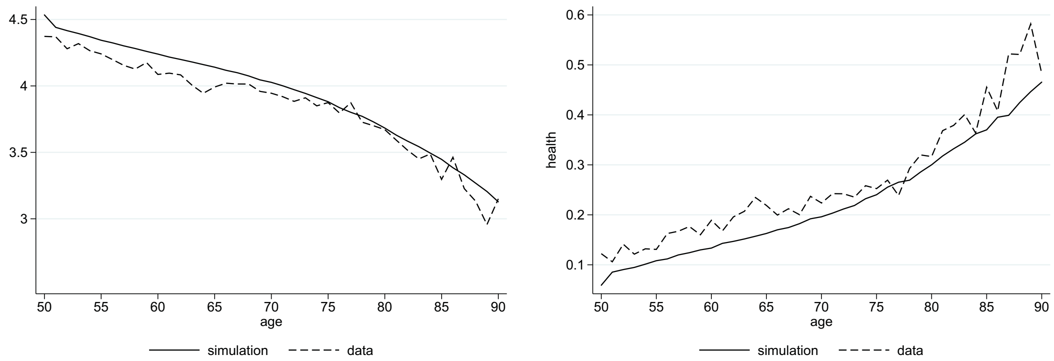
$$\hat{f}_i = T_i^{-1} \sum_t \left(\log H_{it} - \hat{\pi}_1^H age_{it} - \hat{\pi}_2^H age_{it}^2 - \hat{\pi}_3^H age_{it}^3 - \sum_{k=1}^{K} \hat{\delta}_k^H 1\left\{size_{it} = k\right\} - \hat{\mu}^H U_t \right)$$

with T_i number of waves individual i participated in. I then divide the distribution of f_i in two parts: below and above the first quartile, identifying low versus middle and high health levels at age 50. This heterogeneity captures that part of health that is assumed to be predetermined at age 50 and that does not vary with age. I solve the model for these two different 'types' of individuals, meaning that this binary variable enters the state space.

The process for health, specified as in Equation 2.23, is able to replicate quite well the observed evolution of health. Using as initial conditions the sample of males living with a partner and born between 1946 and 1955, I simulate 20,000 history of shocks for health and mortality risk (see the end of this section for details on mortality rates conditional on health) and compute health at each age. In Figure 2.4a I report mean observed (dashed line) and simulated (solid line) health, whereas in Figure 2.4b I report observed and simulated fraction of individuals with health below the first quintile of the unconditional health distribution.

Figure 2.4: Mean health distribution and fraction of individuals with health below the first quintile of the unconditional health distribution in the data and in the simulation.

(a) Mean health distribution. **(b)** First quintile of the health distribution.



Wage process

The second column of Table 2.3 reports parameter estimates for the wage process. Wages are increasing in health up to the first quartile of health (3.5) when the relationship becomes almost flat. They are increasing in age up to age 67 and decreasing afterwards. The estimated autoregressive parameter can be compared with wage process estimates obtained using ELSA linked with National Insurance data in Crawford and O'Dea (2016). They estimate couple wage processes for three different education groups and their autoregressive parameter ranges from 0.87 for low educated to 0.95 for high educated couples, where the education level is the one of the male. The parameter value

I obtain using only ELSA data is 0.864, at the lower bound of their range of estimations (see Table 2.4). They do not account for selection, however having earnings history data the problem of selection should be less severe. An additional aspect to be considered is that wage persistence can be different at the end of working life, i.e. after age 50. When using these estimates in model simulation, I set household size to two, unemployment rate to 4.9% and the fixed effect equal to the average fixed effect for individuals of the reference cohort (those born between 1946 and 1955).

Mortality risk

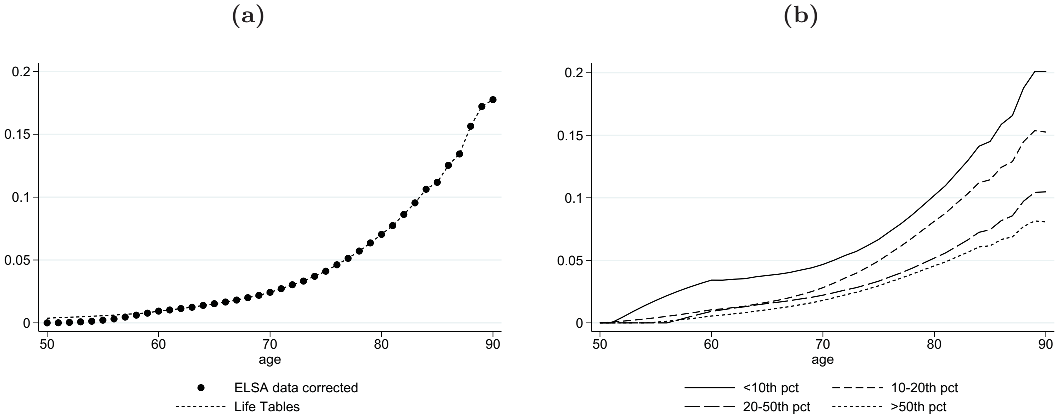
Mortality rates computed using ELSA data are lower than comparable mortality rates from the life tables (see Figure D.1 in the Appendix). This might be due to non-random attrition and/or initial selection into participation: older and unhealthier individuals might be more likely to exit the panel and healthier individuals might be more likely to enter the panel. In what follows I assume that mortality risks perceived by the individuals are consistent with the life tables and I correct mortality rates estimated from ELSA data accordingly.

In particular, mortality rates conditional on health - computed following the procedure explained in Section 2.4.2 - are those for males born between 1946 and 1955 living with a partner. However, individuals of that particular cohort are observed up to age 67 and extrapolation, used to predict mortality rates up to age 90, overpredicts mortality rates at older ages with respect to life tables. Therefore, the estimated probability of dying by $t+1$, $Pr(death_{t+1})$ - computed using estimated conditional probabilities $Pr(death_{t+1}|H_t = i)$ and estimated probabilities of being in each health category $P(H_t = i)$ controlling for cohort effect and household type - is re-scaled for each age in order to match life tables mortality rates.

In Appendix D.1 I report the results from a robustness analysis in which the aggregate mortality rates computed from ELSA data using the entire sample of males are used within the model, instead of the re-scaled conditional probabilities.

In Figures 2.5a and 2.5b are reported the aggregate mortality rate and the mortality rates by health level after having re-scaled the data profiles to match the life tables. From Figure 2.5a it can be seen that aggregate mortality rate is slightly underestimated at younger ages with respect to the life tables. Transition probabilities conditional on health, age and of being alive in t shown in Figure 2.5b are those implemented in the model to account for survival uncertainty.

Figure 2.5: Mortality rates computed using ELSA data and corrected to match life tables. Aggregate figure in panel (a) and by health level in panel (b).



2.5.3 Model calibration

The preference parameters are calibrated as reported in Section 2.4.3 in order to replicate assets, participation, hours worked and IB rate profiles. Figure 2.6 reports simulated versus data profiles for assets. Mean and median assets are most of the time within the confidence intervals[18].

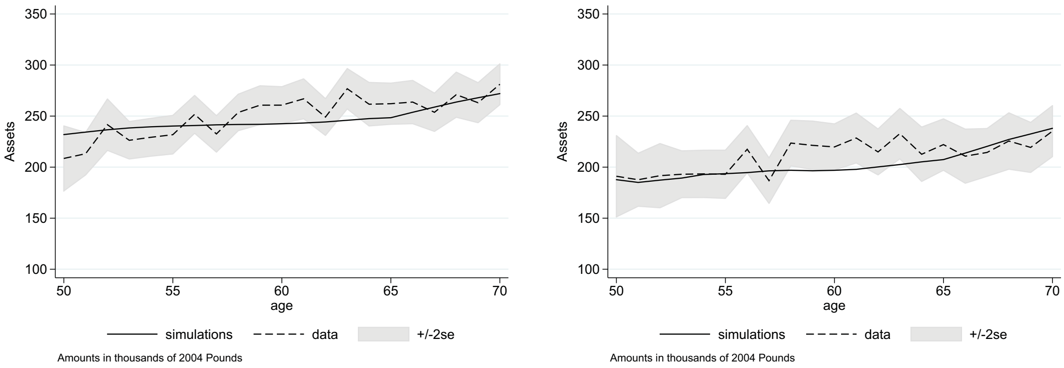
In Figure 2.7 I report participation and hours worked profiles. Aggregate participation is slightly higher than the one observed in the data for those younger than 60 (Figure 2.7a), and this is evident also in Figure 2.7c in which simulated participation is disaggregated by health levels. The model captures differences in participation for individuals with different levels of health quite well. This might be an advantage of using a continuous and comprehensive as well as objective measure of health.

The calibrated model replicates the profile for hours worked less precisely (Figure 2.7b): the average hours worked becomes flat at 1500 hours after age 65, whereas in the data at age 65 it sharply declines at 1000 hours per year. It seems that financial incentives provided by state and private pensions are not enough to explain the drop in hours worked at age 65.

[18]Simulated consumption profile is shown in Figure F.1, it is slightly decreasing up to SPA when it becomes almost flat.

Figure 2.6: Life-cycle profiles for mean and median assets. Simulations versus data.

(a) Mean **(b)** Median



The fraction of individuals receiving IB in the simulation is very close to data profile (see Figure 2.7e). In Figure 2.7f I report the estimated probability of receiving the benefit in t conditional on having received the benefit in $t-2$ in the data and in the simulations. The model is able to replicate 80% of the persistence observed in ELSA data even if no other direct mechanisms of persistence (such as costs of leaving the benefit) are modelled other than persistence coming from health and wages.

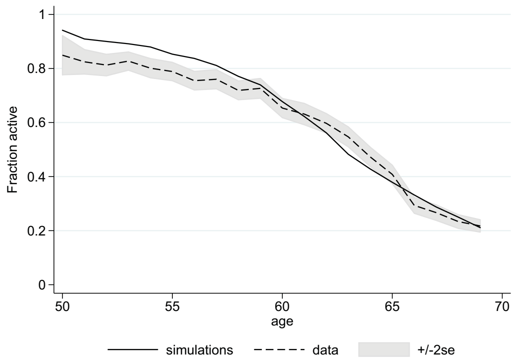
In addition to IB persistence, other facts not directly targeted in the calibration phase and well replicated by the model are worth mentioning. Looking at the assets distribution, the model replicates quite well the first tertile of the distribution, and slightly underestimates the second tertile (Figure F.3 in Appendix F). When conditioning on health (above or below the median), the model replicates quite well the heterogeneity in assets accumulation by health levels, however it underestimates the higher quantiles of assets for those in better health (Figure F.4b in Appendix F).

Finally, in the model individuals endowed with a private pension fund decide when to annuitize the amount in the fund. While in the model only accrual in DC plans is considered, in the data accrual can be both in DC and DB funds and in the initial distribution of the state variables DB accrual is converted into DC accrual[19]. The model generates a distribution of claiming age close to the one observed in the data when merging DC and DB fund holders (see Figure F.5 in Appendix F).

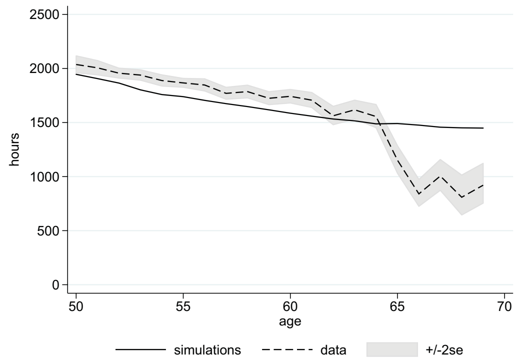
[19]Details are reported in Appendix E.

Figure 2.7: Life-cycle profiles of participation, hours and IB claiming decisions. Simulations vs data.

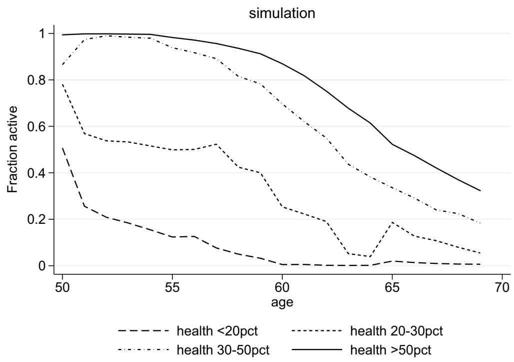
(a) Participation



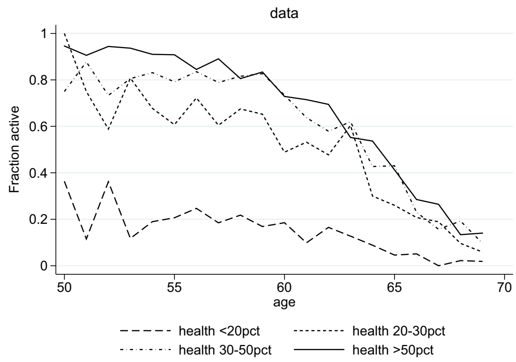
(b) Mean hours



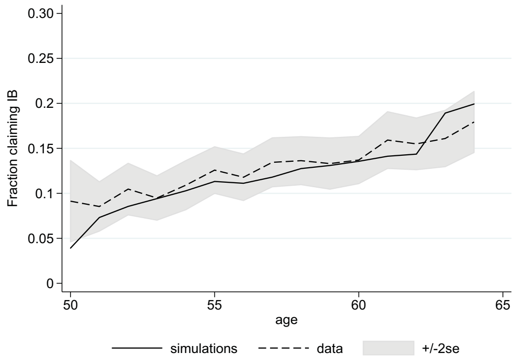
(c) Participation by health



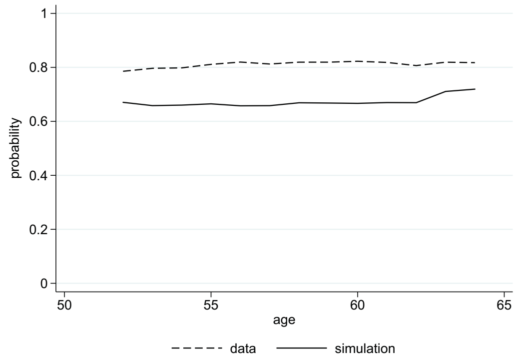
(d) Participation by health



(e) Incapacity Benefit



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



Chapter 3

The effect of policy reforms and the insurance value of Incapacity Benefit

The model presented in Chapter 2 describes and replicates some facts observed in the data, such as age profiles for assets, participation, participation by health, IB rate and IB duration. One of the advantage of a dynamic model of individual behaviour is the possibility to simulate individual decisions under different policy scenarios. In particular I am interested in understanding how changes in the structure of disability insurance programme affect labour supply decision and the distributional consequences of these changes. This exercise is of particular interest in the UK context, for which existing literature have found mixed evidence on the effect of recent disability benefit reforms (see Section 1.5.2 for a brief survey). Recent reforms act in two different directions, on the one hand the eligibility conditions in terms of health requirements have been tightened by introducing a stricter health assessment, on the other hand the return to work of claimants has been stimulated by providing help in finding a suitable work and financial support for the first year of receipt.

To better understand the consequences of such reforms on labour supply participation and income distribution, I perform several policy experiments. First, I eliminate Incapacity Benefit: this simulated scenario allows to investigate the disincentive provided by IB, its redistribution effect, its insurance value and the substitution effect with other benefits that top up income of low earners. The second policy experiment simulates individual responses to strengthening or lessening the health eligibility condition to receive the benefit. The health assessment is one of the main instrument advocated by policy makers to control benefit inflow and this is why understanding its labour supply and distributional effects is of particular interest. The third experiment investigates the effectiveness, in terms of controlling benefit inflow and outflow, of a policy intervention consisting in a set of measures to promote labour market inclusion of workers with health conditions limiting their working capacity. I assume that the mechanism through which this policy intervention affects individual behaviours is by reducing the fixed cost of work they face when entering the labour market. Finally, I compute non-participation elasticity and IB rate elasticity to benefit generosity by simulating agents labour supply and benefit take-up responses changing the amount of IB they would receive had they claimed for the benefit, and I compare these elasticities with previous results in the literature.

In the policy experiments no adjustments to ensure revenue neutrality are implemented. The underlying assumption is that revenue costs, whether positive or negative, resulting from changes in the tax and benefit system are offset by fiscal policies that affect only those younger than 50. Revenue neutrality implemented by adjusting other fiscal parameters, such as the tax rate on income, might generate confounding effects when looking at labour supply effects of DI. Moreover, I focus on a subsample of individuals, males aged over 50 living with a partner, and it is unrealistic to think that a reform to DI programme affects only this specific subsample of the population: in this context, it is not obvious how to implement a revenue neutral policy reform. In a robustness analysis in Appendix D.2 I recompute labour supply effects of removing IB when the total spending on IB generated by the model in the baseline scenario (total amounts paid as IB) is redistributed as an annual lump sum payment to everybody between age 50 and 64. The main findings hold in this case as well.

3.1 Incapacity Benefit removed

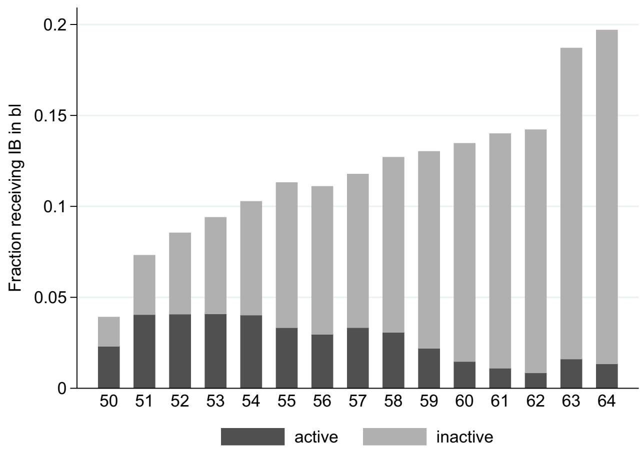
The first policy experiment investigates the effects of removing Incapacity Benefit on several outcomes, such as labour supply, disposable income, private pension claiming age, other benefits take-up and welfare. Considering the potential disincentive provided by IB, removing the benefit might result in an increase in participation if part of the beneficiaries have some working capacity left. However, in the absence of the programme those for which work is too costly lose a source of replacement income. This policy experiment allows to quantify these effects, accounting for the interaction between IB and other welfare programmes. Moreover, it allows to get a measure of the insurance value provided by IB.

3.1.1 Effect on labour market participation

Static perspective: individual-year observations if receiving IB in baseline

In Figure 3.1 I consider individual-year observations for those receiving IB in the baseline scenario and I show the fraction that become active in the reformed scenario (when IB is removed). According to this 'static' perspective, the fraction of those becoming active decreases with age. Employment rate is 47% for those aged 50 to 54, 25% for those aged 55 to 59 and 8% for those aged 60 to 64, in the absence of the programme. Overall, 22.4% of the people receiving IB in baseline would be active if the benefit is removed[1]. Considering individuals with health below the 10th percentile the employment rate is about 2%, whereas for individuals with health close to the eligibility threshold (between 20th and 30th percentile) the employment rate would be 41% in the absence of the programme.

Figure 3.1: Participation in the reformed scenario for those receiving the benefit in baseline.



Note: The height of the bars is the fraction receiving IB in the baseline scenario by age.

[1] Recall that employment rate is zero in baseline because IB recipients are prevented to work.

The literature comparing denied and allowed applicants in the US provides estimates on how large participation rate would have been if beneficiaries had not received the benefit. Considering in particular the causal estimates in Maestas *et al.* (2013) and French and Song (2014), the compared groups have been both out of the labour market for several months due to the requirements of being in unemployment in order to apply for DI, this means that the waiting period is likely to have deteriorated their human capital and their work readiness. Moreover, the identifying assumption allows to quantify the effect of DI on participation for those individuals that are at the margin of programme entry, i.e. those with health close to the eligibility threshold. In this framework Maestas *et al.* (2013) and French and Song (2014) estimate an increase of 28 and 26 percentage points in employment respectively, for those at the margin of programme entry (the effect is decreasing with age), which are estimated to be about 23% of total claimants.

In the UK institutional context there is no need to stay out of the labour market for several months waiting for acceptance or denial, thus to have a measure comparable to the one computed for the UK, it seems important to account for this additional aspect and compute employment loss due to DI with respect to the employment potential before benefit application. Indeed, Autor *et al.* (2015) suggest that previous results for the US have to be considered as a lower bound of the effect of DI on labour supply precisely because the long waiting period for application evaluation is likely to have a negative effect on both control and treatment. Accounting for this crucial aspect, they estimate a 48 percentage points decrease in employment three years after the application for those at the margin of programme entry aged between 18 and 64.

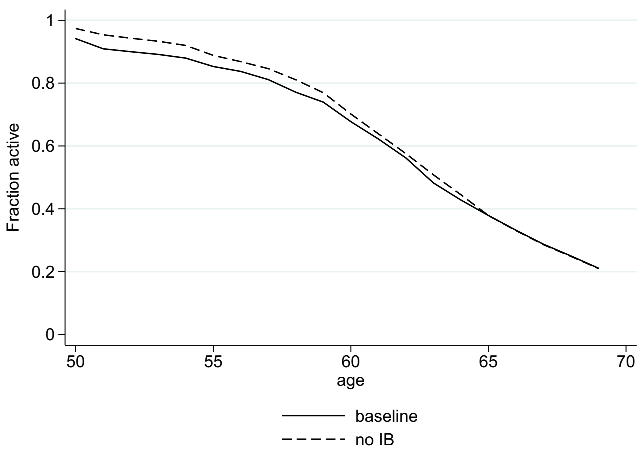
The policy experiment I propose quantifies the effect of IB on participation in an employment rate 41 percentage points higher for those with health close to the eligibility threshold (which are about 27% of total claimants) and aged between 50 and 64. The two figures are remarkably close, even if one might expect a lower effect on participation of IB in the UK because of its low flat rate amount as opposed to a more generous earnings related social security benefit in the US. Autor *et al.* (2015) investigate heterogeneity in the effect by sex and age, results show a significant difference for those aged 50 to 59 (about 70 percentage points) and those aged 60 to 64 (about 10 percentage points, not significantly different from zero). From model simulation the percentage points reduction is 57 for the first age band and 23 for the second, again confirming a similar pattern in the two institutional contexts.

Dynamic perspective: behaviour over the life cycle

What is partially missing from Figure 3.1 is the possibility that labour supply behaviour changes even in years in which IB is not received and for those never receiving IB in baseline. This is because, in the absence of IB, agents would need to internalise the increased health risk and to intertemporally substitute consumption and labour supply in order to self-insure by saving or working more when in good health, with respect to the baseline policy.

Dynamic effects on labour supply imply that individuals may be affected by the policy change even if they do not receive IB benefit in the baseline. Figure 3.2 illustrates how aggregate participation increases as a result of IB removal, by age. Confirming previous findings, much of this effect is concentrated on individuals aged between 50 and 64, when participation increases by 4%.

Figure 3.2: Aggregate participation in baseline and reformed scenarios.



On average, 31% of this increase is due to changes in annual labour supply behaviour of individuals not receiving IB in baseline. This is shown in Figure 3.3. In addition to the increase in labour supply among those receiving IB in baseline (dark gray bars, already reported in Figure 3.1), two other patterns emerge: among those changing behaviour, 21% were inactive in baseline and become active in the reformed scenario; the remaining 10% were active in baseline but do not work in the reformed scenario. This is consistent with an increase in participation when younger and in better health to insure against the increased health risk, and a correspondent decline in participation when old, because of higher accumulated wealth.

Figure 3.3: Changes in labour supply between baseline and reformed scenario.

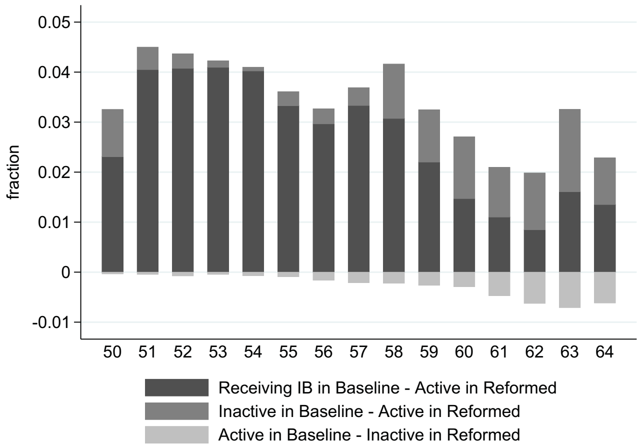
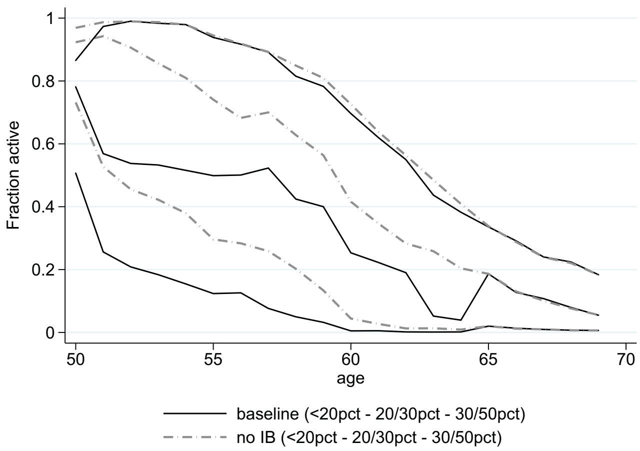


Figure 3.4: Participation by health level in baseline and reformed scenarios.

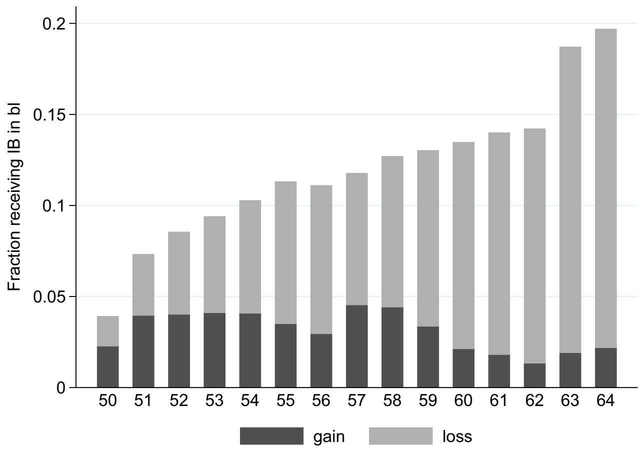


As expected, these effects are stronger for people in the lower health quantiles. In Figure 3.4 I report participation profiles in the two scenarios for those with health below the 20th percentile, between the 20th and the 30th percentiles, and between the 30th percentile and the median of the unconditional health distribution. The main effect is a delay in the age at which individuals with poor health exit the labour market. For example, those with health below the 20th percentile reach a participation rate of 20% about five years later, from age 53 in the baseline scenario to age 58 in the reformed one. The larger effect emerges for those at the margin of programme entry with health close to the eligibility threshold, that is equal to the 30th percentile of the unconditional health distribution.

3.1.2 Effect on disposable income

As for participation, I start from individual-year observations for those receiving IB in the baseline scenario. I compute disposable income both in the baseline and the reformed scenario for each of these observations and in Figure 3.5 I show the fraction gaining or losing in terms of within period disposable income. As for participation the fraction gaining in terms of within period disposable income is decreasing with age. At younger ages the gain is mainly due to labour income, whereas after age 60 is mainly due to increased pension wealth. Overall, 26.2% experience a positive change.

Figure 3.5: Disposable income in the reformed scenario for those receiving the benefit in baseline.



Note: The height of the bars is the fraction receiving IB in the baseline scenario by age.

To evaluate the 'dynamic' effect on disposable income of eliminating IB, I compare the present discounted value (PDV) of disposable income in the two scenarios. In Table 3.1 I report the distribution of the absolute difference between PDV in baseline and reformed scenario for the entire sample (upper panel) and for those who have received IB for at least one year in baseline (bottom panel). When considering the entire sample, the median is zero and the change is less than £10 in absolute value for 28.5% of the people. This means that about a quarter of the sample is not affected by benefit removal. Focusing instead on the subsample of those receiving IB for at least one year in baseline, the median of the change in PDV of disposable income is about -£1,166 and there is no concentration at zero. Those receiving the benefit in baseline are affected negatively (left tail) or positively (right tail) in terms of disposable income by benefit elimination. To give an intuition on whether the loss in PDV of disposable income is meaningful, I compute the PDV of IB assuming it is received each year from age 50 to age 64: the amount at age 50 is about £36,000. The 5th quantile of the change in disposable income in Table 3.1 is about -£22,400, that corresponds to the PDV of IB if the benefit is received from age 50 to age 56, or from age 55 to age 64.

Table 3.1: Distribution of the absolute change in present discounted value of disposable income between baseline and reformed scenario.

All sample						
5th	10th	25th	50th	75th	90th	95th
-11,392	-5,602	-712	0	291	5,548	10,535

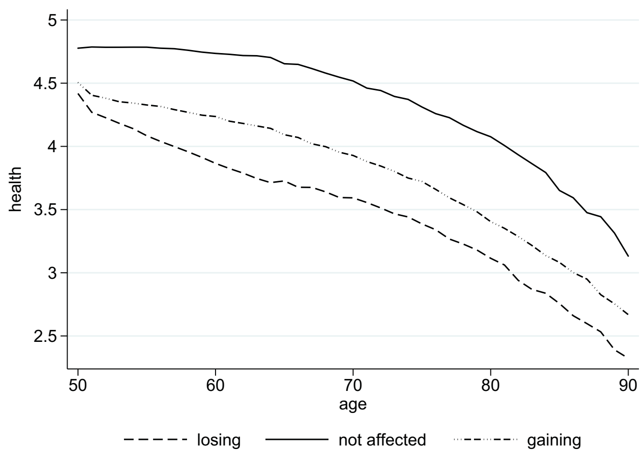
If receiving IB in baseline at least once						
5th	10th	25th	50th	75th	90th	95th
-22,396	-14,388	-5,470	-1,166	4,557	12,111	18,163

One might expect that a large part of these heterogeneous effects is due to differences in health levels and health patterns. In Figure 3.6 I report mean health by age for three groups of individuals: those unaffected by the reform, those gaining and those loosing in terms of PDV of disposable income[2]. Agents not affected by benefit removal are in good health and are less likely to experience negative health shocks. Those affected by benefit removal have on average lower health than the others. Among them, those in better health are more likely to become active and gain in terms of disposable income, whereas for those in worse health work is too costly and they end up losing in terms of disposable income.

[2]The same pattern is found when looking at the entire health distribution by age and comparing other moments such as the median or the 20th percentile.

A loss in PDV of disposable income among those in bad health suggests that, as expected, the most affected from IB removal are those to which the programme is targeted. To better investigate the distributional consequences of IB removal I report in Table 3.2 the distribution of average disposable income between age 50 and 64 (the period in which IB can be claimed) for the entire sample and for those with average health below and above the median between age 50 and 64.

Figure 3.6: Profile of mean health for those losing, gaining or not being affected by benefit removal in terms of PDV of disposable income.



The top panel suggests that major changes come from the bottom of the distribution: for example the 5th quantile is reduced by 4% whereas the 95th quantile is increased by less than 1%. As a result, inequality in the distribution slightly increases with IB removal. The second and the third panels, that split the sample among those with average health above or below the median, suggest that the effect observed in the aggregate distribution is mainly driven by those with a lower level of health. The statistics reported in Table 3.2 do not account for the variability within individual and over time. When the averages of disposable income between age 50-54, 55-59 and 60-64 are considered, heterogeneous effects among age groups emerge. At younger ages the negative effect of benefit removal can be seen only at the bottom of the distribution, whereas the other quantiles increase as a result of an increased labour supply. At older ages instead the entire distribution shifts towards lower values of disposable income.

The significant increase in participation after benefit removal comes with the cost of an increased inequality and a large drop in disposable income, in particular in period of low health and among older individuals. The distributional consequences presented in this section are on top of the income top up provided by means-tested benefits, which role is investigated in the next section.

Table 3.2: Distributional consequences of IB removal: average disposable income distribution between age 50 and 64 in baseline and reformed scenario.

policy	*average disp. income*							
	p5	p10	p25	p50	p75	p90	p95	Gini
	All							
with IB	11,217	13,271	17,153	21,173	25,031	28,295	30,299	0.154
without IB	10,807	13,086	17,075	21,165	24,996	28,295	30,328	0.157
	average health below the median							
with IB	9,970	11,254	14,295	18,538	22,662	26,017	27,989	0.171
without IB	9,473	10,832	14,250	18,445	22,556	25,995	28,011	0.176
	average health above the median							
with IB	16,237	17,501	20,133	23,280	26,630	29,635	31,599	0.114
without IB	16,226	17,499	20,112	23,284	26,646	29,642	31,599	0.114

3.1.3 Welfare analysis

Changes in the present discounted value of disposable income are not informative of the value of Incapacity Benefit. In order to evaluate the welfare effect of disability benefit removal, I calculate the amount of assets that an agent needs to receive at age 50 to be indifferent between the two scenarios.

Removal of IB is a loss for everyone, in expected terms, but has important distributional consequences which I illustrate reporting the results for different combinations of health and wealth levels in Table 3.3. To make amounts comparable across assets level, the compensating variation is expressed as a fraction of total liquid assets at age 50.

The average amount an individual would need to be compensated for benefit removal is about £6,400 or 3.2% of total liquid assets at age 50. The fractions of total assets reported in Table 3.3 suggest the presence of a large heterogeneity in the value of IB for individuals having different levels of health and wealth at age 50. On average, the lower the level of health and the level of assets at age 50 the higher the welfare compensation needed to be indifferent between the two scenarios. In particular, individuals with assets in the first quartile and health in the first quartile of the health distribution at age 50 would need 44% of their initial assets to be compensated for benefit removal, a fraction significantly higher with respect to that needed by individuals with the same wealth level but health in the second quartile (6%). Zooming on this group of individuals, the average 44% increase in initial assets translates in an average compensation of £8,000 at age 50.

Table 3.3: Average welfare compensation as fraction of total assets at age 50 to make individual indifferent between the baseline and reformed scenario.

	Total assets at age 50 (quartiles)			
health at age 50	1st	2nd	3rd	4th
(quartiles)	< £122,000	-£213,000	-£350,000	-£350,000
1st	0.44	0.08	0.07	0.03
2nd	0.06	0.03	0.03	0.02
3rd	0.01	<0.01	<0.01	<0.01
4th	0.01	<0.01	<0.01	<0.01

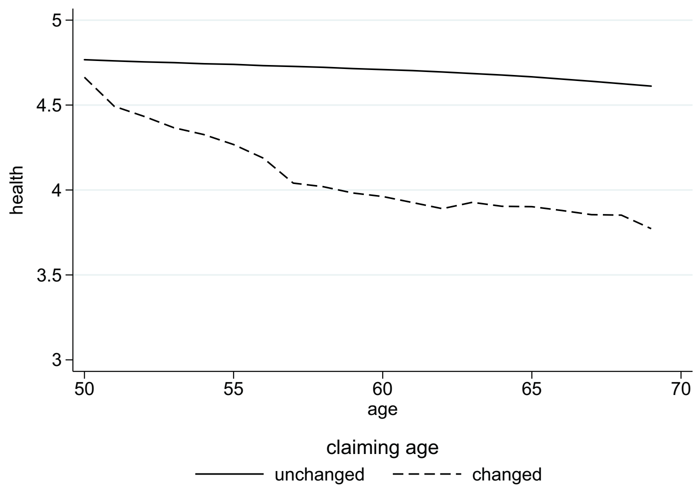
3.1.4 Interaction with other policies

As already pointed out in Section 1.5, IB is not the only insurance mechanism against health risk. In what follows I present some evidence on the insurance provided by savings in illiquid assets. I then focus on the role played by means-tested benefits, such as income support and pension credit, that top up income of low earners and have additional premiums for those with health limitations. The richness of the model allows to investigate whether the take-up of these benefits increases after IB removal and to some extent whether there exists a substitution effect among them.

Effect on private pension
The comparison of Figures 3.1 and 3.5 illustrates that the gains after age 60 cannot be explained only with an increase in labour income because labour supply remains almost unchanged after IB removal. One possible explanation is a change in claiming age for private pension, that is a change in savings in illiquid assets. The removal of

IB might have several implications. First, it might provide incentives to labour market attachment for those with a private pension and with a sufficiently high health level: on top of a resulting increase in current disposable income, additional contributions to private pension fund give rights to higher pension benefits in the future. Second, a increased health risk resulting from removing disability insurance might influence the timing of private pension claim: low wealth individuals might decide to anticipate benefit draw down when in poor health[3]; not constrained individuals instead might postpone benefit draw down and save in illiquid assets to insure themselves against the increased health risk.

Figure 3.7: Median health pattern for individuals claiming private pension at the same age or at a different age in baseline and reformed scenario.



The counterfactual experiment shows that about 17% of individuals having a private pension fund change claiming age. Among them 66% postpone withdrawal from private pension and the remaining anticipate pension draw down. The 17% that change their behaviour are those affected by benefit removal and they are characterized by a lower level of health (Figure 3.7). On average the pension amount at age 70 for agents postponing private pension claim is 21% higher with respect to the benefit they receive in

[3]This explanation is consistent with Figure F.2 in Appendix F showing that those in bad health (receiving IB for at least one year) claim earlier for private pension with respect to the others, suggesting that private pensions represent an insurance mechanism, in particular after age 55 when an agent can start to claim the benefit.

baseline. For those anticipating benefit draw down in the reformed scenario, the private pension benefit at age 70 is 12% lower.

All the guessed mechanisms are consistent with counterfactual results. Removing IB provides incentives to those at the margin (not in good health but with an health level that allows them to work) to work more and postpone private pension draw down to get an higher benefit. Moreover, for those in bad health private pension wealth represents an insurance against health deterioration risk. However, no clear differences in terms of health and wealth levels emerge between those anticipating and those postponing claiming age and further analyses are needed to understand which is the mechanism driving these changes in claiming age.

Income support and Pension credit

In the simulated baseline scenario the take-up of income support (IS) among individuals aged between 50 and 59 is about 1.3% and it increases by 12% after benefit removal. For individuals with initial wealth in the first quartile, the percentage receiving IS is about 5% and it increases by 12% in the reformed scenario.

Looking at Pension Credit and in particular at the Guarantee Credit (GC) component, the baseline model predicts a take-up of 5.6% and it increases by 1.3% after IB removal[4]. For individuals with initial wealth in the first quartile, the average take-up is 22% in baseline and it does not increase in the reformed scenario. If I consider the second quartile, GC take-up is zero in the baseline scenario and it increases to 1% when IB is removed.

In both cases there is some evidence of replacement between IB and IS and PC

[4]In ELSA data, among males living with a partner in the relevant age group, 2% report IS recipience and 4.3% (of those aged 60 to 64) report GC recipience. These statistics are close to simulation results: the model predicts 1.3% receiving IS and 5.6% receiving GC. To assess to what extent these numbers are in line with those obtained from administrative data I use the following data sources: the DWP tabulation tool to recover the number of claimants of IS (aged 50 to 59) and GC (aged 60 to 64) in 2004 having a partner; 2001 Census data to recover the fraction of couple-households aged over 50 (88.7%); the Mid-2004 Population Estimates (Sources: Office for National Statistics, General Register Office for Scotland, Northern Ireland Statistics and Research Agency) to retrieve the number of males aged 50 to 59 and 60 to 64. Using these information I compute the percentage of males living with a partner aged 50 to 59 receiving IS and the percentage of males living with a partner aged 60 to 64 receiving GC. The resulting statistics are 2.4% for IS and 13.8% for GC. While the percentage receiving IS in the data and in the population are quite close, the data seems to underestimate those receiving GC.

respectively. However, as analyses on the distribution of disposable income have revealed, this is not enough to compensate individuals from IB removal[5].

3.2 Varying the health eligibility rule

One dimension through which the government has tried to reduce benefit inflow and the disincentive provided by IB is by strengthening the health assessment. In my model, such assessment is summarised by the IB eligibility health threshold H_d. To understand the effect of these policies, I simulate the model under several health thresholds. Effects on labour supply are investigated by looking at changes in the number of years worked between age 50 and 69 (by assumption at age 70 individuals retire) and at changes in the fraction claiming IB, whereas to gauge distributional consequences of the different scenarios I consider the distribution of average disposable income between age 50 and 64.

Figure 3.8: Changes in years worked and in the fraction claiming IB resulting from changes in the health assessment.

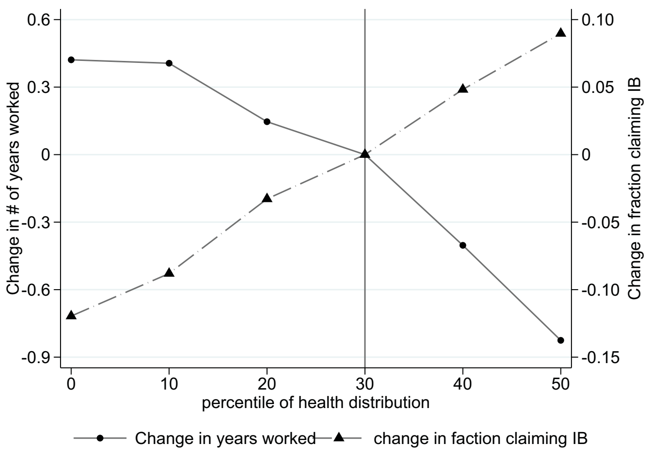


Figure 3.8 shows on the x-axis the health thresholds measured as percentiles of the health distribution, on the y-axes the absolute change in the number of years worked between age 50 and 69 relative to the baseline scenario (left axis) and the absolute change

[5]As the means used to determine benefits eligibility include also the income of the spouse, the deterministic function that approximates the spousal income might affect means tested benefit eligibility and amount.

in the fraction claiming IB between age 50 and 64 (right axis).

The vertical line represents the calibrated threshold in the baseline model, corresponding to the 30th percentile of the unconditional health distribution, which is the reference scenario. In the baseline the average number of years worked between 50 and 69 is 12.46. The extreme case of health threshold set to zero corresponds to the first experiment, i.e. elimination of IB. A reduction of the threshold to the 20th percentile causes an increase in average years worked of 0.15 (1.2%). The percentage claiming the benefit between age 50 and 64 moves from 12% to 8.7%. Assuming that the effect is entirely due to changes in the behaviour of individuals not receiving the benefit anymore (3.3%), the 0.15 years of work in the aggregate corresponds to an increase of 4.5 years for this group. The other extreme (not reported in the chart) corresponds to eliminating the health eligibility condition. This has a huge effect on participation, the absolute change is -2.7 years corresponding to a 22% drop in years worked.[6]

To evaluate the aggregate savings or costs associated with each benefit threshold I perform a simple back-of-the-envelope calculation starting from the expenditure for IB in 2003/2004 tax and benefit year[7]. Among males receiving IB in 2004, 51% were aged 50 to 64 and they are responsible for the 62% of the total expenditure on IB[8]. Under the simplifying assumption that the expenditure on IB is equally distributed between couples and singles, expenditure on IB for those aged 50 to 64 is assumed equal to 62% of total expenditure on IB, that is about 3,250 millions of £. When the threshold is reduced

[6]Banks *et al.* (2015b) estimate that the option of receiving IB when the health eligibility condition is removed reduces years worked by 5.5%, which is much lower than the 22% reduction that I find. They estimate a reduced form retirement model including among controls measures of the option value of remaining in paid work capturing the incentives provided by state and private pensions as well as state provided disability insurance. The difference is probably due to different working assumptions: in their modelling assumption when an individual takes the disability pathway to retirement he is assumed to receive the benefit up to state pension age, as if disability benefit was an absorbing state. In the fully dynamic specification that I propose, individuals can enter end exit from the benefit in every period. In addition, there are no direct mechanisms creating persistence in IB receipt other than persistence in health and wages. As pointed out in presenting the model fit, the model is able to replicate 80% of the persistence observed in the data. The missing 20% together with a more flexible benefit take-up suggest that the difference among the two results is justifiable and the effect I find on participation can be thought of as an upper bound.

[7]Source: DWP, Benefit Expenditure and Caseload Tables, March 2014 (`https://www.gov.uk/government/publications/benefit-expenditure-and-caseload-tables-2014`).

[8]Source: DWP tabulation tool.

from the 30th to the 20th percentile of the health distribution, the expenditure on IB for males aged 50 to 64 is reduced by about 894 millions of £ (1% of total expenditure).

Table 3.4: Distributional consequences of changes in the health eligibility threshold: average disposable income distribution between age 50 and 64 in baseline and reformed scenarios.

policy	p5	p10	p25	p50	p75	p90	p95	Gini
	average disp. income							
zero	10,807	13,086	17,075	21,165	24,996	28,295	30,328	0.157
10th quantile	11,468	13,466	17,178	21,177	25,005	33,940	30,321	0.153
20th quantile	11,344	13,355	17,153	21,136	25,009	28,266	30,280	0.154
baseline	11,217	13,271	17,153	21,174	25,031	28,295	30,299	0.154
40th quantile	11,095	13,036	16,990	21,045	24,960	28,253	30,226	0.156
50th quantile	11,008	12,851	16,669	20,865	24,821	28,136	30,140	0.159

Table 3.4 reports the distribution of the average disposable income between age 50 and 64 for different levels of the health threshold. It emerges that the lower the threshold (the stricter the eligibility conditions) the lower the dispersion in the distribution. In particular, the distribution shifts towards higher values as the threshold is reduced. When average income is considered, the effect that dominates is the reduction in the disincentive to work as a result of less generous benefit structures. The relationship is reversed in the extreme scenario in which health threshold is zero, that coincides with removing IB: when average disposable income is considered much of the effect emerged when the threshold is reduced from the 10th percentile to zero. A more generous health threshold increases average income dispersion and triggers higher disincentives to work as well as a shift of the income distribution towards lower values.

A different picture emerges looking at the income distribution of individuals between 60 and 64 years old (not reported in the table). Moving the threshold from the 30th to the 20th percentile of the health distribution has negligible effects on the distribution of income, whereas a reduction of the threshold to the 10th percentile causes a shift of the distribution towards the left.

I conclude that reducing the threshold might help target efficiency and induce those with remaining working capacity to return to work (in the aggregate average disposable income between 50 and 64 increases) but at the cost of increased inequality, in particular among the olders.

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

Starting from the pilot implementation of Pathways-To-Work in 2003 and even more from the introduction of the Employment and Support Allowance, policy makers in addition to reduce the inflow by tightening eligibility conditions have been committed to promote the outflow from IB. Under the new regime, when claiming for ESA, applicants get the assessment rate for the first 13 weeks. After that, if entitled, they are placed into the work-related activity group or the support group. Beneficiaries in the work-related activity group have to participate to regular interviews with an adviser aimed at improving work readiness by helping individuals with job search and with the management of health related problems within a work context. Those in the support group are severely limited in their activities and are not required to pursue any activity.

Inspired by this aspect of the reform, I consider a public intervention aimed at enhancing labour market inclusion of people with disabilities, such as legislation to the work environment, labour market measure, non-discrimination and accessibility in the work place. Assuming that the mechanism through which this policy affects individuals is by reducing their fixed cost of work, that is the costs associated with labour market entry or re-entry, I look at the effect on DI take-up and DI duration of a proportional reduction in the fixed cost of work ϕ_P. In the model the fixed cost of work is specified as a time cost and it is linearly increasing in age. I assume that if an individual has health below the eligibility threshold for receiving IB, his fixed cost of work is proportionally reduced by 30 and 50% in the reformed scenario.[9]

In Figure 3.9 I report changes in benefit take-up by age in baseline and when ϕ_P is reduced by 30 and 50% respectively. The fraction claiming for IB is reduced by 10% in the first reformed scenario and by 26% in the second. The effect on duration is shown in Figure 3.10 for baseline and reformed scenarios. I consider for each individual the longer IB spell between age 50 and 64, the average duration of the spells in baseline is

[9] This means that not only those receiving IB are affected, but all individuals at risk of claiming IB. Moreover the proportional reduction implies that the absolute reduction in the cost of work is increasing with age.

3.5[10]. A 30% reduction in ϕ_P implies a 6% reduction in IB duration, whereas average duration decreases by 11% as a result of a 50% reduction in the cost of work.

Figure 3.9: IB profiles in baseline and reformed scenarios.

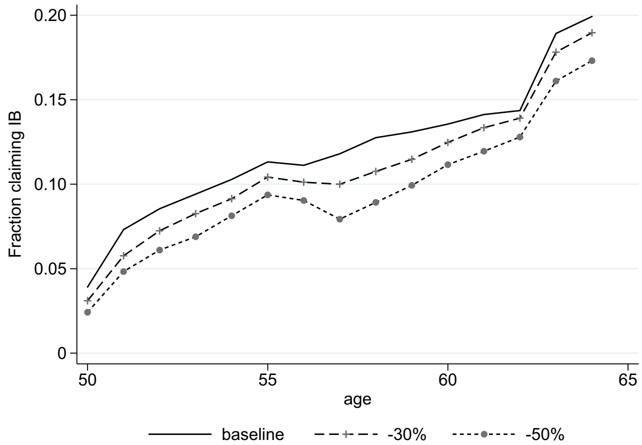
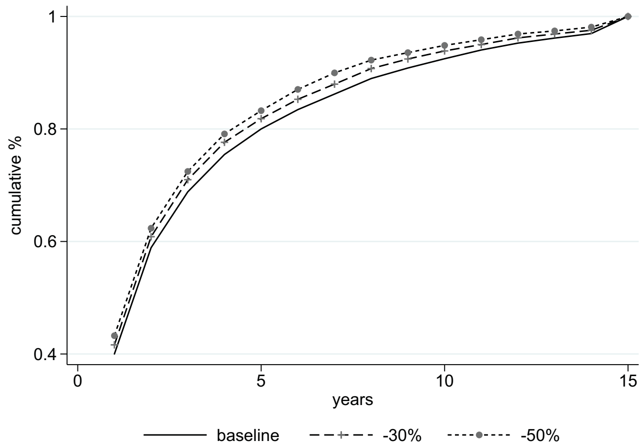


Figure 3.10: Cumulative percentage of IB duration in baseline and reformed scenarios.



With respect to a tightening of the eligibility conditions to receive the benefit, this policy intervention allows to reduce the inflow into IB and reduce benefit duration without increasing inequality. A reduction of the fixed cost of work for those at risk of

[10]The average duration is 4.6 if I consider as measure of duration the number of years receiving the benefit between age 50 and 64.

IB entry reduces the dispersion of the distribution of average disposable income between age 50 and 64 and, overall, it increases average disposable income.

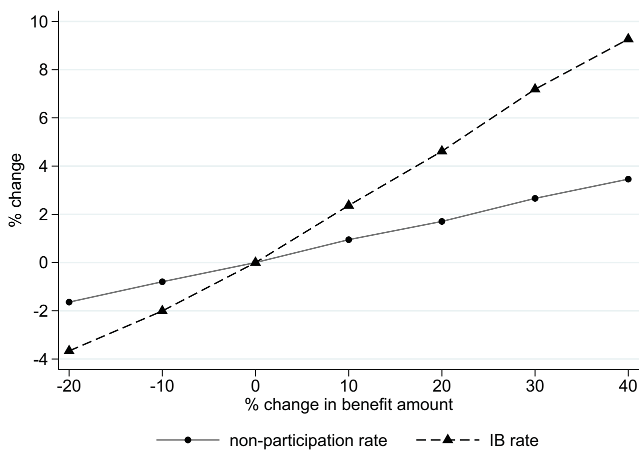
The revenue cost associated with each reformed scenario can be computed as explained in Section 3.2. When the fixed cost of work is reduced by 30% the expenditure on IB for males aged 50 to 64 is reduced by about 325 millions of £, the reduction is instead 678 millions of £ when the fixed cost of work is halved.

3.4 Varying the benefit amount

In the UK institutional context the IB amount is not earnings related even if eligibility requires having paid enough contributions. The amount is flat and quite low, representing a valuable replacement income only for low earners. In this policy experiment I simulate individuals' behaviour under different IB amounts to gauge responses in terms of labour supply decision and to look at distributional consequences of benefit generosity.

To compare model predictions with the existing literature I compute non-participation rate and IB rate by varying the benefit amount from a 20% reduction to a 40% increase relative to baseline. Results are reported in Figure 3.11. The implied elasticity of non-participation to the labour market is 0.09 and the implied elasticity of IB inflow rate is 0.24, for individuals aged 50 to 64. These statistics are low with respect to the range of estimates in the literature.

Figure 3.11: Relationship between IB amount and non-participation/IB rate between age 50 and 64.



Bound and Burkhauser (1999b) survey elasticities estimated for the US institutional context. In addition to the elasticity of labour market non-participation rate $(0.2 - 1)$, two other elasticities are computed: the elasticity of DI application $(0.2 - 1.3)$ and the elasticity of DI award $(0.3 - 0.4)$ with respect to DI benefit level. In the case of the UK it is not necessary to withdraw from the labour market to claim the benefit, moreover I do not model benefit application, therefore the elasticity of IB inflow rate has to be compared with the elasticity of DI award. Mullen and Staubli (2015), using Austrian data and exploiting reforms during 1990s and 2000s, find an elasticity of DI inflow rate to benefit generosity of 1.2, and an elasticity of 0.7 in the period in which the replacement rate was lower $(2004 - 2010)$.

Two reasons might explain the lower elasticities found for the UK. First, the elasticities I compute refer to males aged between 50 and 64, but if I consider only individuals aged 50 to 54 the elasticity of non-participation to the labour market is about 0.34 and the elasticity of IB inflow rate is 0.41, which are values in the range of earlier estimates. This is consistent with previous findings suggesting that elasticities are decreasing as age increases[11]. Second, the peculiar aspect of IB of providing a low flat rate benefit amount results in mainly low skilled workers with lower employment opportunities entering the benefit. Therefore, given the characteristics of the target population one might expect the demand for IB to be rather inelastic to marginal changes in the benefit amount. This is consistent with findings in Mullen and Staubli (2015) according to which the elasticity of DI inflow with respect to benefit generosity is lower for low skilled and poorer workers[12].

Heterogeneity emerges among health levels[13]. Those in worst health (below the 10th percentile) show an elasticity of IB participation of 0.20 and those with health close to the eligibility threshold (between the 20th and the 30th percentile) an elasticity of IB

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

[12]The replacement rate of IB in the UK is about 15% of average earnings after 1995 reform (Banks *et al.*, 2015b), whereas Mullen and Staubli (2015) report an average replacement rate of 57% of the average wage over the best 15 years in Austria and Autor and Duggan (2003) compute replacement rates from 23 to 74% in 1999 for the US.

[13]The same heterogeneity is found also by Low and Pistaferri (2015) for the US.

participation of 0.24. For what concerns labour force non-participation, those at the bottom and at the top of the health distribution show an elasticity close to zero. The more responsive are those close to the eligibility threshold with an elasticity of 0.19.

The heterogeneity of the elasticities by health is the result of the interaction of several endogenous effects that operate, sometimes with different signs, on individuals with different initial conditions. While it is hard to disentangle the contribution of each channel, interesting information can be learned by simulating the model under different assumptions, eliminating various sources of heterogeneity one at a time[14].

For example, the time cost of being in bad health (ϕ_H) plays an important role: the higher ϕ_H the lower the elasticity of IB inflow and the elasticity of non-participation with respect to benefit generosity for a given level of health. For a sufficiently high time cost of being in bad health, individuals non-participation elasticity when bad health is close to zero meaning that they do not have working capacity left.

Furthermore, there is an interesting substitution between means tested benefits (IS and GC) and IB for low health workers when IB benefits are increased, which contribute to a higher elasticity of IB inflow for low health workers than higher health workers.

Finally, individual heterogeneity also significantly contributes to generate heterogeneity in the elasticities. In particular, elasticities computed for individuals with homogenous health levels show large heterogeneity by different levels of initial assets: the lower the level of assets the higher the elasticities. Therefore, the reported elasticities of IB inflow and non-participation are the result of a composition effect of heterogeneous elasticities by wealth levels.

Table 3.5: Distributional consequences of changes in the health eligibility threshold: average disposable income distribution between age 50 and 64 in baseline and reformed scenarios.

policy	p5	p10	p25	p50	p75	p90	p95	Gini
	average disp. income							
-20%	11081	13125	17080	21168	24995	28272	30277	0.156
baseline	11217	13271	17153	21174	25031	28295	30299	0.154
+20%	11480	13398	17196	21175	25056	28289	30282	0.153
+40%	11690	13637	17255	21214	25077	28314	30273	0.151

[14] The details of these simulations are available upon request from the author.

Given the small elasticity of IB rate to benefit generosity, I expect the effect on disposable income to be small. This is confirmed by results reported in Table 3.5. The distribution of average disposable income between age 50 and 64 is less dispersed as benefit amount increases. This is the result of an increase of the lower quantiles, while the upper ones remain almost unchanged. The revenue cost associated with a 20% increase in the benefit amount is about 135 millions of £.

Conclusions

In this work, I develop and calibrate a life-cycle model of labour supply, DI claiming and savings behaviours for the UK. I model the decisions of males living with a partner and approaching retirement age facing uncertainty on wage realization, health and life expectancy. The model is able to replicate quite well average and median assets profiles, labour market participation and its heterogeneity by health level, as well as the fraction receiving Incapacity Benefit by age and benefit persistence over time. Health is measured on a continuous scale and it is based on a large set of objective indicators collected in ELSA, covering the health domains measured in the health assessment to receive IB. Both the mean and the distribution of health evolution over time is well reproduced by the specified process for health, that has a deterministic component depending on age and a stochastic component allowing both persistent and transitory shocks on the health stock.

The model is used to investigate labour supply responses and distributional consequences of alternative structures of the DI programme. The policy experiments reveal the presence of unused working capacity among IB recipients, in particular among those between 50 and 55 and at the margin of programme entry. The benefit has important distributional consequences and reduces income inequality caused by health effects on productivity. The insurance value provided by the benefit is largely heterogeneous across health and wealth groups: it is more valuable for individuals in bad health at age 50 and with wealth in the first quartile of the distribution at age 50. Strengthening the health assessment to receive IB might help target efficiency and increase labour market participation overall, but at the cost of an increase in the dispersion of the income distribution. A policy reducing the cost of re-entering the labour market by 30% (for those with health below the IB threshold) is promising, as it increases labour supply by 3% without increasing inequality. Finally, marginal changes in the generosity of the benefit amount seems not to have an overall large effect on labour supply. Focusing instead on individuals younger than 55 the effects are larger: the elasticity of non-participation to the labour market with respect to benefit generosity is 0.34 and the elasticity of IB rate is 0.41.

The policy results are valid for cohabiting males. I specify a unitary model in which the head of the household maximizes his utility, and the income of the partner (a deterministic function of the male's characteristics) enters the budget constraint. Speculating on results extension to single males, one might expect single males to be more responsive to IB changes: they could not rely on an additional source of income as it is the case for cohabiting males. However, entitlements to means tested benefits might attenuate the elasticity of the demand for IB and it is not a priori clear which mechanism would dominate. In addition, the distribution of initial conditions would be different, in particular single males are more likely to be in poor health and less wealthy than males in couples.

The results presented in this dissertation are conditional on the plausibility of the modelling assumptions. In particular the model assumes that health is perfectly observed, preventing an evaluation of changes in the effectiveness of the screening process. However, to include this aspect in the model one would need data on the application process, the examiners evaluation and the final decision. Information exploited in other studies (Low and Pistaferri, 2015; Ball and Low, 2014) to identify the probability of false applications and the probability of false rejections is the presence of self-reported health conditions limiting working capacity. ELSA respondents are asked about the presence of health problem that limits the kind or amount of work they can do, however cross tabulation of this variable with IB receipt reveals that only 10% of males aged over 50 report no health problem that limits work while receiving DI benefit and 20% report some problems even if not receiving the benefit. This feature of the programme needs further investigation, in particular for what concerns the reformed DI programme (ESA) that tightens the health eligibility rule and, with the reassessment of IB claimants, increases sharply the rejection rate.

A second limitation is that health is treated as an exogenous process. This assumption may be acceptable when considering older individuals, for whom health capital is in large part predetermined. Many studies have shown the effect of labour supply decision when old on several health outcomes (Rohwedder and Willis, 2010; Bonsang *et al.*, 2012; Celidoni *et al.*, 2016; Rohwedder and Willis, 2010) and the importance of individuals' behaviours in terms of prevention, nutrition and physical and cognitive activities to prevent health deterioration (Crowther *et al.*, 2002; Haveman-Nies *et al.*, 2003; McLaughlin *et al.*, 2010).

Finally, my model ignores the joint decision of couples living with a partner. Two concerns might arise with respect to this modelling choice. First, couples are generally richer than singles and in the UK the DI programme represents valuable replacement income in particular for low earners. However, large fraction of individuals aged 50 to 64 interviewed in ELSA live with a partner. A complementary analysis considering singles is left for future research. Second, the presence of the partner enters the model through the budget constraint only and the joint behaviour of the couple is not modeled. This might be relevant when investigating the effect of health on labour supply because the presence of the partner and her economic status might influence the couple decision to participate in the labour market and to claim for IB. On the one hand the allocation of time devoted to work within the couple might vary with the health level of the partners, on the other hand if one couple member is in bad health the other might decide to stay out of the labour market and give informal care to the sick partner. This latter decision is influenced by preferences and by the financial support provided to the caregiver by the welfare system.

References

Adam, S., Bozio, A. and Emmerson, C. (2010) Reforming disability insurance in the UK: Evaluation of the Pathways to Work Programme, `http://econ.sciences-po.fr/sites/default/files/file/AdamBozioEmmerson2010_0.pdf`.

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., Maestas, N., Mullen, K. J. and Strand, A. (2015) Does delay cause decay? the effect of administrative decision time on the labor force participation and earnings of disability applicants, Working Paper 20840, National Bureau of Economic Research.

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

Banks, J., Blundell, R., Bozio, A. and Emmerson, C. (2012) Disability, health and retirement in the United Kingdom, in *Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms* (Ed.) D. A. Wise, University of Chicago Press, pp. 41–77.

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

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

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

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, `http://ms.cc.sunysb.edu/~hbenitezsilv/jpam_2011new.pdf`.

Black, D., Daniel, K. and Sanders, S. (2002) The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust, *American Economic Review*, **92**, 27–50.

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. (2016) Comparing retirement wealth trajectories on both sides of the Pond, *Fiscal Studies*, **37**, 105–130.

Bonsang, E., Adam, S. and Perelman, S. (2012) Does retirement affect cognitive functioning?, *Journal of Health Economics*, **31**, 490–501.

Bound, J. (1989) The health and earnings of disability insurance applicants, *American Economic Review*, **79**, 482–503.

Bound, J. (1991) The health and earnings of rejected disability insurance applicants: Reply, *American Economic Review*, **81**, 1427–1434.

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

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

Bound, J., Cullen, J. B., Nichols, A. and Schmidt, L. (2004) The welfare implications of increasing disability insurance benefit generosity, *Journal of Public Economics*, **88**, 2487–2514.

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

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

Bound, J. and Waidmann, T. (1992) Disability transfers, self-reported health, and the labor force attachment of older men: Evidence from the historical record, *The Quarterly Journal of Economics*, **107**, 1393–1419.

Bradley, C. J., Neumark, D. and Barkowski, S. (2013) Does employer provided health insurance constraint labour supply adjustments to health shocks? New evidence on women diagnosed with breast cancer, *Journal of Health Economics*, **32**, 833–849.

Burkhauser, R. V., Butler, J. and Gumus, G. (2004) Dynamic programming model estimates of social security disability insurance application timing, *Journal of Applied Econometrics*, **19**, 671–685.

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

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

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

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

Celidoni, M., Dal Bianco, C. and Weber, G. (2016) Retirement and cognitive decline. a longitudinal analysis using share data, mimeo.

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

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

Crawford, R. and O'Dea, C. (2016) Household portfolios and financial preparedness for retirement, `https://dl.dropboxusercontent.com/u/9281649/ForWebsite/CrawfordODea_Submitted.pdf`.

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

Crowther, M. R., Parker, M. W., Achenbaum, W. A., Larimore, W. L. and Koenig, H. G. (2002) Rowe and kahn's model of successful aging re-visited: positive spirituality - the forgotten factor, *The Gerontologist*, **42**, 613–620.

Currie, J. and Madrian, B. C. (1999) Health, health insurance and the labor market, *Handbook of Labor Economics*, **3**, 3309–3416.

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. (2013) Medicaid insurance in old age, Working Paper 19151, National Bureau of Economic Research, forthcoming in The American Economic Review.

Disney, R., Emmerson, C. and Wakefield, M. (2003) Ill health and retirement in Britain: A panel data-based analysis, Working Paper WP03/02, The Institute for Fiscal Studies.

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

Duggan, M. G. and Imberman, S. A. (2009) Why are the disability rolls skyrocketing? the contribution of population characteristics, economic conditions, and program generosity, in *Health at Older Ages: The Causes and Consequences of Declining Disability among the Elderly* (Eds.) D. M. Cutler and D. A. Wise, University of Chicago Press, pp. 337–379.

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.

García-Gómez, P. (2011) nstitutions, health shocks and labour market outcomes across Europe, *Journal of Health Economics*, **30**, 200–213.

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

Grossman, M. (1972) On the concept of health capital and the demand for health, *Journal of Political Economy*, **80**, 223–255.

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

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.

Gruber, J. and Wise, D. A. (Eds.) (1999) *Social Security Programs and Retirement Around the World*, University of Chicago Press.

Gruber, J. and Wise, D. A. (Eds.) (2004) *Social Security Programs and Retirement Around the World: Micro Estimation*, University of Chicago Press.

Gruber, J. and Wise, D. A. (Eds.) (2007) *Social Security Programs and Retirement Around the World: Fiscal Implications of Reform*, University of Chicago Press.

Gustman, A. L. and Steinmeier, T. L. (1986) A structural retirement model, *Econometrica*, **54**, 555–584.

Hai, R. and Heckman, J. J. (2015) A dynamic model of health, education, and wealth with credit constraints and rational addiction, Working Paper 2688384, Becker Friedman Institute for Research in Economics, available at SSRN: `http://ssrn.com/abstract=2688384` or `http://dx.doi.org/10.2139/ssrn.2688384`.

Halpern, J. and Hausman, J. A. (1986) Choice under uncertainty: a model of applications for the Social Security Disability Insurance program, *Journal of Public Economics*, **31**, 131–161.

Haveman, R. H. and Wolfe, B. L. (1984) The decline of male labor force participation: Comment, *Journal of Political Economy*, **92**, 532–541.

Haveman, R. H. and Wolfe, B. L. (2000) The economics of disability and disability policy, *Handbook of Health Economics*, **1**, 995–1051.

Haveman-Nies, A., de Groot, L. C. and van Staveren, W. A. (2003) Dietary quality, lifestyle factors and healthy ageing in europe: the seneca study, *Age and Ageing*, **32**, 427–434.

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

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

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

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

Kapur, K. (1998) The impact of health on job mobility: A measure of job lock, *ILR review*, **51**, 282–298.

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

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

Kreider, B. (1999) Social Security Disability Insurance: Applications, awards, and lifetime income flows, *Journal of Labor Economics*, **17**, 1223–1235.

Kreider, B. and Riphahn, R. T. (2000) Explaining applications to the U.S. disability system: A semiparametric approach, *Journal of Human Resources*, **35**, 82–115.

Lahiri, K., Song, J. and Wixon, B. (2008) A model of social security disability insurance using matched SIPP/administrative data, *Journal of Econometrics*, **145**, 4–20.

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

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

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

Lumsdaine, R. L. and Mitchell, O. S. (1999) New developments in the economic analysis of retirement., *Handbook of Labor Economics*, **3**, 3261–3307.

Madrian, B. C. (1994) Employment-based health insurance and job mobility: Is there evidence of job-lock?, *The Quarterly Journal of Economics*, **109**, 27–54.

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.

McLaughlin, S. J., Connell, C. M., Heeringa, S. G., Li, L. W. and Roberts, J. S. (2010) Successful aging in the united states: Prevalence estimates from a national sample of older adults, *The Journal of Gerontology*, **65B**, 216–226.

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

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

Morciano, M., Hancock, R. M. and Pudney, S. E. (2015) Birth-cohort trends in older-age functional disability and their relationship with socio-economic status: Evidence from a pooling of repeated cross-sectional population-based studies for the uk, *Social Science & Medicine*, **136-137**, 1–9.

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

Nakajima, M. and Telyukova, I. A. (2012) Home equity in retirement, available at SSRN: `http://ssrn.com/abstract=1898330` or `http://dx.doi.org/10.2139/ssrn.1898330`.

Nakajima, M. and Telyukova, I. A. (2016) Housing and saving in retirement across countries, in *Contemporary Issues in Microeconomics* (Eds.) J. E. Stiglitz and M. Guzman, Palgrave Macmillan UK, London, pp. 88–126.

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

Parsons, D. O. (1980a) The decline of male labor force participation, *Journal of Political Economy*, **88**, 117–134.

Parsons, D. O. (1980b) Racial trends in male labor force participation, *American Economic Review*, **70**, 911–920.

Parsons, D. O. (1991) The health and earnings of rejected disability insurance applicants: Comment, *American Economic Review*, **81**, 1419–1426.

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.

Rohwedder, S. and Willis, R. J. (2010) Mental retirement, *Journal of Economic Perspectives*, **24**, 119–138.

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

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

Scholz, J. K. and Seshadri, A. (2013) Health and wealth in a lifecycle model, `http://www.ssc.wisc.edu/~aseshadr/WorkingPapers/Health_and_Wealth.pdf`.

Slade, F. B. (1984) Older men: disability insurance and the incentive to work, *Industrial Relations*, **23**, 260–269.

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

Stock, J. and Wise, D. (1990) Pensions, the option value of work, and retirement, *Econometrica*, **58**, 1151–1180.

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

Trevisan, E. and Zantomio, F. (2016) The impact of acute health shocks on the labour supply of older workers: Evidence from sixteen European countries, *Labour Economics*.

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

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

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

Wise, D. A. (Ed.) (2012) *Social Security Programs and Retirement around the World: Historical Trends in Mortality and Health, Employment, and Disability Insurance Participation and Reforms*, University of Chicago Press.

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

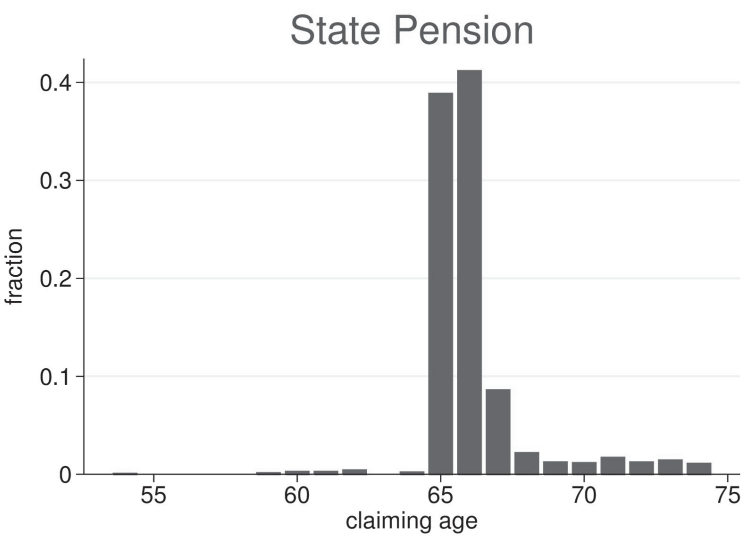
Wise, D. A. (Ed.) (forthcoming) *Social Security Programs and Retirement Around the World: The Capacity to Work at Older Ages*, University of Chicago Press.

Appendix A

Pensions' claiming age

The state pension age for males in UK is 65 and even if incentives to delay state pension claim have been increased since 2004, the majority of individuals claims state pension at age 65. In Figure A.1 I report the age at which individuals are first observed receiving state pension using ELSA data. Having biannual data does not allow to distinguish whether the benefit is first received in the interview year or the year before.

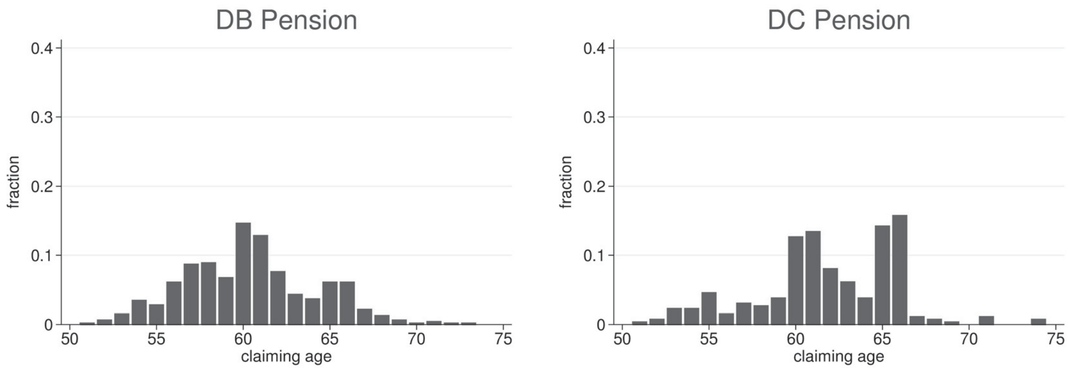
Figure A.1: Age at which individuals are first observed receiving state pension benefit. ELSA wave 1 to 6.



For what concerns private pensions, as in the US a progressive shift from defined benefit (DB) plans to defined contribution (DC) plans has been observed in the UK.

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

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



(a) **(b)**

In Figure A.2, using ELSA data, I report the age at which I first observe individuals receiving a DB or a DC pension respectively. The distribution of age for DC pension recipients is shifted to the right with respect to the one for DB recipients as expected. In the model I consider only Defined Contribution Pension, and in particular I convert accrual in DB plan as if it was accrued in a DC fund (i.e. it gives rights to an equivalent benefit). I do not model the decision between DC and DB plan, individuals are assumed to be endowed with a DC plan (or no private plan) when entering the model at age 50.

Appendix B

Health and wage process

B.1 Moments derivation

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

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

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

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

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

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

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

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

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

B.2 Selection in the wage profile

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

I solve this problem in two steps. First, I impute potential wages for non-workers estimating a regression model with selection by using full maximum likelihood (Heckman, 1979). Second, I estimate the wage process in Equation 2.24 using both observed and potential (imputed) wages.

In the first step I regress observed wages on a large set of individual characteristics: a polynomial in age, family size, educational level, time fixed-effects, a set of health controls[1], whether the individual is a smoker, whether he has a private pension plan, wealth quintiles and homeownership.

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

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

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

family structure I control for the presence of a partner, the presence of children, partner's health, whether the partner is above state pension age and whether the partner is above minimum age to withdraw from private pension plans.

The imputation model is used to obtain potential wages for non-workers. I then use offered wages (accepted and potential) to estimate the parameters of the wage process specified in Equation 2.24[2]. The estimated effect of health on wages when accounting for selection (i.e. using as dependent variable accepted and potential wages) is lower than the estimated effect when using as dependent variable only accepted wages. This is counterintuitive because one expects the health effect to be downward biased when selection is not accounted for. Further investigation is needed on this point. I solve the model, look at model fit and perform policy experiments using both the wage process estimated accounting for selection and the wage process estimated using only observed wages and results are very similar.

[2]In the literature two other approaches have been used. Low *et al.* (2010) and Low and Pistaferri (2015) correct for selection by applying a two-step Heckman correction to the wage process and deriving moment conditions to estimate the parameters of the error component model accounting for the correlation between selection equation and wage equation. This is done under the assumption of a non-stationary persistent component. French (2005) applies an iterative procedure that consists in replicating the selection observed in the date within the model in such a way that fixed-effects profile estimated using observed data is replicated by the fixed-effects profile obtained using simulated data.

Appendix C

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[1] is implemented instead. IB is a benefit targeted to sick individuals with temporary or long-term inability to work. The benefit can be received up to state pension age. It is taxable and the amount of the benefit depends on the weeks of sickness: a lower short term rate up to week 28, a higher short-term rate for weeks 29-52 of sickness and finally a higher long-term rate until state pension age. In the model a flat rate is applied and it is assumed that the benefit is received for twelve months, given that decisions are taken annually. Retirement Pension can be received starting from state pension age (65 for males). If contribution conditions are met the pensioner receives a flat rate basic pension. In addition, if pensioners have contributed to the State Earnings Related Pension Scheme (SERPS) an earnings-related pension is also payable. Both components are taxable.

In the second category - non-contributory and non-means-tested benefits - those relevant for this analysis are Attendance Allowance (AA) and Disability Leaving Allow-

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

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

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

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

Since September 2003 a means-tested income support scheme very similar to the one presented above was available to people aged 60 and older (Minimum Income Guarantee), but starting from October 2003 it has been replaced with Pension Credit. The introduction of the programme aimed at increasing the take-up of income support among the pensioners. In the tax function implemented in the model I consider the post-reform scenario. Thus I present below the main characteristics of PC. The PC consists of two elements: the Guarantee Credit (GC) meant to top up income to an 'appropriate minimum guarantee' and the Savings Credit (SC) meant to reward those who save for retirement. To be eligible to GC, individuals must be aged 60 or older and there are no capital limits. The tested income is the same as for IS with the exception that the tariff income is of £1 every £500 instead of every £250 and it is computed for capital above £6,000. The applicable needs are computed according to the basic allowance and the premium reported in Table C.1 (as for IS housing costs are not considered). Eligibility to SC requires being 65 or older and having means above the savings threshold, a reduced 40% taper rate applied to means above the threshold. The maximum amount receivable is reported in Table C.1. The income taken into account is the same as for GC except WTC that are deducted.

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

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

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

Table C.2: 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 C.3: Income tax schedule

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

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

For those aged less than SPA National Insurance payments are levied on earnings between a lower limit (£4,628) and the upper earnings limit (UEL £30,940) at a rate of 11%. Those having gross earning below the lower limit do not pay social insurance contributions, whereas those with earnings above UEL are subject to a rate of 1%, see Table C.5. These rules apply to those who are contracted in.

Finally, in Table C.6, I report the thresholds to compute Second Tier State pension accrual for 2003/2004 tax and benefit year.

Table C.4: Personal tax allowances and credits

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

Table C.5: National Insurance payments

Band	Rate on gross earnings
0-4628	0
4629-30940	0.1
30941-	0.01

Table C.6: Threshold to compute STSP accrual (tax year 2003/2004).

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

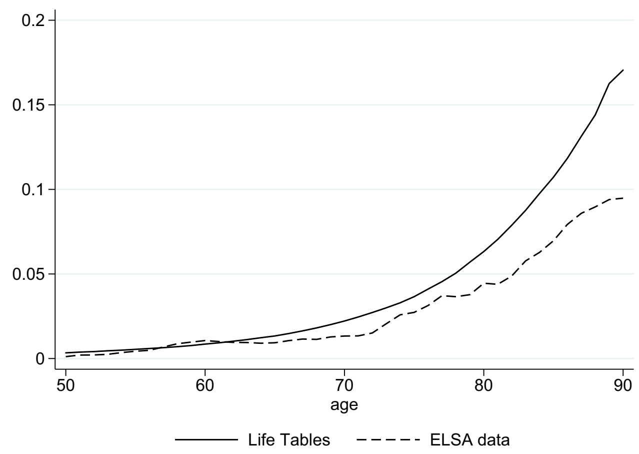
Appendix D

Robustness analyses

D.1 Mortality risk

In Section 2.5.2 the assumptions on mortality risks and their derivation are discussed. I assume that mortality rates' expectations are consistent with those of the life tables. As a result the implemented mortality process is obtained by re-scaling data probability in order to match life tables.

Figure D.1: Comparison between data and life tables mortality rates.



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

To assess the effect of this specific modelling assumption on model fit and model implications, I keep the set of calibrated parameters fixed and I solve the model using data mortality risk for males (the implemented death probabilities are those shown in Figure D.1 with a dashed line). As expected, with the alternative definition of mortality risk the survival probabilities increase, in particular for those in worse health, and this results in a modest decrease of average health when older. Mean and median assets are slightly lower, and participation slightly increases for those with health below the median. This means that the higher life expectancy for those in relatively worse health induces them to work more and save more in illiquid assets (pension wealth). Finally, hours worked and Incapacity Benefit rates do not show significant changes.

Table D.1: Effect of removing IB on employment. Baseline model specification with corrected mortality risk (1) and model with alternative mortality specification using raw data (2).

If receiving IB in baseline scenario		
	(1)	(2)
overall employment rate	0.224	0.232
employment rate if aged 50-54	0.470	0.464
employment rate if aged 55-59	0.249	0.270
employment rate if aged 60-64	0.080	0.087
Entire sample		
	(1)	(2)
overall increase in employment rate	0.040	0.044
fraction due to those receiving IB in baseline	0.690	0.704
fraction inactive that become active	0.210	0.205
fraction active that become inactive	0.097	0.091

Focusing on model implications, I perform the first policy experiment presented in Chapter 3 with both model specifications, in particular I compare changes in employment rate implied by the two models when IB is eliminated from the benefit system. Results are reported in Table D.1. The first panel reports statistics for those observed receiving IB in the baseline scenario, for this group employment rate is 22.4% when the corrected mortality rates conditional on health level are used and 23.3% when the aggregated data mortality rates are used. The latter employment rate is almost 4% higher than the former. When the entire sample is considered, employment rate increases by

4% percent in the first model specification and by 4.4% in the alternative model specification. I conclude from this comparison that results are only marginally affected by the assumption on mortality risk.

D.2 Revenue neutrality

The policy experiments presented in Chapter 3 are not revenue neutral. Given that I am considering the effect of the reform on a subsample of the population it is not obvious how to implement revenue neutrality without confounding the effects of interest. In this robustness analysis I redistribute the reduction in total expenditure generated by benefit removal as a 'universal' non taxable lump sum payment to each individuals up to SPA (above which IB cannot be received).

Table D.2 compares policy experiment results on labour supply responses in the absence of IB when no redistribution is implemented (Column 1 - the results are the same reported in Chapter 3) and when a broad revenue neutrality is implemented (Column 2).

Table D.2: Effect of removing IB on employment. Policy experiment reported in the main text (1) and 'revenue neutral' policy experiment (2).

If receiving IB in baseline scenario		
	(1)	(2)
overall employment rate	0.224	0.217
employment rate if aged 50-54	0.470	0.462
employment rate if aged 55-59	0.249	0.238
employment rate if aged 60-64	0.080	0.076
Entire sample		
	(1)	(2)
overall increase in employment rate	0.040	0.033
fraction due to those receiving IB in baseline	0.690	0.556
fraction inactive that become active	0.210	0.211
fraction active that become inactive	0.097	0.233

The first panel reports statistics for those observed receiving IB in the baseline scenario, for this group employment rate is 22.4% in the first column and 21.7% in the second column. When the entire sample is considered, employment rate increases by 4% percent in the first experiment and by 3.3% in the revenue neutral experiment. As expected the lump sum payment attenuates the effects of IB removal, in particular the fraction of those changing behaviour is lower and among them a larger percentage became inactive in the reformed scenario (from 10% in the first column to 23% in the second) meaning that the 'universal' non taxable transfer introduces a work disincentive among those approaching retirement age. However, it can be conclude that the main findings are confirmed.

Appendix E

Accrual derivation

The two variables to recover from pension wealth derived variables are accrual in state pension and accrual in private pension. As stated in the model formulation, I assume that everyone is entitled to Basic State Pension full amount, whereas earnings related state pension (Second Tier State Pension) is derived from the value of STSP provided in the released data, assuming the individual retires in the interview year.

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

Appendix F

Additional Tables and Figures

Figure F.1: Simulated mean consumption profile.

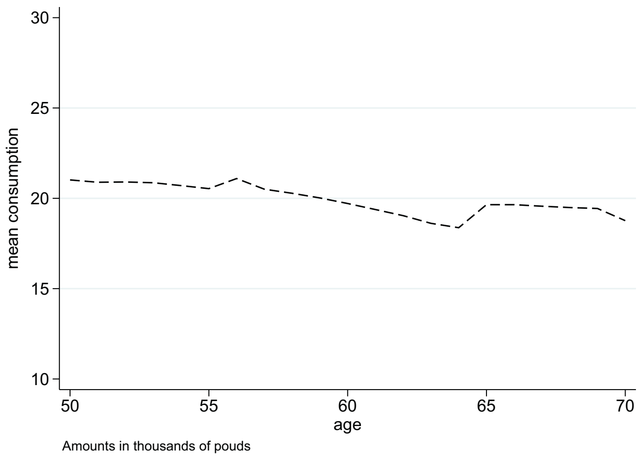


Figure F.2: Model simulation: cumulative frequencies of private pension claiming age conditional on having received IB for at least one year.

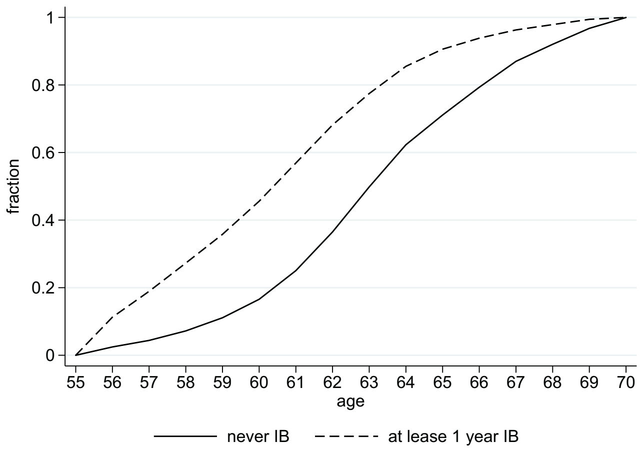


Figure F.3: First and second tertile of assets distribution by age. Simulations versus data.

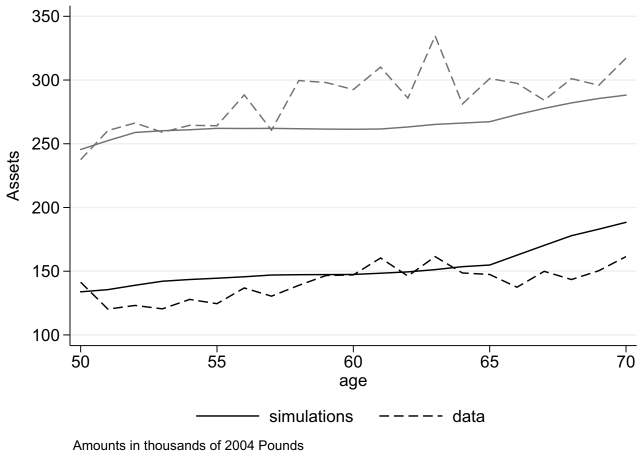


Figure F.4: First and second tertile of assets distribution, for health below (a) or above (b) the median. Simulations versus data.

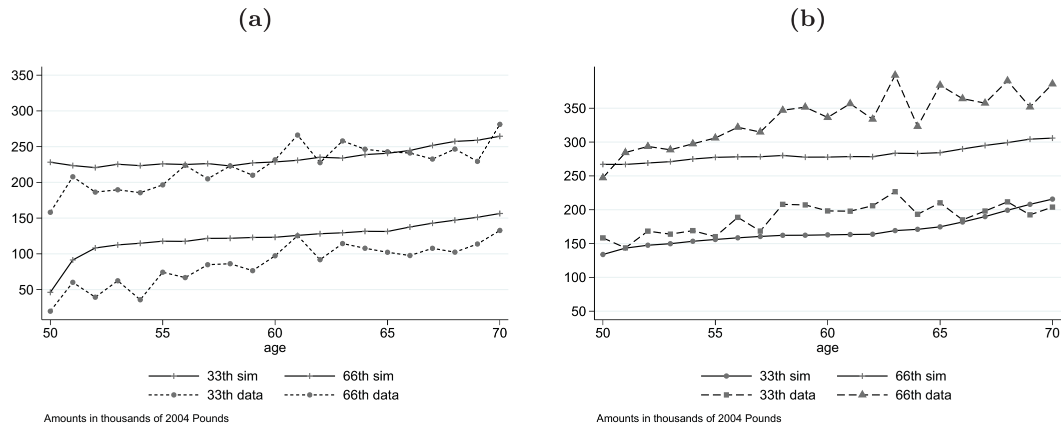


Figure F.5: Cumulative fraction claiming private pension.