

WORKING PAPER

Pension reform disabled

**NORGES BANK
RESEARCH**

20 | 2017

SIGURD MØLSTER
GALAASEN



NORGES BANK

Working papers fra Norges Bank, fra 1992/1 til 2009/2 kan bestilles over e-post:
FacilityServices@norges-bank.no

Fra 1999 og senere er publikasjonene tilgjengelige på www.norges-bank.no

Working papers inneholder forskningsarbeider og utredninger som vanligvis ikke har fått sin endelige form. Hensikten er blant annet at forfatteren kan motta kommentarer fra kolleger og andre interesserte. Synspunkter og konklusjoner i arbeidene står for forfatterens regning.

Working papers from Norges Bank, from 1992/1 to 2009/2 can be ordered by e-mail
FacilityServices@norges-bank.no

Working papers from 1999 onwards are available on www.norges-bank.no

Norges Bank's working papers present research projects and reports (not usually in their final form) and are intended inter alia to enable the author to benefit from the comments of colleagues and other interested parties. Views and conclusions expressed in working papers are the responsibility of the authors alone.

ISSN 1502-819-0 (online)

ISBN 978-82-8379-004-7 (online)

Pension reform disabled*

Sigurd Mølster Galaasen[†]

Abstract

Old-age pension reform is on the agenda across the OECD, and a key target is to delay retirement. Most of these countries also have a disability insurance (DI) program accounting for a large share of labor force exits. This paper builds a quantitative life-cycle model with endogenous retirement to study how DI and old-age pension (OA-pension) systems interact with health and wages to determine retirement age, with particular focus on the macroeconomic effects of OA-pension reforms. Individuals face uncertain future health status and wages, and if in bad health they are eligible for DI if they choose to retire before reaching the statutory retirement age. I calibrate the model to the Norwegian economy and explore the effects of raising the statutory retirement age and cutting OA-pension on labor supply and public finances. The main contribution of the paper is that I, in contrast to standard macro pension models, include DI as another endogenous margin of retirement. I show that failure to account for this margin might severely bias the analysis of OA-pension reforms. (JEL E2, E6, H31, H55, J26)

Keywords: Retirement, disability insurance, life-cycle, pension reform.

*This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the authors and do not necessarily reflect those of Norges Bank. This paper was previously circulated under the title "Disability Insurance, Pension Reforms and Retirement Behavior". The project received financial support from Statistics Norway and the Norwegian Research Council. I thank Mariacristina De Nardi, Eric French, Jonathan Heathcote, Steinar Holden, Asbjørn Rødseth, Robert Shimer, Kjetil Storesletten and conference and seminar participants at the Frisch Centre, Statistics Norway, University of Oslo, Norges Bank, 5th Nordic Summer Symposium in Macroeconomics, EEA ESEM 2014, IIPF 2014 and 13th International Workshop on Pensions, Insurance and Savings for useful comments.

[†]Norges Bank. E-mail: Sigurd-molster.galaasen@norges-bank.no

1 Introduction

Pension reform is on the agenda in most developed countries. The predicted increases in life-expectancies are putting public finances under pressure. A key target in most reforms is to increase labor supply. Despite this fact, little is known about macroeconomic labor supply elasticities of pension rules, and even less about the interaction with disability insurance (hereafter DI). In this paper I bridge this gap.

DI take-up across the OECD is high and, in many places, rising. On average, 6 percent of the working-age population rely on DI, with rates exceeding 10 percent in Norway and Sweden, and countries such as the UK and the US have seen enrollment rates increase by more than 50 percent since 1990 (OECD, 2010).¹ As the OECD data shows, DI recipients are likely to be old (median take-up rate across countries for age group 50-64 is roughly 15 percent) and never return to employment (the probability of returning to work is close to zero). Consequently, DI is an important path into retirement. Nevertheless, most pension reform studies neglect this dimension entirely, and the few that don't, typically assume exogenous DI enrollment.² This paper, on the other hand, studies pension reforms using a standard heterogeneous agent life-cycle model, augmented with health risk, public disability insurance, and endogenous retirement. In particular, agents in poor health have the option to retire early with DI. By including this dimension, I allow for an important alternative retirement margin. Failure to account for this dimension might bias the labor supply response to pension reforms.

Consider a pension reform that reduces old-age pensions (hereafter OA-pension). Agents typically respond by increasing their savings and postponing retirement. However, policy makers face a trade off when deciding how to treat those that rely on DI when transferred to OA-pension. Unhealthy individuals cannot easily delay retirement, and there are potentially good reasons to protect them from benefit cuts. On the other hand, if those on DI are protected, the incentive to retire through

¹For the US: 3.8 to 6.1 percent. For the UK: 4.2 to 7 percent

²For models with exogenous DI, see Díaz-Giménez and Díaz-Saavedrac (2009) and Erosa et.al (2011)

the DI system goes up, inducing an increased flow of workers to DI, offsetting the initial labor supply increase. Policy makers thus face the trade-off between maintaining insurance and limiting retirement incentives. The findings in Börsch-Supan et.al. (2005) provide support for the importance of such a trade-off. Using the 'Survey of Health, Ageing and Retirement in Europe' (SHARE), they find that the large variation in disability take-up across countries is not due to demographics and health, and conclude that institutional differences must be a key explanation.

In this paper I focus on the Norwegian economy, which has one of the highest DI take-up rates in the OECD. More than 40 percent of the population relies on DI when they are transferred to OA-pension at age 67.³ Over the last two decades the Norwegian pension system has also undergone two large reforms (in 1989 and 2011), both of which has important interactions with the DI system. In 1989 an early-retirement program (hereafter ERP) was launched, covering about 60 percent of the work-force. During the 1990s, the program gradually lowered the old-age pension eligibility age from 67 to 62, making OA-pension an alternative to DI.⁴ In 2011 the OA-pension became longevity-adjusted (increasing life-expectancy leads to benefit cuts), but disabled individuals are currently partly protected.⁵ Consequently, the institutional design of the Norwegian welfare state is an interesting starting point for studying the interplay between DI, retirement behavior and pension reform.

The core of the model is a dynamic theory of life-cycle consumption, saving and labor supply. Individuals decide how much to consume, save and when to stop working, facing idiosyncratic health, earnings and mortality risk. I use this framework to study how welfare programs interact with individuals' earnings and health shocks to determine career lengths, with particular focus on the effects of pension reform. This setup is closely related to French (2005). However, I make two important extensions. First of all, I model a DI system, thus allowing for interac-

³See White Paper No. 5, Ministry of Labour, 2006, page 31.

⁴The program, named AFP, covers a fraction of private sector workers, and everyone employed in the public sector. Today, roughly 60 percent of the work force is entitled to early retirement (Holmøy and Stensnes, 2008)

⁵See Government Proposition 130 L, 2011. In 2018, the government will re-evaluate the protection of the disabled.

tion effects between different welfare programs. Furthermore, I add two forms of preference heterogeneity needed to account for the data (as will be discussed below).

The main mechanism in the paper is as follows: Poor health increases the disutility of work. Unhealthy agents are eligible for DI if they retire, and retirement is permanent (cannot return to work). In the benchmark model, calibrated to the Norwegian economy, there are unhealthy individuals choosing not to retire early with DI. When an OA-pension reform increases the incentive to claim DI, some of these individuals become DI recipients. The main reform experiment (mimicking the 2011 Norwegian reform) removes an OA-pension earnings-test and reduces benefits by 20 percent from age 67 (due to increased longevity) but those coming from DI are protected from cuts. Despite large improvements in work incentives, aggregate labor supply in age group 51-69 falls with 2.1 percent relative to a no-reform scenario, due to a substantial increase in DI take up. If those on DI is not protected, however, the reform produces an old-age labor supply increase of 7.7 percent. To illustrate the importance of the main mechanism, I also consider a naïve accounting exercise, implicitly assumed in all other studies. That is, a model in which by assumption no substitution goes on; all DI retirement is exogenous (poor health force agents to retire). In this exercise the reform induces an increase in old-age labor supply of 4.1 percent. Consequently, not accounting for an endogenous DI margin might substantially bias the analysis of pension reforms.

A second contribution of the paper is the estimation of two types of preference heterogeneity: across agents (disutility of working) and across welfare programs (DI take up generates a utility cost). Consider first heterogeneous disutility of working. A key parameter in the model is the wage-offer profile among older agents. Due to selection one cannot estimate the wage profile directly on observed wages. However, since participation is endogenous in the model, one can consistently estimate the wage offers, providing a structural adjustment of the selection bias in the data (as in French, 2005). Even though selection is potentially acute, one particular dimension of the data indicates that it is not. Comparing labor income at age 51 for all individuals with income at age 51 for

the highly selected group who still works at age 67, I find that the latter group has only 7 percent higher labor income. Preference heterogeneity provides a straightforward way to account for this premium. In a version of the model calibrated with homogeneous preferences the simulated difference is 21 percent. I therefore augment the model with heterogeneous disutility of work, and identify the degree of heterogeneity by targeting the 7 percent premium in the data.

In addition, the model introduces a novel mechanism with respect to uptake of DI versus early-retirement benefit. DI retirement comes with a utility cost labeled *stigma*, a reduced form for all costs related to DI, including pure social stigma, cost of application, cost of going to the doctor etc. In contrast, I assume that this cost is not attached to uptake of early-retirement benefit, since age is the only eligibility criterion.⁶ Consider now the labor supply effects of a reform that introduces an ERP. In addition to lowering the eligibility age for OA-pension, the reform also offers a way for unhealthy agents to bypass the stigma cost attached to DI. This can potentially induce an unhealthy agent, who (despite being eligible) did not take up DI prior to the reform, to retire with early-retirement benefits. Consequently, the labor supply elasticity of the reform increases. This mechanism is needed to account for the large drop in employment following the introduction of the ERP in Norway. To quantify the magnitude of this cost, I exploit the empirical old-age employment rates, before and after the early-retirement reform in the 1990s.

Because key preference parameters are estimated by targeting aggregate retirement moments I can use retirement patterns in sub-groups as a check of the model. In particular, the model does a good job in replicating the very heterogeneous retirement patterns across education groups we observe in the data, both in the cross section and over time (before and after the introduction of the ERP). Moreover, among workers that are part of the ERP, the model replicates the fraction that receives early-retirement pension before age 67.

The paper proceeds as follows: The next section goes through related literature. Section 3 outlines the life-cycle model, while section 4 contains

⁶In order to be a member of the ERP, the worker has to be employed in a firm that participate in the program. As noted, roughly 60 percent of workers meet this criterion.

the calibration. Section 5 goes through the calibrated model and policy experiments. Section 6 concludes.

2 Related literature

The paper contributes to the macroeconomic literature on social security reforms, and is the first attempt to focus on the interaction between the DI and OA-pension systems. The pioneering work by Auerbach and Kotlikoff (1987) has been succeeded by a vast number of papers on social security reforms in overlapping-generations model. Most of these papers do not move beyond assuming exogenous retirement.⁷ Recently, both theoretical (see for example Ljungqvist and Sargent, 2014) and quantitative macro papers have focused on retirement behavior. Some early quantitative contributions are Hirte (2002), Fehr et.al. (2003) and Eiseensee (2006), all abstracting from DI. Díaz-Giménez and Díaz-Saavedrac (2009) and Imrohoroglu and Kitao (2012) are the ones most closely related to my paper.⁸ The former study explores the effects of a Spanish pension reform in a model with endogenous retirement, income uncertainty and disability. Unlike my paper, DI retirement is treated as a pure exogenous process. The second study consider endogenous retirement, as well as health and medical expenditure risks, in a model for the US. However, the public disability program (SSDI) is not modeled.

Many microeconomic life-cycle models has been developed to study social security and endogenous retirement (see for example Rust and Phelan, 1997 and Blau and Gilleskie, 2008). French (2005) estimates a full structural life-cycle model with income and health risk, and finds that the actuarial unfairness and work disincentives in the pension system seem to explain the U.S. retirement pattern well. Removing the earnings test causes individuals in the model to spend an extra year in the labor force. Some papers have studied retirement through disability insurance (Low et al., 2010 and Low and Pistaferri, 2015). However, these studies

⁷See e.g. Imrohoroglu et.al (1995), Conesa and Krueger (1999) , De Nardi et.al (1999)

⁸Erosa et.al (2011) builds on Imrohoroglu and Kitao (2012), to study cross country differences in labor supply late in the life cycle. Their study includes disability retirement as an exogenous event.

treat old-age retirement as exogenous at age 62. In independently developed work, Laun and Wallenius (2015) find, as this study, that including health and disability insurance is important when analyzing pension reforms.⁹ Bound et.al. (2010) and Iskhakov (2010) estimate DP-models of retirement with latent health indicators. However, they abstract from wealth and assume that consumption equals income every period. These studies are highly restrictive in that respect.

The model assumes that there is an economic margin among disabled individuals, i.e. the becoming a DI recipient is an endogenous decision. Assuming that the unhealthy can work is in line with the life-cycle models in e.g. French (2005) and Low et al. (2010). The former study has an explicit two-state health process, but no DI-program, while the latter models DI eligibility as a negative income shock.¹⁰ The importance of economic incentives is supported by the findings in Bratberg (1999). Using Norwegian registry data to estimate a multinomial logit model, the study finds that income opportunities is a significant determinant of labor force status, even after controlling for measures of health status. In a recent study, Bratsberg et al. (2010) report that job loss more than doubles the risk of permanent disability and accounts for three out of ten new DI claims in Norway.

3 Model

The economy is populated by J overlapping generations and a model period corresponds to a calendar year. Each cohort is a constant fraction of total population, which grows at a constant rate of one percent. Cohorts consist of a continuum of bachelor households starting their eco-

⁹Kitao (2014) develops a life cycle model of unemployment and disability, building on these studies. However, also this paper assumes mandatory old-age retirement. In contrast, my paper focuses explicitly on the interaction between DI and OA-pension retirement.

¹⁰A clear physical medical diagnosis among the disabled could cast doubt on the relevance of economic factors. In many cases, however, the health condition is self-perceived without any observable physical illness, making it harder to determine the true work limitation. Norwegian official medical diagnosis from 2000-2003 reveals that disability insurance was granted on mental, muscular, and skeletal disorders in 60 percent of the cases (see Mykletun and Knudsen, 2009).

economic life at age $j = 22$. Lifetimes are uncertain; agents survive to age j with probability p_j and die before age $j = J$. At the beginning of their economic life, agents are of different types, characterized by: education (no college, college), preferences towards work and income levels. Education heterogeneity maps into education specific life-cycle profiles for earnings and preference heterogeneity maps into idiosyncratic disutility of work. Different income types are represented by a fixed effect in earnings realized at age $j = 22$. Health and income risk generate within-type heterogeneity over the life-cycle. At each age j individuals are either in good or bad health, and health condition follows a two-state first order Markov process. Within-type income inequality is produced by idiosyncratic shocks to earnings, represented by an AR(1) process. Markets are incomplete and individuals hence rely on private saving and public social security for partial insurance against idiosyncratic health and income risk. In addition, there is a link between income, health and mortality. The low income-education type has a higher probability of going from good to bad health, and unhealthy individuals face a lower survival probability.

Agents choose how much to consume and save, and labor supply along the extensive margin. Retirement is an absorbing state and an option only for those eligible for public pension, either through disability pension or through old-age pension.¹¹ Eligibility ages are R and $D < R$ in the old-age pension and disability insurance program, respectively. Once entitled to OA-pension at age R , disability insurance is no longer an option and all DI-recipients are transferred into OA-pension. The benefit formulas are the same in both programs, but OA-pension is subject to an earnings test if working, while disability pension is received conditional on retirement. Health condition is observable and only unhealthy agents can retire with disability pension. When claiming DI, there is zero rejection probability, and eligibility is not reassessed if health condition

¹¹This is not a restrictive assumption. When simulating the model calibrated in section 4 while allowing for self-financed retirement (i.e. retiring either when (i) not entitled to DI nor OA-pension, or (ii) retiring without claiming DI despite being entitled to it), the retirement profiles are essentially unchanged. The reason is that retiring when not entitled to DI implies (i) lower future OA-pension, and (ii) no possibility of claiming DI in the future.

improves in the future.

The value function while employed is

$$V_j^W(a_j, m_j, e_j, I_j^b) = \max_{a_{j+1}, c_j} \{u(c_j, I_j^w = 1, I_j^b) + \beta s_{j+1, I_j^b} \mathbb{E}_j[V_{j+1}(a_{j+1}, m_{j+1}, e_{j+1}, I_{j+1}^b)]\}, \quad (1)$$

where s_{j+1, I_j^b} is the (health dependent) conditional survival probability. The expectation is over next period earnings and health condition, conditional on information available at age j . State variables are current assets, pension wealth, earnings and health condition (a, m, e, I^b) , and I^w is an indicator function, taking value 1 if working and 0 otherwise, while I^b is an indicator for bad health. The value function during retirement is

$$V_j^{NW}(a_j, m_j, I_j^b) = \max_{a_{j+1}, c_j} \{u(c_j, I_j^w = 0, I_j^b) + \beta s_{j+1, I_j^b} \mathbb{E}_j[V_{j+1}^{NW}(a_{j+1}, m_{j+1}, I_{j+1}^b)]\}, \quad (2)$$

and the unconditional value function is

$$V_j(a_j, m_j, e_j, I_j^b) = \max(V_j^W(\cdot), V_j^{NW}(\cdot)), \text{ if eligible for retirement}$$

$$V_j(a_j, m_j, e_j, I_j^b) = V_j^W(\cdot), \text{ if not eligible for retirement.}$$

Period utility of consumption and leisure is given by

$$u = \ln(c) - \delta_w I^w - \delta_b I^b I^w, \quad (3)$$

with δ_w denoting the disutility of work (heterogeneous across agents with mean μ_δ and dispersion σ_δ), and δ_b the disutility of bad health if working (common to all agents).

A utility cost labeled stigma is attached to receiving DI, interpreted as a reduced form, capturing all costs related to DI uptake. The cost is parameterized as additive cost of claiming, z_1 , and a flow cost, z_2 incurred every period receiving DI (until one is transferred to OA-pension at age R). Formally, stigma costs are included in equation 3 once on DI. Alternatively, these two costs can be combined to form an age-specific

fixed cost of DI-retirement at age j , \hat{z}_j , increasing in the number of years expected to be on DI:

$$\hat{z}_j = z_1 + z_2 \sum_{i=j}^{R-1} \beta^{i-j} \mathbb{E}_j \left[\prod_{l=j}^i s_{j+1, I_j^b} \right], \quad (4)$$

where the second term is the value of current and future flow cost, discounted back to age j , taking into account the expectation about future (health dependent) survival probabilities.¹²

Health risk is age-dependent and health condition is either good or bad. Let h_{j, I_j^b, I_{j+1}^b} denote the probability of a transition to health condition I_{j+1}^b next period, given health condition I_j^b at age j . The health condition parameter if healthy is $I^b = 0$, otherwise $I^b = 1$, and bad health is not an absorbing state (i.e. $h_{j,1,1} < 1$). The low income-education type has a higher risk of going from good to bad health at each age, denoted $\hat{h}_{j,0,1} = \chi h_{j,0,1}$.

At age j the agent receives an endowment of efficiency units per hour of $\ln(n_j) = q_j + e_j$. The first component captures the deterministic age-earnings profile, and the second component the stochastic endowment process, governed by:

$$\begin{aligned} e_j &= \alpha + v_j \\ v_j &= \eta v_{j-1} + \varepsilon_j, \quad v_{21} = 0, \end{aligned} \quad (5)$$

where $\alpha \sim N(0, \sigma_\alpha^2)$ is a fixed effect obtained age $j = 22$, and z is an AR(1) with innovation $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Given a wage rate, w , per efficiency unit, annual pretax labor income is (full time work corresponds to $H = 1725$ hours per year)

$$y_j = \exp(q_j + e_j) H w. \quad (6)$$

Both the age component and the stochastic component of earnings will be estimated separately for each education group.

¹²Note that in the value functions I do not have a separate DI claiming decision. This done to ease the computational burden. However, the assumption is not restrictive. First, working agents are not entitled to DI. Second, when simulating the model calibrated in section 4 allowing for retirement without claiming DI, no agents choose this option.

Pension wealth m_j is a function of annual pension accumulation. Each year over the working periods pension points m_j accumulates according to:

$$m_{j+1} = \begin{cases} m_j + \frac{M(y_j)}{20} & \text{if } j \leq 20 \\ m_j + \frac{\max(0, M(y_j) - m_j)}{20} & \text{if } j > 20 \end{cases}, \quad (7)$$

where annual pension accumulation M is a progressive function of labor income. In the actual Norwegian system total pension points m_j is the average of the twenty best earnings years. It is computationally infeasible to keep track of twenty years of earnings history. Hence, the function is an approximation in which pension wealth is revised upwards only if the pension claim is above average pension claims (as in French, 2003 and French, 2005).¹³ Upon retirement a pension benefit is calculated from $b_j = B(m_j)$.

Agents are born with zero assets and have access to a capital market yielding a risk-free rate of return on savings, denoted r . A zero borrowing constraint is imposed and assets evolves according to the sequence of budget constraints

$$\begin{aligned} a_{j+1} &= (1 + r(1 - \tau^a))a_j + T(j, y_j, b_j) - (1 + \tau^c)c_j \\ a_j &\geq 0, \end{aligned} \quad (8)$$

where τ^a and τ^c are capital gains and consumption tax rates and $T(j, y, b)$ after tax labor and pension income. The maximization problem in equations ?? and 2 is constrained by equations 7 and 8.

The economy is small and open. Agents face an exogenous world market interest rate r and wages grow at a constant rate g . A wage tax (τ^p) is levied on the firm (corresponds to the employer's contribution to the social security payroll tax).¹⁴ In order to transform the economy to a stationary one, I write all non-stationary variables as growth adjusted and replace the left-hand side of equation 8 with $a(1 + g)$.

¹³This approximation leads to a downward bias in pension benefits. To correct for this, the pension benefit increased, such there is no bias for an individual with an average path of income shocks over the life cycle.

¹⁴It is straightforward to microfound the constant wage growth g , using Cobb-Douglas production function in combination with perfect capital mobility and exogenous technological growth.

Each period the government collects taxes on capital, labor and pension income. If agents die with positive assets, the government collects the entire amount. The expenditure side consists of pension outlays and unproductive consumption X . In the benchmark model equilibrium, the budget is balanced by adjusting X . Later, when computing long-run effects of the reforms, the government uses the consumption tax rate to close the budget.¹⁵

The stationary equilibrium conditions of this small open economy are simple to characterize. Factor prices are constant, and when exogenous public consumption closes the budget, tax rates are constant as well. When tax rates are endogenous, equilibrium requires that maximizing individual behavior is consistent with the tax rate that balances the government budget.

3.1 Solution algorithm

Decision rules are found by backwards recursion. Since the retirement decision is absorbing I solve the problem in two steps. First I solve the retirement problem. Since there is no income uncertainty when retired, the problem is a simple consumption-savings problem with one endogenous state variable (asset) and two exogenous (pension benefit, health status). This gives me the retirement value function. Due to the discrete nature of labor supply, the value function is not necessarily globally concave (see e.g. Low et al., 2010), and working with the first-order condition can lead to local and not global optima. The value function for an employed agent is therefore solved by straightforward discretization of the state and control space. Health status is by definition discrete. For current and next period asset I use the same non-equally spaced grid, with denser grids for lower asset levels. The auto-regressive part of the income process is approximated by a two-state first order Markov process, following Tauchen and Hussey (1991). Fixed effects are represented by two types, a low and high income type, and the magnitude of α is set to match σ_α^2 , the dispersion of preference heterogeneity is approximated

¹⁵When unproductive consumption X is treated as exogenous, it grows at a constant rate equal to aggregate growth (i.e. indexed to productivity and population growth)

by 4 types. The pension wealth is discretized using a linear grid. When next period pension wealth is between grid points, it is determined by a lottery over the two neighboring grid points. Number of grid points are set to 200 and 15 for asset and pension wealth respectively. The model is simulated with 40000 individuals for each pair of income, preference and education type.

4 Calibration

Parameters are calibrated to male observations. Average conditional survival probabilities \hat{s}_j are taken from Statistics Norway's life table for males in 1987, and age is capped at 102. Social security rules reflect two different regimes, i.e. the systems in place in the 1987 and 2005. Regarding the preference parameters $(\beta, \mu_\delta, \sigma_\delta, \delta_b, z_1, z_2)$ I use the following calibration strategy: 5 out of a total of 6 parameters $(\beta, \mu_\delta, \sigma_\delta, \delta_b, z_1)$ are calibrated by matching 5 model moments simulated in the 1987 regime with data moments from the same time period. The calibration of the remaining parameter, the stigma cost z_2 , exploits the introduction of the ERP in the 1990s, and matches the drop in employment from 1987 to 2005. In both the 1987 and 2005 regimes the model simulation assumes a stationary state, implying that the the individual cross-sections and life cycles are the same.

The interest rate is set to $r = 0.04$ and wage growth to $g = 0.015$. I set capital gains tax rate, τ^a , equal to 28 percent, which is the current flat rate in Norway. The payroll tax rate, τ^p , is 13 percent. To obtain the consumption tax rate, τ^c , I compute the average rate based on the 2006 National Account. Total household indirect taxes divided by total household consumption gives $\tau^c = 0.19$. Labor and pension income is taxed according to the function T .

A. Social security 1987 regime The annual pension accumulation M in (7) is given by:

$$M(y) = \begin{cases} \max(0, \frac{y-G}{G}) & \text{if } y < K_1 G \\ K_1 - 1 + \frac{\min(y, 12G) - K_1 G}{3G} & \text{if } y \geq K_1 G \end{cases}, \quad (9)$$

where G is the unit of measurement, the *basic amount*, in the Norwegian social security system, calibrated to 15 percent of average annual labor income in age group 40-44 (see section 4D). Upon retirement, the function B transforms pension wealth to a annual pension benefit b according to:

$$B(m) = G + S(m). \quad (10)$$

Hence, b consists of an amount G that is independent of labor earnings, and an earnings-dependent *supplementary benefit*, $S(m) = GmK_2$. In the 1987 regime, the kink point K_1 in (9) and the replacement rate K_2 in (10) are 8.0 and 0.45, respectively. The basic amount G is indexed to wage growth g , which means that pensions are also indexed to wage growth. The eligibility age in the OA-system is $R = 67$. Since the focus in this paper is on old-age retirement, I set the eligibility age in the DI-system to $D = 51$. Both DI and OA-pension are determined by (7) and (9)-(10). However, those who qualify for OA-pension are allowed to work and receive benefits and if younger than 70, benefits are tested against earnings. When the sum of pension benefit and labor earnings ($b + y$) exceed 80 percent of previous earnings (denoted y^{prev}), benefits are reduced such that $b = \max(0, 0.8y^{prev} - y)$. In the Norwegian system previous earnings is the average of the past three years.¹⁶ Note that the benefits taxed away by the earnings test are lost, i.e. there is no upwards readjustment of future benefits.

To be entitled to the entire supplementary benefit $S(m)$, agents must work until age 70. Retirement at age 67, 68 and 69 induces a reduction in m of approximately 11%, 7%, and 3.5% respectively. See Appendix B for details and the exact expression. Note that the benefit for individuals

¹⁶For computational reasons I simplify by assuming that previous earnings is the average of deterministic labor earnings, i.e. $y_j^{prev} = wH\frac{1}{3}\sum_{i=j-3}^{j-1}\exp(q_i)$.

retiring on DI is calculated as if they retired at age 67.

B. Social security 2005 regime Pension entitlements are given by the same equations as in the 1987 regime, i.e. (7) and (9)-(10), but now with $K_1 = 6$ and $K_2 = 0.42$. In addition, there is no downwards readjustment of $S(m)$ associated with retiring between age 67-69. Moreover, there has been a relaxation of the pension earnings-test for ages 67-69. As of 2002, benefits are reduced by 40 percent of earnings above 2G; a clear improvement of work incentives (i.e. a reduction in implicit tax on labor income).¹⁷

In 1989, the opportunity to retire with OA-pension was extended to include people of age 66, and since 1998 the eligibility age has been 62. Roughly 60 percent of the work force is covered by this ERP (Holmøy and Stensnes, 2008). I divide households into two groups at age 22, those that are entitled to OA-pension from age 62 (60 percent) and those that must wait until age 67 (40 percent). The pension benefit is derived from (9) and (10), with $K_1 = 6$ and $K_2 = 0.42$ and computed as if the retiree worked until age 67. Hence, there is no benefit penalty associated with early retirement. Due to a *pro rata* earnings test, the implicit tax rate on labor income is higher before age 67 than after. For individuals between age 62-66 that are entitled to OA-pension and receive labor income, the pension benefit is reduced by the same proportion as the ratio of current labor income to previous income y_j^{prev} .¹⁸ Since there is no intensive labor supply margin in the model, the *pro rata* test basically implies that the entire benefit is taxed away if the agent continue working. As in the 1987 regime, benefits lost through the earnings test do not raise future benefits.

The DI rules are the same as in the 1987 regime and DI-recipients are transferred to OA-pension at age 67.

¹⁷Assuming a constant life-cycle labor income of 5.5G, after-tax income goes up by 27 percent relative to the 1987-regime. For the low and high education types in the model, after-tax income at age 67 increases on average with 40 and 18 percent, respectively.

¹⁸If current labor income is e.g. 90 percent of previous income, benefits are reduced by 90 percent.

C. Tax function The tax function $T(\text{age}, y, b)$ takes into account the progressivity in the Norwegian system and special tax rules for retirees. In general, the system differentiate between a general income tax and a social security tax. The income tax consists of a marginal tax rate scheme with three brackets. Income in the first bracket is taxed at a marginal rate of 0.28, while income in the second and third bracket is taxed at rates 0.37 and 0.40. The threshold levels are 6G and 9.8G, respectively. In addition, an (income-dependent) amount roughly between 0.5G and 1.46G is made tax exempt. The social security tax is levied on total labor and pension income, and is 0.078 on labor income and 0.03 on pension income. Finally, special tax rules applies for individuals who receive pension income. They can deduct an additional earnings-tested amount of 0.25G. If total earnings and pension income is below a threshold of roughly 2.1G, then no tax is paid, and total income tax is limited to 55 percent of income above this threshold. As a consequence, a retiree who receives only the minimum pension pays no income tax. The tax-favorable treatment of pension income contributes to an even more redistributive pension system, but also to a further weakening of work incentives, especially among low-income households.

D. Income process I define two education groups, those with a college degree and those with no college degree, and estimate an income process separately for each group. The stochastic component of the process in (5) is estimated using a 1997-2008 panel of hourly wage for Norwegian males in age group 30-50 with annual labor income above 1G. The sample age restriction is chosen to avoid the most severe selection problems. In the estimated income equation, log of wages are determined by a time effect, age, and a stochastic part. Let (i, j, t) index individual i of age j in period t . The income process is given by:

$$\ln(y_{i,j,t}) = \gamma_t + \text{age}_j + \omega_{i,j,t}, \quad (11)$$

Table 1: Income process estimates

	No college	College
ρ	0.736	0.676
$var(\alpha)$	0.049	0.072
$var(\varepsilon)$	0.003	0.005
$var(u)$	0.008	0.009

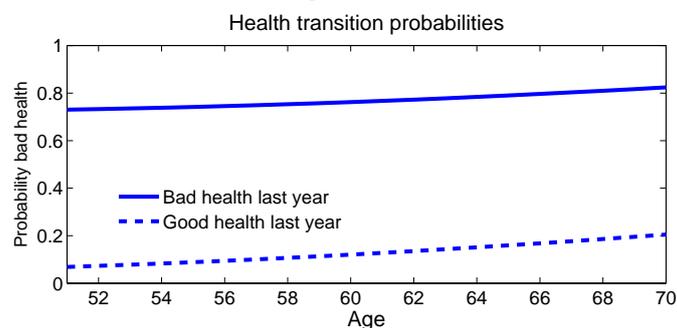
where $y_{i,j,t}$ is hourly wage, and γ_t and age_j are time and age dummies. The disturbance term $\omega_{i,j,t}$ is determined by:

$$\begin{aligned}\omega_{i,j,t} &= \alpha_i + v_{i,j} + u_{i,j} \\ v_{i,j} &= \rho v_{i,j-1} + \varepsilon_{i,j}, \quad v_{i,21} = 0,\end{aligned}\tag{12}$$

which is the empirical analog to the idiosyncratic income component in (5). In addition to the fixed effect and the AR(1) shock, an i.i.d. transitory shock $u_{i,j}$ is added to the empirical process. When simulating the model, the transitory shock is set to zero, which is consistent with interpreting the shock as measurement error. Parameter estimates (reported in table 1) are obtained by first running OLS on (11) and then fitting the income process to match the covariance structure of the OLS residuals. See appendix C for details.

For the deterministic age component of efficiency units, q , I take (11) and regress it on the same panel of hourly wage for males, for the broader age group 22-69. Due to selection, I cannot estimate the wage profile directly on observed wages. True wage growth is confounded with spurious wage growth caused by differences in the level and growth rate of wages between those who exit and remain in the labor market. However, since retirement is endogenous in the model, I can consistently estimate the offered wage profile using the model, providing a structural adjustment of the selection bias in the data (as in French, 2005). When calibrating the model I set efficiency units in age group 51-69 such that the model delivers the same average wage profile, conditional on working, as in the data. To be consistent with the social security system in place over the data sample period (1997-2008), the model correction is

Figure 1:



done in the 2005 regime.

Efficiency units are set to zero for individuals older than 69, implicitly assuming that agents do not work after reaching age 70. A relatively small fraction (7 percent) are observed participating at that age and an even smaller fraction work full time.¹⁹ Appendix C deals with details about the data, estimation strategy, and identification.

To pin down the social security replacement rate, the basic amount (G) is set such that the average annual labor income across agents in age group 40-44 is $6.6G$. From 1997 to 2008, the average full-time monthly wage in units of G for male workers of age 40-44 has fluctuated between 0.55 and 0.56.^{20,21}

E. Health The lack of panel data for health status in Norway precludes external calibration of the health process. I therefore adopt the health transition matrix in French (2005) estimated on US self-reported health status from the PSID. Figure 1 displays the conditional probability of being unhealthy for ages 50-69. The estimation accounts for both measurement error and individual heterogeneity.²² I assume that the health process starts at age 51. Prior to this age, all agents are in good health.

¹⁹Under the current Norwegian Working Environment Act, workers are protected from unfair dismissal, e.g. due to age. The upper age limit for this protection is 70.

²⁰I calculated yearly wage as $12 \times$ (monthly wage).

²¹Source: Table 05218 in Statbank, <http://statbank.ssb.no/>. Full-time equivalent monthly pay, part-time and full-time working males, age 40-44.

²²The figure corresponds to the smoothed versions in figure 1 in French (2005). The health measure is the response to the question: "Do you have any physical or nervous condition that limits the type or amount of work you can do?". The estimation is found in French (2001, mimeo) and used in French (2005).

The probability of becoming unhealthy at age 51 is set equal to 0.073. In French (2005), the unconditional probability of bad health at age 51 is 0.173. However, as explained below, the empirical employment profile is normalized by age 50 employment, thereby leaving out individuals not in the labor market at age 50 (roughly 11 percent, both in 1987 and 2005). The downward adjustment of $h_{50,good,bad}$ is based on the assumption that these have bad health.

When adopting the US health process, one possible concern is the cross-country variation found in surveys of self-assessed health status. These studies typically ask respondents to rate their own health according to a 5 point scale, ranging from "very good" to "very poor". However, using the European SHARE survey, Jürges (2007) finds that, after accounting for more objective measures of health,²³ most of the variation comes from reporting style, i.e. the connotation of the health categories differs across countries. The health measure used in French (2005) is based on the response to the question: Do you have any physical or nervous condition that limits the type or amount of work you can do?. It is likely that this measure is more in line with objective health measures than the conventional 5 point scale. Figure 10 in appendix A compares the life-cycle health profile in French (2005) with the Norwegian life-cycle profile, in which good health is measured as the fraction of people in a cross-section with 'good' or 'very good' health by age, using four waves of the Norwegian Level of Living surveys (1998-2008).²⁴

The Norwegian Institute of Public Health (2007) reports that males with income below the median had about 60 percent higher mortality rate than the above median group during the 1990s. Using three waves of the Norwegian Level of Living survey (1995, 1998, 2002), the same study documents large and significant differences in self-reported health between income groups. Roughly twice as many males in the bottom 1/3 of the (age-adjusted) income distribution rate their own health as

²³These (quasi) objective measures are 15 different physical conditions reported by the respondents

²⁴The Level of Living survey is a biannual cross-sectional representative survey with rotating themes. Health was the theme in 1998, 2002, 2005, 2008 and 2012 (see Statistics Norway, 2014) In these surveys, respondents are asked to rate their own health according to a 5 point scale: very good, good, neither good nor poor, poor and very poor.

below 'good', compared to the upper 2/3 (30 percent vs. 15 percent).²⁵ In the model, I allow for this correlation between income and health. However, because the measure of income includes transfers such as DI and unemployment benefits, these results could to some extent reflect a relationship between being on social security and having poor health. To control for this, I redo the analysis using four waves of the Level of Living survey (1998-2008), restricting the sample to males between 45-55 with more than 1G in labor income in the survey year, and in addition leave out those who have received DI or unemployment benefit the same year. The overall sample contains 1773 observations. 19 percent of individuals in the bottom 1/3 of the labor income distribution in the sample rate their own health as poor, compared to 11 percent in the upper 2/3. To account for this in the model, the low income and education type has a 66 percent higher probability of going from good to bad health than the other types ($\chi = 1.66$). Formally, $h_{j,0,1}$ is adjusted upwards for the type with low income and education, and downwards for the rest, keeping the average probability of going from good to bad health unchanged.²⁶

The relationship between bad health and mortality is also taken from French (2005). Assuming that the relative difference in age-specific mortality rates between healthy and unhealthy also holds for Norway, I adjust the health dependent mortality rates such that the average mortality rate is in line with Statistics Norway's life table. To illustrate, suppose that those in bad health at age j have $x\%$ higher mortality than those in good health. It follows that $s_{j+1,1} = (1+x)s_{j+1,0}$, and I then adjust $s_{j+1,0}$ downwards such that the average (over health status groups) mortality rate is equal to \hat{s}_{j+1} .

²⁵The same pattern emerges across education

²⁶When doing the readjustment I assume that 75% of individuals have no college, in line with the empirical ratio for males in age-group 45-55 between 1998-2008. Assuming that 75 percent have no college and given the estimated AR(1) process for wage shocks (which is implemented numerically as a 2-state first order Markov process), the type born with low fixed effect and low education accounts for the bottom 37.5 percent of the labor income distribution at age 50 in the model. The other types account for the upper 62.5 percent of the distribution. If I instead approximate the AR(1) process by a 20-state Markov process, the bottom 37.5 percent is still strongly dominated by the low fixed effect, low education type (94 percent).

F. Discount rate, disutility of work and bad health The discount rate β is pinned down by the ratio of aggregate wealth to labor earnings in age group 30-64. I use the ratio 2.15 which corresponds to the empirical ratio for males in 1993.²⁷ Due to differences between the actual 1993 male population structure and the stationary model population, the actual 1993 structure is used when aggregating the corresponding model moment.

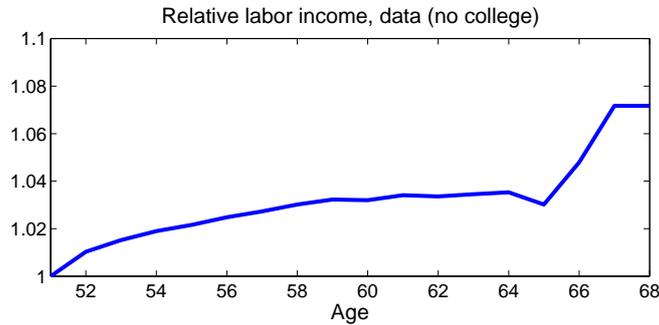
The disutility parameters μ_δ and δ_b target aggregate employment rates.²⁸ The empirical employment moments are calculated using the cross section of male annual labor income, collected from an administrative record covering all residents in Norway. Individuals are classified as retired when income is below the basic amount G (equal to 5300 USD in 1987). Income is defined as *pension effective income* and includes all income contributing to pension claims. In particular, it consists of all labor income, as well as unemployment insurance and sickness absence pay. When calculating the moments I normalize by age 50 employment, thereby leaving out individuals classified as non-participants at age 50 (11 percent). Dropping them is in line with the paper's focus on old-age employment, that is, the model does not attempt to explain retirement during prime-age working life. The disutility of bad health (common to all agents) δ_b is pinned down by the average employment rate at ages 62-66 in 1987.

To illustrate the identification of heterogeneous preferences, consider two versions of the model: with and without heterogeneity. The first version has no preference heterogeneity ($\sigma_\delta = 0$) and μ_δ is calibrated by matching the drop in average employment from ages 62-66 to ages 67-69 in 1987. A striking feature of this version is that it creates huge participation selection on income. The group of individuals still working at age 67, despite being eligible for OA-pension, consists almost entirely of high income types. Comparing the average annual labor income at age 51 with the average labor income at 51 for the sub-sample of individuals

²⁷The wealth data is collected from administrative records covering all residents in Norway. It includes both financial and housing wealth. 1993 is the earliest year with reliable individual level wealth data.

²⁸When aggregating over education types in the model, I use the empirical education distribution in 1987 (see figure 11, appendix A). Aggregation in 2005 takes into account the cross-sectional shift between 1987 and 2005.

Figure 2:



Notes: The graph displays the mean annual labor income at age 51, for individuals observed working at age 51+j, relative to mean annual labor income at age 51 for all individuals (no college degree). Data source: Norwegian administrative data, 1981 to 2008, covering the entire population. The age 51 mean income premium is estimated separately for 1981-1990 and the graph displays the average.

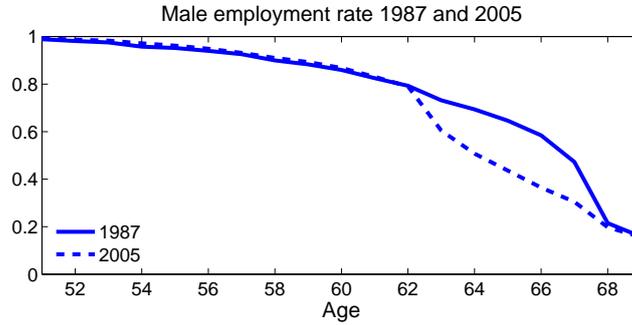
still working at 68, the latter group has on average 21 percent higher labor income. This income premium is much higher than what we observe in the data, where the premium, depicted in figure 2, is 7 percent.

Consequently, the model creates too much selection on income relative to data, indicating that income types explain too much of retirement. The model needs something that makes low income individuals retire late, and this paper considers preference heterogeneity as a solution. With heterogeneity, some agents will, despite low income and high social security replacement rates, choose to retire late simply because they don't dislike working quite as much, compared to other agents. It is a natural extension of the model, and the labor income data provides a clear identification. The degree of heterogeneity is calibrated by reducing the model selection on income to levels observed in the data for the low education group.²⁹

Formally, I assume that $x = \exp(-\delta_w)$ is distributed according to a Beta distribution $x \sim Be(\eta_m, \eta_s)$ with support $x \in (0, 1)$. Note that $-\delta_w$ is measured as disutility, hence $-\delta_w < 0 \Leftrightarrow 0 < \exp(-\delta_w) < 1$. The value

²⁹Adjusting also for the high education group would not change the results much, since the no-college group is by far the largest group. Among agents of age 67, roughly 90 percent have no college degree (in 1987)

Figure 3:



Notes: Male employment 1987 and 2005, normalized with age 50 employment

of x can be interpreted as follows: $(1 - x)$ is the utility cost of working measured as percentage of consumption. The utility function in equation 3, conditional on working, can be rewritten in terms of x ,

$$u = \ln(cx) - \delta_b l^b.$$

In addition to δ_b , the utility function is now parameterized by the mean (η_m) and the scale (η_s) of the Beta distribution, i.e. including preference heterogeneity adds one extra parameter to the model.³⁰ The mean is calibrated as before by matching the drop in average employment from ages 62-66 to ages 67-69 in 1987. The model selection on income is reduced to levels observed in the data by calibrating the scale such that the model simulated with the 1987 regime matches the mean wage at age 51 conditional on being employed at age 67 (for those with no college degree).³¹

G. Stigma Figure 3 displays two employment profiles for Norwegian males, in 1987 and 2005, both normalized by the age 50 employment rate.

The motivation for stigma is the following: In a model calibrated

³⁰The standard parametrization of the Beta distribution is in terms of two shape parameters s_1 and s_2 . I have chosen to re-parameterize in terms of the mean and scale, where these are related to the two shape parameters via $s_1 = \eta_m \eta_s$ and $s_2 = (1 - \eta_m) \eta_s$

³¹In the numerical implementation the Beta distribution is discretized using Gaussian quadrature with 4 nodes.

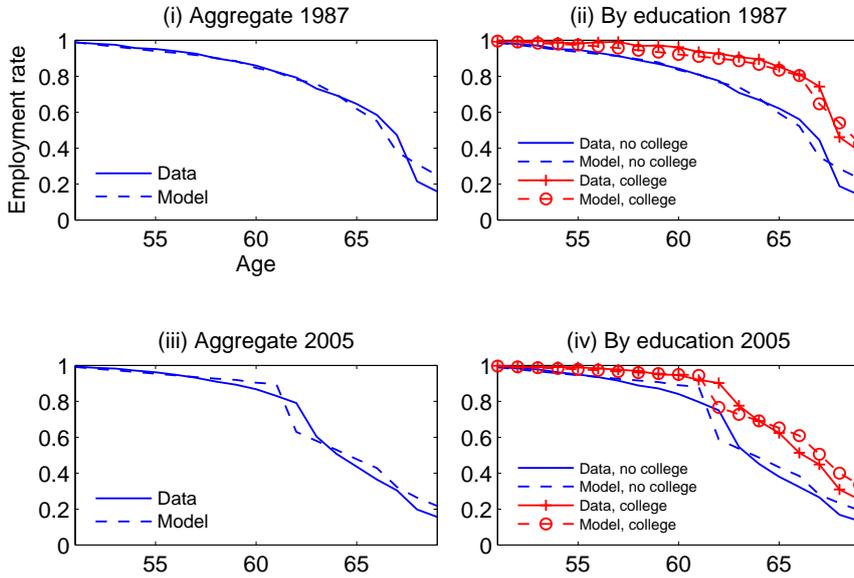
without such cost, the labor supply effect of going from the 1987 to the 2005 regime is not consistent with the data. In fact, labor supply increases (the simulated employment lies above the 1987 profile in figure 3). Two things explain this. First, given the underlying health process, 84 percent of all no-college agents and 76 percent of college agents have at least one year in bad health by age 67, and thus at one point have had the opportunity to retire with DI. Many of those who turn down the DI opportunity, will also turn down early-retirement benefits when implementing the 2005 regime. Second, the relaxation of the earnings test between age 67-69 associated with the 2005 regime induces an increase in labor supply. Overall, the ERP is not very attractive, and the model fails to account for the large downward shift in employment shown in figure 3. This changes with stigma cost. In addition to lowering the eligibility age for OA-pension, the ERP now also offers a way for unhealthy to bypass the stigma of DI.

The fixed cost of applying (z_1) is calibrated to match the average employment rate for ages 62-66 in 2005. The per-period cost (z_2) targets the average employment rate for ages 57-61 in 1987. The reason I model stigma with two parameters is twofold. First, if all stigma comes from the fixed cost z_1 , then retirement prior to age 67 occurs mainly at age younger than 60. If, alternatively, stigma is entirely due to the flow cost z_2 , most agents will retire at age 60-66. Both of these scenarios are inconsistent with the empirical 1987 employment profile, which shows a smoother decline in employment until age 66. By targeting age 57-61 employment, the calibration balances these two forces.

H. Summary of calibration This section summarizes the calibration. The following parameters are calibrated externally: The health process, survival probabilities, the (education specific) age-earnings profile q_j for age $j \in [22, 50]$ and income shock process, tax and pension systems. The following parameters are calibrated internally (targets in parenthesis): β (wealth/earnings ratio), δ_b (average employment rate at ages 62-66 in 1987), η_m (drop in average employment from ages 62-66 to 67-69 in 1987), η_s (mean age 51 wage conditional on still working at 67, in the 1987 regime), the age-income profile q_j for age $j \in [51, 69]$ (data observed

mean wage age 51-69), and the stigma parameters z_1 and z_2 (average employment at ages 57-61 in 1987 rate and average employment rate at ages 62-66 in 2005).

Figure 4: Employment, model vs. data



5 Results

5.1 The benchmark economy

This section presents the calibration results and the fit of the benchmark economy.

The mean and scale of the Beta distribution are calibrated to $\eta_m = 0.791$ and $\eta_s = 4.46$, and the disutility of bad health to $\delta_b = 0.70$. Expressed in terms of percentage of consumption, this correspond to a mean utility loss of 20.9 and 60.6 from working in good and bad health, respectively, with a standard deviation of 17.4 and 8.66. The discount factor is $\beta = 1.002$, and the stigma costs are calibrated to $z_1 = 0.05$ and $z_2 = 0.35$.

Employment The simulated aggregate employment profiles in the 1987 and 2005 regime are depicted in the panels (i) and (iii) in figure 4, along with the corresponding data profiles. Although only targeting average employment rates 1987,³² the model closely follows data employment at most ages. However, the model generates a bit too high employment at

³²The calibration fails to achieve exact match between the model and data moment at age 68.

the very end of the working career. Despite only matching average age 62-66 employment in 2005, the model is able to capture overall level and curvature quite well.³³ The model is calibrated assuming steady state both in 1987 and 2005. As figure 13 (appendix A) shows, the result is not sensitive to this assumption. Generating the 2005 employment profile by simulating the transition starting from the 1987 steady state creates only negligible differences.

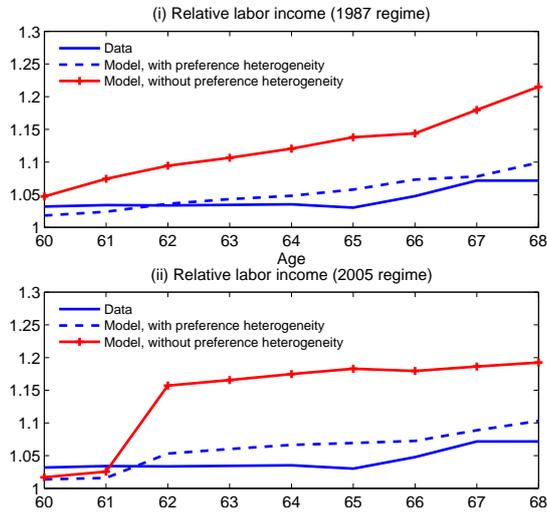
Interestingly, the model does a very good job at replicating the observed heterogeneity in retirement patterns across low and high education groups, depicted in panels (ii) and (iv) in figure 4. Note that the retirement age distribution across education groups are not targeted by the calibration. The most interesting case is the high education group, which due to its small share receives relatively low weight when targeting aggregate employment. Clearly, the model fits the relatively late retirement among individuals with high education. In the 1987 regime the model also captures the level and curvature, as well as the cross sectional drop to 2005. As a further evaluation of the model, consider the group of males still working at age 61 *and* entitled to early-retirement pension in the 2005 regime. In the data we observe that by age 66, 71 percent of these individuals received early-retirement pension in 2005.³⁴ The corresponding model moment is 68 percent.

The effect of preference heterogeneity Figure 5 is the same as figure 2, adding the income premium profiles for the two calibrated models (with and without preference heterogeneity). The model without preference heterogeneity produces a very clear mapping between income level and retirement age. Although the calibration routine only targets the age 67 premium using the simulated counterpart in the 1987 regime, the model with preference heterogeneity is successful in bringing selection on income closer to data also for other ages and for for the 2005 regime.

³³There is retirement timing discrepancy between the model and the data around the OA-eligibility age 67 in 1987 and 62 in 2005. This is due to the way I classify an individual as being retired in the data. An individual is defined as employed if earning more than 1G in the calendar year reaching the OA-eligibility age.

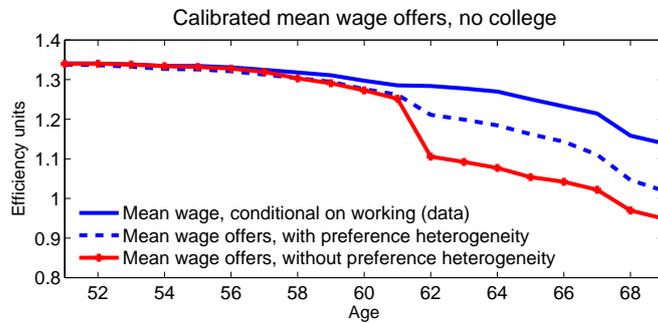
³⁴See Government White Paper no. 5 (2006-2007), p. 49. Among those retiring with early-retirement pension in 2005, 85 percent had full pension (i.e. full retirement)

Figure 5:



Notes: Simulated income premium at age 51, conditional on working at age $51+j$, with and without preference heterogeneity. Upper panel simulated on 1987 regime, lower panel simulated on 2005 regime. Calibration of dispersion in preference heterogeneity: match income premium conditional on working at age 67 in the 1987 regime with corresponding data moment.

Figure 6:



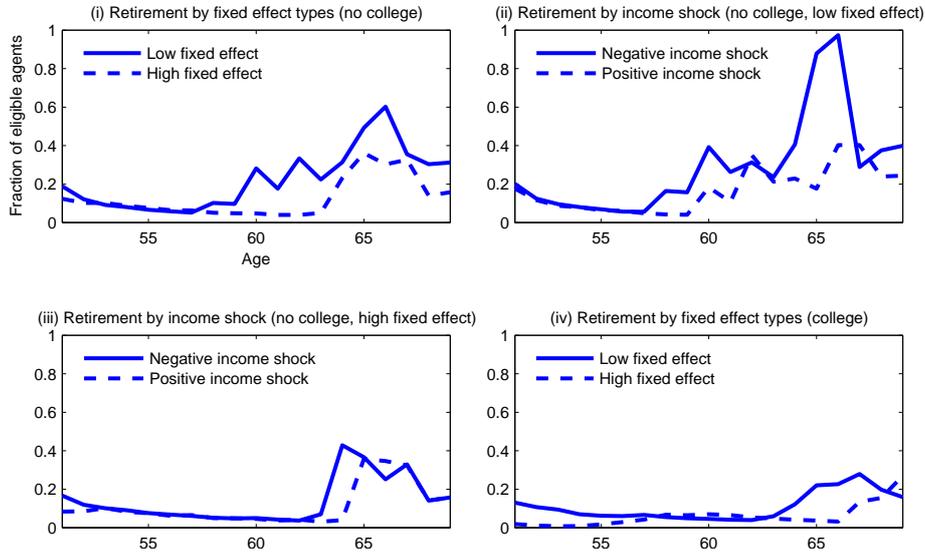
Calibrated mean wage offers are also affected by preference heterogeneity. Figure 6 shows the calibrated age-wage profiles for the low education type, in both model versions. In both versions, the wage offers are calibrated such that the participation selection on wages creates an age-earnings profile, conditional on employment, that corresponds to the empirical profile. Including preference heterogeneity in the model eliminates the sharp drop observed in the model with homogeneous preferences. Hence, retirement is not simply caused by a corresponding drop in wages. Overall, a negative adjustment is nevertheless needed for the model to be consistent with data. The full life cycle profiles are depicted in figure 12 in appendix A.

The retirement decision The disutility of bad health (δ_b) is small enough to prevent DI-retirement from being, de facto, an exogenous event. In fact, in the 1987 regime, 18 percent of agents working at age 66 are entitled to DI. A sufficiently large δ_b would induce agents to retire as soon as they experience poor health and the exogenous health process would in effect determine employment rates. Since 84 percent of no-college and 76 percent of college individuals have at least one year in poor health by age 66, model employment rates would then be too low relative to the data.

Since health alone is not enough to induce labor market exit, the retirement distribution must be explained by other factors too. Consider the 1987 regime. Focusing on those with no college degree, panel (i) of figure 7 shows the fraction of eligible agents choosing retirement at each age, by fixed effect types.³⁵ As noted in section 3, poor health is required for DI-eligibility. In spite of high degree of preference heterogeneity, early retirement is dominated by low-income (low fixed effect) types. Average retirement age is 63.7 and 66.5 for the low and high income type, respectively. Within both income groups there is also selection on current idiosyncratic income shocks, shown in top-right and bottom-left panel. Those with a negative income shocks are much more likely to retire early.

³⁵For example, consider the group defined by low-education, age 58, with low fixed income effect who has 1) not yet retired and 2) in poor health. The top left panel shows that among these agents 11.2 percent choose to retire.

Figure 7:



Notes: Model selection on income types (fixed effect) and income shocks into retirement. Simulated with 1987 regime

A characterizing feature of the model is that it fits the drop in employment around the OA eligibility age 67 in the 1987 regime. This is not a surprise considering the fact that the calibration targets the drop from age 66 to 68. It is, however, interesting to understand why we get a peak in retirement exactly at age 67. Since most of the action takes place among the low-education type, I focus on this group. No-college employment drops from 0.55 to 0.38. As the drop in wages is negligible, the work incentives does not change much from age 66 to 67, given poor health. But at age 67, the unhealthy can retire with OA-pension and thus avoid stigma cost of DI. This is indeed the most important explanation. Simulating the model with stigma extended to age 69,³⁶ reduces the fraction retiring at age 67 from 17 to 9 percent.

³⁶I do this by adding a cost $z_1 + z_2$ to the value of choosing retirement at age 67, 68 and 69.

Table 2: Government budget components

DI and OA-pension, % of Gov't expenditures	20.8
Consumption tax revenue, % of Gov't expenditures	20.3

Notes: Selected government budget components in the initial steady state

5.2 Experiments

This section uses the calibrated model to analyze the macroeconomic implications of pension reforms, justified by increased average life-expectancy at age 67, by a factor of $1/0.8 = 1.25$, from 13.9 years to 17.4.³⁷ Before the demographic shock occurs the economy is in a stationary equilibrium with a matured version³⁸ of the 1987 pension regime in place, assuming a stationary education distribution where 25 percent have a college degree. The key government budget components in the initial equilibrium are summarized in table 2. In the new stationary equilibrium the consumption tax rate adjusts to close the government's budget.³⁹ If no pension reform is undertaken, the consumption tax rate increases by 4.2 percentage points, from 19.0 to 23.2 percent.

Two qualitatively different reforms are considered, one which combines OA-benefit reduction and earnings-test removal (reform 1), and one in which the OA-pension eligibility age is increased (reform 2).

Reducing Old-Age pension benefits and remove earnings test In reform 1, OA benefits are reduced by 20 percent. The expected life-time, pre-tax pension bill of someone who receives OA-pension from age 67 is therefore roughly unchanged, despite increased life-expectancy.⁴⁰ Moreover, the OA-pension earnings test is completely removed, so agents can work *and* receive full OA-pension. These two reform elements mimic the main

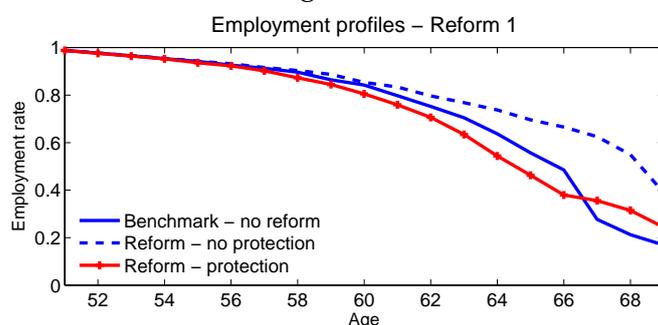
³⁷Even though the agent's life-expectancy improves, it is assumed that health and productivity remains unchanged.

³⁸i.e. the early retirement penalty as described in Appendix B is removed

³⁹Exogenous government expenditures (X) are taken from the initial steady state

⁴⁰This holds exactly if the wage growth rate is used as discount factor (which is the discount factor used by Norwegian government in the actual 2011 pension reform)

Figure 8:



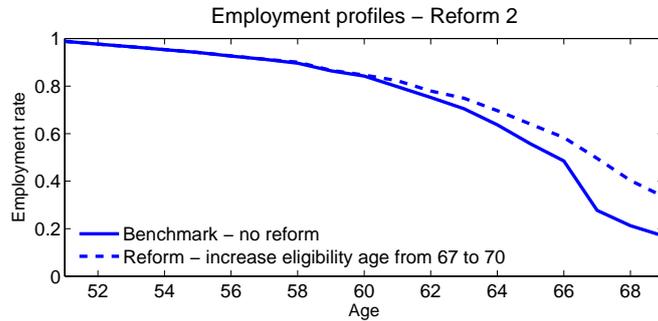
Notes: Employment effects of reform 1: Reducing OA-pension and removing earnings test

features of the Norwegian pension reform of 2011. Finally, individuals transferred from DI to OA-pension at age 67 are protected from benefit cuts. While the benefit cut and earnings-test removal create an incentive for the everyone to delay retirement, the protection creates an incentive for the unhealthy to retire with DI prior to age 67. This captures the trade-off in the Norwegian pension reform between maintaining insurance and increasing labor supply. This version of reform 1 is denoted reform 1 - protection. To single out the degree of substitution between the two social security programs, I also consider a reform in which DI recipients are not protected (reform 1 - no protection).

Figure 8 shows the effects on employment. When there is no protection, the response is positive. In age group 51-69, labor supply goes up by 7.7 percent and the consumption tax rate is 17.3 percent (compared to 23.2 percent in the no-reform scenario). These effects are large, but the reform is also quite drastic. In addition to a benefit cut of 20 percent, we go from a system in which most agents lose their entire OA-benefit if working, to a system in which everyone keep 100 percent. Hence, the labor supply distortion associated with the earnings test is completely removed.

When protecting those coming from DI, the substitution between DI and OA program is so large that the potential fiscal gain is wiped out. Aggregate labor supply now declines by 2.1 percent, while the consumption tax rate is 23.4 percent. Unhealthy individuals who previously chose

Figure 9:



Notes: Employment effects: Increase OA eligibility age from 67 to 70

to work, are now retiring to such extent that it offsets the increased participation by healthy individuals, and as a result, the consumption tax rate is higher compared to the no-reform scenario.

Increasing eligibility age Reform 2 raises the eligibility age for OA-pension by three years, from 67 to 70; justified by the same increase in life-expectancy as in the previous reform. Consider the following *static* response for agents retiring at 67. Healthy agents now have to work, while the unhealthy switch to DI pension if it is worth the stigma cost. The *dynamic* response comes from forward looking agents. Consider a 66-year-old agent in poor health. Before the reform he is working. Conditional on bad health, the budget constraint after the reform remains unchanged, i.e. retirement with DI at age 67 is still an option. If health status changes, the agent is, however, no longer eligible for OA-pension. This might encourage DI retirement at age 66. On the other hand, when raising the OA eligibility age the number of years on DI goes up, causing the discounted sum of stigma costs to go up (from equation 4), and thereby discouraging DI retirement.

Figure 9 shows, however, that the reform causes an increase in age 66 employment, caused by the increased (discounted sum of) stigma cost. Between ages 67-69 unhealthy agents now switch from OA-pension to DI, while healthy agents continue to work. Overall, labor supply goes up by 4.9 percent, while the consumption tax rate falls by 2.2 percentage

Table 3:

Reform	Endogenous DI	Exogenous DI
Reform 1 - no protection	7.7	4.1
	-25.4	-20.1
Reform 1 - protection	-2.1	4.1
	1.0	-12.7
Reform 2	4.9	4.1
	-9.5	-8.6

Notes: Reform effects, percentage change in aggregate labor supply (above) and consumption tax rate (below), relative to the no-reform scenario. First row: Reform combining OA-pension benefit cut and earnings test removal. The second row refers to the same reform, but in which those coming from DI are protected against OA-pension benefit cuts. The third row refers to the reform that increases the OA-eligibility age.

points relative to the no-reform scenario.

Exogenous disability To evaluate the quantitative importance of endogenous DI retirement, it is useful to compare the above results to a model economy with exogenous DI retirement. Bad health is now an absorbing state and the disutility of working when in poor health (δ_b) is infinity. The age-wage profiles are taken from the endogenous DI model. The probability of a bad health shock at age t , conditional on good health at age $t - 1$, is calibrated to match the simulated employment rates at age 51-66 in the 1987 regime of the endogenous DI model.⁴¹ From age 67, the probability is set equal to the age 66 probability. The parameters of the δ_w -distribution, $Be(\eta_m, \eta_s)$, and β target the corresponding simulated moments from the endogenous DI model, i.e. same wealth-to-earnings ratio, age 68 employment rate, and average labor income premium at age 51 for those still working at age 67-69. The calibration produces the same β as before, and a mean utility loss of working of 25.5 percent of consumption, with a standard deviation of 5.6. With exogenous DI the stigma cost does not play a role.

Table 3 summarizes the results. In the exogenous DI model, reform

⁴¹I do this for each education group separately

effects are muted by the fact that many are not able to respond to incentives. Roughly 40 percent have no ability to work at age 66, and given the extrapolation of health risk, almost 55 percent at age 69. Nevertheless, the consumption tax rate goes down, and labor supply goes up in all reform scenarios.⁴² With endogenous DI take-up, in contrast, there is no fiscal gain in reform 1 - protection. On the other hand, both reform 2 and reform 1 - no protection produces a somewhat larger fiscal improvement in the endogenous DI model than in the model with exogenous DI.

Interestingly, the calibration of the exogenous DI model induces a much smaller dispersion in preferences. This is because with exogenous DI there is much less selection on income into retirement, and consequently less need for preference heterogeneity in order to match the labor income premium of late retirees. A low degree of preference heterogeneity pushes up the labor supply elasticity with respect to pension rule (among individuals in good health). This explains why the exogenous DI model produces almost as large fiscal gain as the endogenous DI model in the eligibility-age reform.

Welfare effects of an early retirement program High stigma cost of DI suggests that there may be scope for early retirement programs to improve welfare. These programs offer an easy and quick way of leaving the labor force, without the effort and social stigma attached to the DI route. In this respect, an early retirement program works a lot like removing the stigma cost. Clearly, introducing a universal early retirement program assuming no labor supply response (i.e. the only effect is that DI-recipients switch to early retirement pension), would raise welfare. However, the behavioral response (i.e. early retirement), triggers a tax increase which could outweigh the welfare gain of reduced stigma cost. This is indeed the result. With an ERP, granting universal access to OAPension from age 62, the consumption tax rate goes up by 5.3 percentage points. Almost all agents are made worse off in the new stationary equilibrium. The exception is the type with high disutility of work, which gains up to 0.5 percent of annual consumption. This type accounts for

⁴²In the exogenous DI model, the consumption tax rate in the no-reform scenario is 23.3 percent.

2.1 percent of the total population. For the other 97.1 percent, the type-dependent welfare loss varies between 1.5 and 4.3 percent of annual consumption.

6 Conclusion

Increasing life-expectancies are putting pressure on old-age pension systems, and reforms are necessary. A key target is to make people work longer and postpone benefit uptake. When designing a reform, policy makers must consider the interaction with other welfare programs that provide income insurance. In this paper I have looked at the interaction between DI and OA-pension and the effects on labor supply and public finances of reforming the OA pension system. Across the OECD, a large share of people between 50-64 receive DI, making it an important path into retirement. In Norway, which is the basis of this study, more than 40 percent of the population relies on DI when transferred to old-age pension at age 67. The main contribution of this essay is to extend the standard macro pension model by considering DI as an alternative retirement margin.

A concern when reforming OA-pension is how to treat those coming from DI. I find that protecting DI-recipients from benefit cuts induces a large increase in DI-take up. In one reform experiment, old-age employment falls despite large improvements in work incentives. The increased labor supply among healthy individuals is outweighed by a surge in DI take-up among unhealthy individuals. Conducting the same reform experiment in a model with exogenous DI take-up (implicitly assumed in all other studies) results in increased labor supply. Consequently, there is a risk of biasing the analysis of pension reforms if DI as a retirement margin is not properly accounted for.

References

- [1] Auerbach, A.J., and L.J. Kotlikoff (1987): *Dynamic Fiscal Policy*. Cambridge University Press, Cambridge.

- [2] Blau, D.M. and D.B. Gilleskie (2008): "The role of retiree health insurance in the employment behavior of older men", *International Economic Review*, 49(2), 475-514.
- [3] Börsch-Supan, A., Hank, K. and H. Jürges (2005): "A new comprehensive and international view on ageing: Introducing the Survey of Health, Ageing and Retirement in Europe", *European Journal of Ageing*, 2(4), 245-253.
- [4] Bound, J., Stinebrickner, T. and T. Waidmann (2010): "Health, economic resources and the work decision of older men", *Journal of Econometrics*, 156(1), 106-129.
- [5] Bratberg, E. (1999): "Disability Retirement in a Welfare State", *The Scandinavian Journal of Economics*, 101(1), 97-114.
- [6] Bratsberg, B., Fevang, E. and K. Røed (2010): "Disability in the Welfare State: An unemployment problem in Disguise?", IZA Discussion Paper No. 4871.
- [7] Conesa, J.C. and D. Krueger (1999): "Social Security Reform with Heterogeneous Agents", *Review of Economic Dynamics*, 2(4), 757-795.
- [8] De Nardi, M., Imrohoroglu, S. and T.J. Sargent (1999): "Projected U.S. demographics and social security", *Review of Economic Dynamics*, 2, 575-615.
- [9] Díaz-Giménez, J. and J. Díaz-Saavedrac (2009): "Delaying retirement in Spain", *Review of Economic Dynamics*, 12(1), 147-167.
- [10] Eisensee, T. (2006): "Essays on public finance: Retirement behavior and disaster relief", Stockholm University, Monograph Series No. 54.
- [11] Erosa, A., Fuster, L. and G. Kambourov (2011): "A Theory of Labor Supply Late in the Life Cycle: Social Security and Disability Insurance", Unpublished Working Paper.

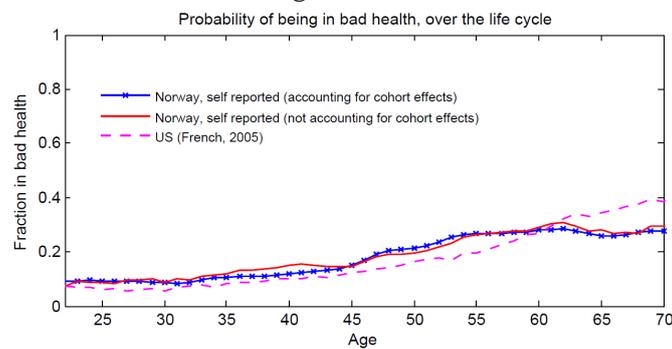
- [12] Fehr, H., Sterkeby, W.I. and Ø. Thøgersen (2003): "Social security reforms and early retirement", *Journal of Population Economics*, 16, 345-361.
- [13] French, E. (2001): "How severe is measurement error in health status: Evidence from the PSID", (Mimeo)
- [14] French, E. (2003): "The Effects of Health, Wealth, and Wages on Labor Supply and Retirement Behavior", Chicago Fed Working Paper 2000-12.
- [15] French, E. (2005): "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour", *The Review of Economic Studies*, 72(2), 395-427.
- [16] Hirte, G. (2002): "Welfare and Macroeconomic Effects of the German Pension Acts of 1992 and 1999: A Dynamic CGE Study", *German Economic Review*, 3(1), 81-106.
- [17] Holmøy, E. and K. Stensnes (2008): "Will the Norwegian Pension reform reach its goals? An integrated micro-macro assessment", Statistics Norway Discussion Paper No 557.
- [18] Imrohoroglu, S. and S. Kitao (2012): "Social Security Reforms: Benefit Claiming, Labor Force Participation and Long-run Sustainability", *American Economic Journal: Macroeconomics*, 4(3), 96-127.
- [19] Imrohoroglu, A., Imrohoroglu, S. and D.H. Joines (1995): "A life cycle model of social security". *Economic Theory*, 6, 83-114
- [20] Iskhakov, F. (2010): "Structural Dynamic Model of Retirement With Latent Health Indicator", *Econometrics Journal*, 13(3), 126-161.
- [21] Jürges, H. (2005): "True Health vs Response Styles: Exploring Cross-country Differences in Self-reported Health", *Health Economics*, 16(2), 163-178.
- [22] Kitao, S. (2014): "A life-cycle model of unemployment and disability insurance", *Journal of Monetary Economics*, 68, 1-18.

- [23] Laun, T. and J. Wallenius (2015): "A life cycle model of health and retirement: The case of Swedish pension reform", *Journal of Public Economics*, 127, 127-136.
- [24] Ljungqvist, L. and T.J. Sargent (2014): "Career length: Effects of curvature of earnings profiles, earnings shocks, taxes, and social security", *Review of Economic Dynamics*, 17(1), 1-20.
- [25] Low, H. and L. Pistaferri (2015): "Disability risk, Disability Insurance and Life Cycle Behavior", *American Economic Review*, 105(10), 2986-3029.
- [26] Low, H., Meghir, C. and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle", *American Economic Review*, 100(4), 1432-1467.
- [27] Mykletun, A. and A. Knudsen (2009): "Tapte arbeidsår ved uførepensjonering for psykiske lidelser. En analyse basert på FD-trygd", The Norwegian Institute for Public Health, Report 2009:4, (Norwegian only).
- [28] Norwegian Institute of Public Health (2007): "Sosial ulikhet i helse. En faktarapport". Næss Ø., Rognerud. M and B.H. Strand (ed.), Report 2007:1.
- [29] Norwegian Ministry of Labour (2006): "Opptjening og uttak av alderspensjon i folketrygden", Green paper No. 5 (2006-2007), (Norwegian only)
- [30] Norwegian Ministry of Labour (2011): "Endringer i folketrygdloven (ny uføretrygd og alderspensjon til uføre)", Prop. 130L (2010-2011), (Norwegian only)
- [31] OECD (2010): "Sickness, disability and work: Breaking the barriers", Organisation for Economic Co-operation and Development, Paris, 2010.
- [32] Rust, J. and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets", *Econometrica*, 65(4), 781-831

- [33] Statistics Norway (2005): NOS D 362 Wage statistics 2005, Oslo.
- [34] Statistics Norway (2014): Level of living 2009-Cross sectional study-Working conditions. Study Documentation.
- [35] Tauchen, G. and R. Hussey (1991): "Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models", *Econometrica*, 59, 371-396.

A Figures

Figure 10:



Notes: comparing the fraction in bad health over the life-cycle, US (French, 2005) and Norway. Data source: Norwegian Level of Living (1998,2002,2005,2008), cross-sectional surveys. Good health defined as self-reported health status 'good' or 'very good'. The profile is smoothed, using a three year moving average.

Figure 11:

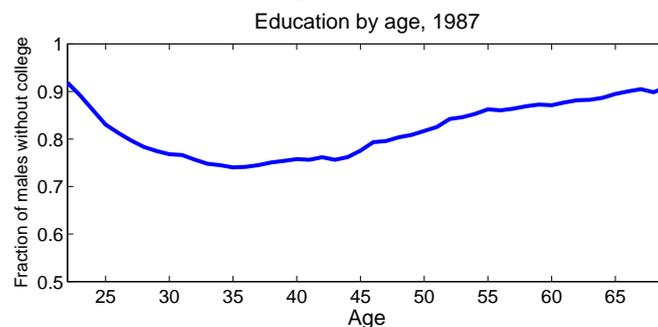
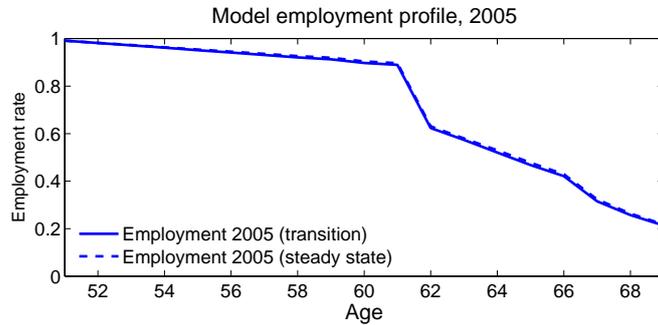
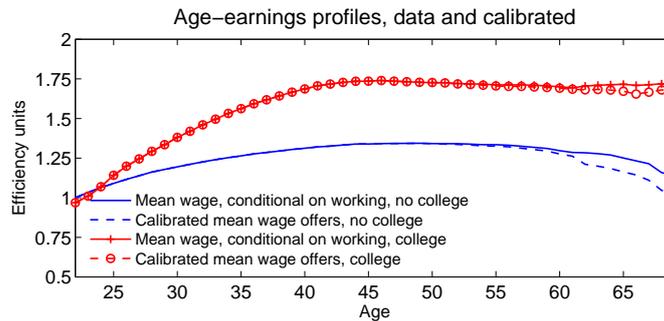


Figure 13:



Notes: Solid line refers to the model simulated 2005 cross-sectional employment profile, when simulating the transition starting in 1987, from a steady state with the 1987 pension regime in place. The transition is initiated with the immediate implementation of the 2005 pension regime, and the model is simulated forward until year 2005.

Figure 12:



Notes: Normalized by age 22 mean wage, no college

B Social security systems in 1987 and 2005

The penalty when retiring between 67-69 in 1987, reflects fact that 40 years of earnings above G is required to qualify for full supplementary pension $S(m)$ in the Norwegian National Insurance System (NIS). Because the system was introduced in 1967, the maximum history attainable was 20 years in 1987, which would only give 50% supplementary pension (20/40). The following scheme was implemented to compen-

sate for lack of full earnings history. Let L denote an individual's number of years with earnings above 1G since the introduction of NIS in 1967, and F the number of years required to get full supplementary pension. Upon retirement, the retiree is entitled to a supplementary pension of $S(m)\frac{L}{F}$. For individuals older than 50 years in 1967, $F = 20$, while $F = \min(0, 20 + 50 - j)$ for individuals of age $j \leq 50$. This implied that individuals were granted full pension as long as they earned more than 1G at each age up to and including 69. These compensation rules applied, however, only to pension entitlements up to $m = 4.0$. With $m > 4$ the retiree would be entitled to $\frac{L}{F}S(4) + \frac{L}{40}S(m - 4)$. Each generation hence faced a cohort specific pension system. Due to the stationary environment in the model, each cohort face the same pension scheme. The penalty of retiring at age 67-69 in the 1987 regime in the model is derived from the perspective of the age 62 cohort in 1987 (i.e. $F = 28$), and the supplementary pension after adjusting for the penalty is given by $\frac{L}{28}S(\min(m, 4)) + \frac{L}{40}S(\max(m - 4, 0))$, where $L = 28$ if retiring at age 70 and $L = 25$ if retiring at age 67.

In contrast, in the 2005 regime all cohorts in the model are entitled to full supplementary pension from age 67. This is in line with the the actual pension scheme facing cohorts of age 65 or less in 2005.

C Income process

The data source used for estimating the age-wage profiles and the income process consists of monthly earnings (consisting of basic salary, fixed and variable additional allowances, bonuses and commissions, and overtime pay) and contractual monthly working hours from 1997-2008. Data is collected once a year during the 3rd quarter. Hourly wage is computed as the ratio of monthly earnings (excluding overtime pay) to monthly hours. The sample covers all employees in the public sector, and large private sector firms (100-150 employees, depending on industry), while employees in small and medium sized firms are sampled each year with a sampling rate between 10-50 percent, depending on industry. Between 50-65 percent of all private sector employees are sampled (see Statistics Norway, 2005).

The income process is estimated by fitting the observed covariance structure of residuals $(\omega_{i,j,t})$ obtained from running OLS on equation 11. Dropping time and individual specific subscripts, and noting that j denotes age, the empirical income process used in the model (equation 12) is:

$$\begin{aligned}\omega_j &= \alpha + v_j + u_j \\ v_j &= \rho v_{j-1} + \varepsilon_j, \quad z_0 = 0\end{aligned}$$

The vector of parameters to be estimated is $\theta = \{var(\alpha), var(u), var(\varepsilon), \rho\}$. Computing variances and covariances gives

$$\begin{aligned}var(\omega_j) &= var(\alpha) + var(v_j) + var(u_j) \\ cov(\omega_j, \omega_{j+k}) &= var(\alpha) + \rho^k var(v_j) \\ cov(\omega_j, \omega_{j+1}) - cov(\omega_j, \omega_{j+k}) &= \rho(1 - \rho^{k-1})var(v_j),\end{aligned}$$

where $var(\omega_j)$ is the age j variance of the residuals and $cov(\omega_j, \omega_{j+k})$ the covariance of residuals between age j and $j+k$. For a given k , the identification of ρ is given by

$$\frac{1 - \rho^{k-1}}{1 - \rho} = \frac{\sum_j [cov(\omega_j, \omega_{j+1}) - cov(\omega_j, \omega_{j+k})]}{\sum_j [cov(\omega_j, \omega_{j+1}) - cov(\omega_j, \omega_{j+2})]}$$

where the summation is over all ages $j \in [30, 50 - k]$. The variance $var(u)$ is given by

$$var(u) = \frac{1}{N_k} \sum_j [var(\omega_j) - cov(\omega_j, \omega_{j+k}) - (1 - \rho^k)var(v_j)]$$

where $N_k = 50 - k - 30 + 1$ and

$$var(v_j) = \frac{cov(\omega_j, \omega_{j+1}) - cov(\omega_j, \omega_{j+2})}{\rho(1 - \rho)}$$

Identification of $var(\alpha)$

$$var(\alpha) = \frac{1}{N_k} \sum_j [var(\omega_j) - var(v_j)] - var(u)$$

Identification of $var(\varepsilon)$

$$var(\varepsilon) = \frac{1}{N_k} \sum_j \left[var(\omega_{j+1}) - var(\omega_j) + (1 - \rho^2) var(v_j) \right]$$

I do this for $k = 3, \dots, 9$, producing 7 values for θ (one for each k). I take the mean over all k 's to obtain estimates of $var(\alpha)$, $var(u)$, $var(\varepsilon)$, and ρ .