Sectoral Choice with Human Capital and Accumulation of Pension Benefits

Erik Ø. Sørensen*
Department of Economics
Vrije Universiteit Amsterdam

Work in progress, August 2007.

Abstract

Universal pension plans and large public-sector workforces affect accumulation and allocation of human capital. The benefit reforms and re-training programs being considered in many countries are likely to affect behavior in ways that can only be analyzed within forward-looking models of lifetime labor supply. Using Norwegian panel data on three birth cohorts, this paper develops and estimate a life-cycle model of public- and private-sector employment. The model of sequential career-choices builds on Keane and Wolpin (1997), extending the accumulation of skills to be sector-specific and allowing unobserved, non-deterministic depreciation of skills. The Norwegian retirement benefits are modeled in a way that builds on the discretization approach of Rust and Phelan (1997), and identification is aided by exploiting how current career-choices affect future expected benefits. I find important heterogeneity in skill accumulation. The model is used to analyze the effect of a pension reform on sector-specific labor supply. The reform has large effects on labor supply, but the sectoral effects are small.

*email:sameos@gmail.com. I am grateful for encouragement and very helpful advice from my advisors Christopher Ferrall and Kjell G. Salvanes. I have also benefited from helpful comments by Christian Belzil, Gerard J. van den Berg, Anne M. Danø and Roine Vestman.

This paper constructs a model of life-cycle labor supply in which individuals accumulate sector-specific skills by attending school or gaining work experience, and in which individuals run the risk of losing skills that are not used. Individuals can choose to specialize in one of the two sectors, with an expected cost of switching sectors that depends on the skill technology and career history. I estimate the model on a panel of three Norwegian birth-cohorts, using the incentives of sector-specific retirement benefits to aid identification. This framework allows evaluation of pension reforms and adult education initiatives. These policies are under consideration in a number of countries as welfare states prepare for ageing populations and react to increasing public pressure to modernize and downsize their public sectors.

Preparing for demographic changes about to come, universal pension plans financed out of current taxation has been the target of reform plans in most developed countries. Such reforms will have consequences for more than government budgets; implicit claims on future benefits constitute a large fraction of households net wealth (Domeij and Klein 2002) and have been claimed to have important effects on life-cycle labor supply (Rust 2001). In addition to the universal plans, large public sector workforces have often accepted generous future supplementary benefits in exchange for a lower level of current earnings. Changes to the universal plans will interact with the public sector supplementary benefits and could have large effects on recruitment to the public sector and labor mobility between the public and the private sector. Quantification of these effects can only be made within a life-cycle model of labor supply.

I follow Keane and Wolpin (1997) in modeling the occupational choice problem as a net present value maximization, with the decision to work being discrete. Individuals take into account the effect current career choices have on future retirement benefits. The paper departs from the baseline model of Keane and Wolpin in making the occupational choice one between the public and private sector rather than a choice between blue-collar/white-collar and military employment. This is necessary in order to analyze policies directed at the public sector employment and benefit reforms. The public workforce also accounts for a very large proportion of total employment in most European welfare states. In Norway it accounted for as much as 33% of employment in 2000 (Statistics Norway 2001, Table 244). There is also evidence that wage-education profiles are different in the public and private sector in Norway even after controlling for a large number of other factors (Hægeland, Klette, and Salvanes 1999).

The paper shares the goal of Rust and Phelan (1997) of modeling the effect of retirement benefits on behavior. It also borrows some of the methodology of modeling the dynamics as controlled Markov processes on rather coarse grids. The state-space can be factored such that some of these Markov processes can be pre-estimated. It differs from Rust and Phelan (1997) in having an explicit model of how skills accumulate and in including career-choices over the full life-cycle: formal schooling and National Service, accumulation of work experience and retirement at old age.[1] This makes it possible to study a wider range of policies.[2] The model of individual behavior makes it possible to integrate out the effect of past choices on the distribution of unobservables, such that it is not necessary to make the restrictive assumption that all non transient state variables can be observed by the econometrician.

The three birth-cohorts of Norwegian males I use span the life-cycle, and following Lee (2005) I use the stochastic structure of the model to integrate out over the unobservables for time that pass before the cohorts under study reaches the window of observation (1986–2000). Observers have claimed that the incentives introduced by the public sector supplementary benefit schemes have the effect of locking people into the public sector when they reach middle age. This difference in retirement benefits between the two sectors provides exogenous variation in incentives that helps identify the model.

The retirement benefits system in Norway has two main parts, an universal scheme, the "National Insurance" (*Folketrygden*, somewhat resembling the US Social Security), and supplementary systems provided by contracting between employers and employees. The National Insurance is a strongly redistributive program, financed mostly by pay-as-you-go taxation. Even those who have never worked, and therefore not contributed to the pension fund, are entitled to a basic pension and a special supplement guaranteeing a yearly minimum income of about USD 11,000 with little taxation applied. Even above this, the accumulation of claims to future benefits is highly non-actuarial. Employers provide additional supplementary pension schemes. Only the private sector supplementary schemes are approximately actuarial. The public sector schemes are financed by all public employees at a rate of 2% of gross earnings, and the benefits offers a defined benefit level after consolidation with the National Insurance.

[1] Former studies of the Norwegian pension benefits schemes have focused either on forecasting effects of changes in the fine structure of early retirement schemes, and have not modeled the full life-cycle. See for instance Hernæs, Sollie, and Strøm (2000); Bratberg, Holmås, and Thøgersen (2001); Brinch, Hernaes, and Strøm (2001); Jia (2005).

[2] The present paper concentrates on reform of the retirement benefits but future work will look at other issues as well.

This consolidation implies that people with low incomes, little public-sector experience and part-time workers subsidize the pensions of high-wage full-time employees who stay on in the public sector throughout their career. The benefit level is calculated from experience and the final salary. Going to one public sector employer to another does not interfere with the public sector insurance schemes, one is seamlessly transferred to the scheme of the new public employer as if all past public-sector contributions were to the last employer's fund. This provided the rationale for standardization of public sector pension schemes (Norges offentlige utredninger 2000).

In Section 1 I present a model of sectoral choice and retirement in an environment where there is no aggregate uncertainty but where there are idiosyncratic shocks to earnings and human capital accumulation. The following sections present the data, explain how the model is estimated, the structural estimates and a policy experiments. An appendix provides additional details about the Norwegian retirement benefit system with some excessive detail, to make it possible to evaluate the relevance of the incentives embedded in the present model.

1 Modeling career choices

The full level of detail in the Norwegian system is impossible to embed in an empirical model of optimizing individuals. I work with the simplifying assumption of no aggregate uncertainty, and abstract from the consolidation of married couples pensions. Modeling the interactions of spousal choices at the level of detail needed to model responses to policy changes in retirement benefits is currently not a tractable problem in life-cycle models, in which households cannot be taken as stable units (although important progress has been made by van der Klaauw and Wolpin (2005)). I do model the incentives for single men only, and apply these rules to men irrespective of their marital status. Since men tend to be older than their wives, have more work experience and higher wages, this is probably a reasonable first approximation to the marginal pension benefits. I assume that individuals have perfect foresight about changes to the benefit schemes.

I also attempt to produce the simplest possible model that can retain some of the incentives in human capital and pension benefits accumulation. This means that all variables will take on a discrete number of values, and many of them will be binary. Whereas it has been shown that complicated models can be estimated when appropriate approximations are made (French 2005; van der Klaauw and Wolpin 2005), this typically depends on the limiting restrictions of all unobservable variables having no history-dependence, and a strongly restricted structure on the unobserved permanent heterogeneity. I formulate a simple model, but one that allows for both unobserved differences in skill-technology and incomplete observation of the relevant state of individuals. Choices are made to maximize the net present value of the future income stream (as in Keane and Wolpin (1997)). Hence there is no insurance motive to the retirement benefits in the model.

1.1 States and choices

Time is characterized by two indices, t index the age of the individual, $t \in \{0, \ldots, T\}$, and $d \in \mathcal{D}$ index demographic groups. The different demographic groups, generations of birth cohorts, are indexed because the environment is non-stationary: government policy and skill prices change over time. Together, d and t determine the calendar year. The state of an individual, denoted by θ, is

$$\theta_t \equiv (\alpha_{t-1}, \text{elig}_t, \boldsymbol{x}_t, \text{pp}_t^f, \boldsymbol{s}_t, u_{4t}, \boldsymbol{\epsilon}_{wt}, \text{edu}_t). \tag{1}$$

I use α to denote the choice taken, hence α_{t-1} is the choice taken in the previous period. The "elig" variables index eligibility for National Service and early retirement. Experience, $\boldsymbol{x} \equiv (x_1, x, x^{NI})$, consists of three variables that represents experience in the public sector, x_1, total experience, x, and experience within the National Insurance system x^{NI} (needed to model those who started working before National Insurance was introduced). The capital within the National Insurance system is kept track of in pp^f, the final pension point score. Sector specific skills are held in $\boldsymbol{s} \equiv (s_1, s_2)$. The productivity in home production is u_4, and $\boldsymbol{\epsilon}_w \equiv (\epsilon_1, \epsilon_2)$ is the realization from the conditional wage distribution. While not relevant for individual choice, $\text{edu}_t \in \{0, \ldots, 9\}$, counts past years of education over and above mandatory education (9 years).

I will return to the exact specification of the the states, for now it is sufficient to note that an action α_t taken from the finite set $\mathcal{A}(\theta_t, t)$ of possible actions induce a distribution over next period states $\theta_{t+1} \in \mathcal{S}(\alpha_t, \theta_t)$. The state to state transitions can

Table 1: Possible and mutually exclusive actions

α	action
1	Work full time in the public sector
2	Work full time in the private sector
3	Go to school
4	household production
5	retirement

be factored as

$$P(\theta_{t+1}|\alpha_t,\theta_t,t,d) = P_x(\boldsymbol{x}(\theta_{t+1})|\alpha_t,\boldsymbol{x}(\theta_t),t,d) \cdot P_{\mathrm{pp}^f}(\mathrm{pp}^f(\theta_{t+1})|\alpha_t, w_{\alpha t}(\theta_t),\mathrm{pp}^f(\theta_t),\boldsymbol{x}(\theta_t),t,d) \cdot P_s(\boldsymbol{s}(\theta_{t+1})|\alpha_t,\boldsymbol{s}(\theta_t)) \cdot P_{u4}(u_{4,t+1}|u_{4,t}) \cdot P(\epsilon_w). \quad (2)$$

The experience sub-process depends only on choices and time, the retirement capital depends on choices, calendar time and the rest of the state only through the wage earned. The skill process depends only on choices, the productivity in home production depends only on the current level and the wage shock is independent of history and the other state variables.

The action set depends on the state because there are restrictions on when retirement is allowed, and some further restrictions will be imposed for computational reasons. There is also the possibility of the government restriction choices by forced and mandatory drafting into the National Service (I will return to this in section 2.3). In the first period, individuals are 16 years old and exiting compulsory education. In the last period T, only retirement is allowed and retirement brings the career to a full stop. The possible actions are enumerated in Table 1.

Experience is sector-specific and acquired by working. In order to limit the size of the state space, experience is acquired in step intervals of 10 years. In order to smooth the effect of these jumps, experience is acquired stochastically: By working one has probability $1/10$ of a step increase in experience. Total experience is bounded at 40 years. The experience vector $\boldsymbol{x} = (x_1, x, x^{\mathrm{NI}}) \in \mathcal{X}$ keeps track of x_1, experience in the public sector and x, total experience, and x^{NI}, national insurance experience:

$$\mathcal{X} = \{(i,j,k) : i,j,k \in \{0,10,20,30,40\} \text{ and } i \leq j,\ k \leq j\}.$$

For everyone who entered the labor market after the introduction of the National Insurance, $x^{\mathrm{NI}} \equiv x$, and there are only 15 points on the grid in total. For the cohorts that made their early career choices without an existing National Insurance system in place, there are more possible points because total experience can be greater than the National Insurance experience. Limiting attention to those who entered ten years before 1967, I allow only the attainment of one unit (10 years) of experience earned before 1967, and this keeps the total number of experience points down to 29.

People leave compulsory schooling without any specific skills, but can acquire sector-specific skills by working or attending school. The specific skills take only two values (high and low), such that $s \in \{0, \overline{s_1}\} \times \{0, \overline{s_2}\}$. Production of specific skills is not deterministic, but attending school there is a probability p_3^s that an unskilled individual acquires skills. Schooling produces general skills, by which I mean that public and private sectors skills jump simultaneously. There is a probability p_+^s that an unskilled individual working in sector will acquire a unit of specific skills in that sector and a probability p_-^s that a skill unit in a sector will be lost by an individual who does not use it.

The National Insurance administration calculates pension benefits conditional on a much larger state space than can possibly be allowed for in a quantitative model. The National Insurance keeps track of an "average pension points" which is the average over the 20 years with highest acquired pension points. Pension points are calculated as a non-linear function of earnings, and the formula for calculating them has changed over time. Since 1992, one earns no pension points for earnings below $1G_t$. For earnings in between $1G_t$ and $8G_t$, one earns $(y/G_t - 1)$ pension points, and above $8G_t$ there is no marginal contribution of earnings to pension points. Since 1992, the value of G_t is indexed to wages, and $G_{2000} = 48400$, approximately 7500 USD. See the appendix for more details on how this has been calculated historically.

In order to model the average-of-twenty-best years exactly, one would need to keep track of an pension point history as a point in $\mathbb{R}_+^{20}$. This is not feasible. The problem of keeping track of continuous variables can easily be solved by an appropriate discretization. Keeping track of how the average over the twenty years of highest earnings evolve is a more difficult problem, and some approximation will have to be found. Researchers studying Social Security in the United States are faced with a similar problem.[3] French (2005) assumes that above average earnings always replace an

[3]The US Social Security Administration keeps track of 35 years of highest past earnings in the

average year after the history is full, while van der Klaauw and Wolpin (2005) assume that *any* current earnings always replace an average year. Since the current average always is higher than the minimum of past earnings (the earnings point actually replaced under Social Security regulation), it would seem that these approaches systematically underestimate the effect earnings late in the career can have on retirement benefits. Rust and Phelan (1997) takes a different approach by estimating a reduced form regression where Social Security Averaged Indexed Monthly Earnings (AIME) depends on current earnings, AIME lagged by one period and age.

Since the previous public sector wage can be predicted by the state-space, I do not track it separately. I follow the method of Rust and Phelan (1997), but without treating the National Insurance capital as a continuous variable. I discretize the accumulated pension average into four levels, and estimate reduced form logit-equations for whether a given earnings level give rise to a jump on the grid, such that the distribution of next period accumulated pension average is a continuous function of current earnings. I return to the details of this procedure in Section 2.5 after having described the available data.

1.2 Utility of choices

Based on their outcome (θ_t, α_t) individuals receive a per period utility of $U(\theta_t, \alpha_t)$:

$$U(\theta_t, 1) = (1 - \tau_{dt}(w_{1dt}(\theta_t)) - 0.02)w_{1dt}(\theta_t) + u_1, \tag{3a}$$

$$U(\theta_t, 2) = (1 - \tau_{dt}(w_{1dt}(\theta_t)))w_{2dt}(\theta_t), \tag{3b}$$

$$U(\theta_t, 3) = u_3 - 1(\alpha_{t-1} \neq 3)u_{3r}, \tag{3c}$$

$$U(\theta_t, 4) = (1 + \xi)^{t-d} u_{4t}, \tag{3d}$$

$$U(\theta_t, 5) = B(\theta, t, d). \tag{3e}$$

Wages are functions of the state, and τ_{dt} is the average tax function. In the public sector, workers also contribute 2% towards future retirement benefits. All agents are risk-neutral. If they work they receive a wage payment and potentially a non-pecuniary value u_1 of working in the public sector. The net present value of the pension benefits are $B(\theta, t, d)$, and this function allows for taxation, mortality and the details of the benefit systems (outlined in Appendix A). If they go to school there is a cost (in

"Average Indexed Monthly Earnings".

money-terms) of $-u_3$ and possibly an extra cost u_{3r} of returning to school from a working or a household production state.[4] Productivity in home production is u_{4t}, this also increases at the rate of technological progress.

Wages are functions of experience, current skills and a conditional wage distribution with discrete support,

$$w_{tj}(\theta) = (1+\xi)^{t-d} r_j \exp\left(s_j + \beta_{j1}x_j + \beta_{j2}x_j^2 + \beta_{j3}x_{i(\neq j)} + \epsilon_j\right).$$

The skill-price is determined outside of the model, but there is (potentially) secular growth in labor productivity. The stochastic element ϵ is distributed on a discrete support with a known covariance-matrix. Note that whereas education produces skills with some probability, education is not in itself a determinant of wages. There is also a technologically neutral growth ξ in labor productivity.

All attempts at modeling the value of retirement have to simplify the legal code and the state space. In this paper I model the total gross flow of benefits as the maximum of the potential National Insurance, b_{dt}^{NI} and the public sector benefits, b_{dt}^{gov}, and define

$$b_{dt}^{\text{gross}} = \max\{b_{dt}^{NI}(\theta_t), b_{dt}^{\text{gov}}(\theta_t)\}. \tag{4}$$

The *net* benefits are

$$b_{dt}(\theta_t) = 1(t < T)u_4(\theta_t) + 1(\text{elig}(\theta_t) = 5)\left[b_{dt}^{\text{tax}} + (1 - \tau_{dt}(b_{dt}^{\text{gross}}))b_{dt}^{\text{gross}}\right]. \tag{5}$$

Home production and retirement can be combined if $t < T$, and the tax system favors old-age pensioners in several ways that are partially off-setting. The net effect of this does not vary much with income (Rønningen and Fredriksen 2002) and can hence be modeled by a constant b_{dt}^{tax}. If the individual is eligible for retirement, he gets the tax-reduction and benefits net of taxes.

The yearly National Insurance component of retirement benefits for a single male is approximated by

$$b_{dt}^{NI} = G_t \max\{1.79, 0.42\text{pp}_t^f \cdot x_t^{NI}/40\}. \tag{6}$$

[4] The function $1(\psi)$ is an indicator function that takes the value 1 if ψ is true and 0 otherwise.

The public sector component is approximated by

$$b_{dt}^{\text{gov}} = \begin{cases} 0.67 \cdot \min\{x_{1t}/30, 1\} \cdot \min\{w_{rt}, 12G_t\} & \text{if retirement from public sector} \\ 0.67 \cdot \min\{x_{1t}/40, 1\} \cdot \min\{\tilde{w}_{1t}, 12G_t\}, & \text{else,} \end{cases} \tag{7}$$

where w_1 is the wage at the time of retirement, x_1 is the experience in the public sector and $\tilde{w}_1$ is the last recorded wage. Since $\tilde{w}_1$ is not kept track of separately in the state-space, I use the current state to predict what the last public sector wage would have been. I make the restriction on choices that no inter-sectoral career changes can be made after the age of 60 I therefore predict public sector wages of those not retiring from the public sector back to the this age netting out the effect of technological progress.

Everyone is eligible for retirement when they reach 67. Everyone working in the public sector are eligible for early retirement, "AFP" when they reach 62. In the private sector, only about 50% of the firms are covered by the programme.[5] Without modeling firm-to-firm transitions in the private sector, only a crude approximation to private sector eligibility can be made. At the age of 60, I randomly let 50% of the private sector employment be eligible for the AFP. The AFP programme itself I approximate by an earlier possible retirement age, but keep the rest of the regulations from the ordinary retirement. See Hernæs et al. (2000) for a more focused analysis of the AFP programme.

Having defined the flow from retirement, the net present value at a given moment in time t, discounted by a discount factor δ and the net survival probabilities, with m_j representing the mortality rate, can then be written

$$B(\theta_t, t, d) = \sum_{i=t}^{95} \delta^{(i-t)} b_{di}(\theta_t) \prod_{j=t}^{i} (1 - m_j). \tag{8}$$

The choice of terminal age at 95 is not important.

1.3 value of choices

The total value of an outcome is affected by the state one is at and the choice one makes, but for expositional clarity it is convenient to introduce $p^m(\theta_t, t, d)$ which is

[5] The exact number is disputed (Midtsundstad 2004).

the probability with which the government enrol an individual in the national service in period t. This could potentially depend, on lagged outcomes, such that individuals can control the draft to some extent.[6] If the National Service is forced on individuals, their possible next-period states are $\theta' \in \mathcal{M}(\theta)$. The transition to a new state is governed by the Markov transition probabilities $P(\theta'|\alpha, \theta, t, d)$. This depends on age, which determines eligibility for national service, and calendar time which reflects changes in the retirement benefit system – but also the effects of increasing wages on the probability of a transition on the final pension points score. For computational reasons, I only update the transition matrix every fifth year. I restrict the National Service to be equivalent to the possible home-production states, $\mathcal{M}(\theta) = \mathcal{S}(4, \theta)$. The expected value of a choice can now be defined as

$$v_t(\alpha, \theta) = U(\alpha, \theta) + \delta(1 - m_t)\Big[p^m(\theta, t+1, d) \sum_{\theta' \in \mathcal{M}(\theta)} \frac{P(\theta'|4, \theta, t, d)}{P(\mathcal{M}(\theta))} V_{t+1}(\theta') + (1 - p^m(\theta, t+1, d)) \sum_{\theta' \in \mathcal{S}(\alpha, \theta)} P(\theta'|\alpha, \theta, t, d) V_{t+1}(\theta')\Big]. \quad (9)$$

The state is fully observed before the choice is made, so the value function V_t is the maximum of the values of the possible actions,

$$V_t(\theta) = \max_{\alpha \in \mathcal{A}(\theta, t)} v_t(\alpha, \theta). \quad (10)$$

The current action determines the set $\mathcal{S}(\alpha, \theta)$ of reachable next-period states. If the current action was to retire or the individual reaches mandatory retirement age, $t = T$, this set is empty and the career ends.

We must allow for some smoothing of the choice probabilities, or there would be a degenerate choice distribution for every θ. This would a lead to a non-smooth objective function in the estimation procedure. Following Ferrall (2004), I therefore impose a ρ smoothing of choices, with $\rho \in [0, 1)$,

$$\tilde{v}_t(\alpha, \theta) = 1(\alpha \in \mathcal{A}(\theta)) \exp\left(\frac{\rho}{1-\rho}(v_t(\alpha, \theta) - V_t(\theta))\right), \quad (11)$$

[6]This would be similar to how Lee (2005) models the accumulation of children. I am, however, not letting choices affect the probability of being drafted since we will find in the next section that the average transition probabilities does not reflect the selection mechanism most people would expect.

and let individuals have exponentially smoothed choice probabilities

$$P_\alpha(\theta, t, d) = \frac{\tilde{v}_t(\alpha, \theta)}{\sum_{\alpha' \in \mathcal{A}(\theta,t)} \tilde{v}_t(\alpha', \theta)}. \tag{12}$$

The extreme case of $\rho = 0$ is the "perfect smoothing" case with all choices being given equal probability. As ρ approaches 1 the choices approach the degenerate case with $\arg\max_\alpha v(\alpha, \theta)$ being chosen with probability one. The value of ρ is estimated. In practice, and with regard to predicted choice probabilities, this method is not much different from following the tradition of adding extreme-value added terms to each choice.[7]

Having solved the backward induction for $P_\alpha(\theta, t, d)$ given $P(\theta'|\theta, \alpha, t, d)$, one can solve forward for the full distribution of states at any time, let $\mu(\theta, t, d)$ be the probability mass of generation d in state θ at age t,

$$\begin{aligned}\mu(\theta, t, g) = \sum_{\theta'} \sum_{\alpha} \Bigg\{ \Bigg[& p^m(\theta', t, d) \frac{1(\theta \in \mathcal{M}(\theta'))}{P(\mathcal{M}(\theta'))} P(\theta|4, \theta', t-1, d) \\ & + (1 - p^m(\theta', t, d)) P(\theta|\alpha, \theta', t, d) \Bigg] P_\alpha(\theta', t-1, d) \mu(\theta', t-1, d) \Bigg\}. \end{aligned} \tag{13}$$

With the assumption that individuals start after leaving middle school without skills and experience, this defines a full distribution of states that can be compared with data. I will return to the empirical application in Section 3.

1.4 Computational restrictions on choices

For computational reasons, I restrict the possible choice of attending school to the first 25 years after leaving middle-school. In order to fit the pension system, I will not let anyone retire (in the sense of claiming the National Insurance and the supplementary benefits) before they are 62 years old. There are few people changing their career from the private sector back to the public sector in the last years of their career, although it might seem as if this would be favorable for those with previous experience in the public sector. Undoubtedly there are problems with switching career the year before retirement that will be difficult to capture in this model, so I will not allow anyone to

[7] Although it can be argued that the method I employ have advantages in situations where the size of the choice set changes over the life-cycle.

switch sectors after the age of 60.

2 Data and ancillary models

Statistics Norway has collected data from various government administrative records, from which the labor economics group at Norwegian School of Economics and Business Administration has constructed an integrated database (Møen, Salvanes, and Sørensen 2003). The entire population of residents in Norway aged 16–74 can be followed with yearly observations in the years 1986–2000. In addition, a full panel of yearly labor earnings since the introduction of the National Insurance scheme in 1967. labor earnings is measured using the National Insurance definition, which includes benefits while sick but excludes unemployment benefits.

For the purpose of this paper, since there is no immigration or emigration in the model, I restrict attention to those born in Norway (or resident at the age of 16 when the choice-model begins). I use three cohorts for estimation, the 1941, 1955 and the 1970 cohort. The 1970 cohort is a natural choice since these people leave compulsory schooling in the first year from which we have data. The 1955 cohort is chosen to have a middle-aged group, and the 1941 cohort since this generation can accumulate exactly 10 years of experience before the introduction of the National Insurance scheme. The model assumes an educational environment that is constant. This is an abstraction, an educational reform in the 1960's meant that compulsory education for the 1941 cohort in fact ended earlier, and that not everyone stayed on in school until they were 16 (Raaum, Salvanes, and Sørensen 2003, 2006). Considering the time in between schooling and observation, I assume this to have only minor impact on the estimates. The population in the oldest cohort is, at the time of our observations, 20498 persons, while the 1955 and 1970 cohorts consists of 30670 and 32422 respectively.

The attachment of workers to firms is observed at a point in time, from 1986 to 1994 this is May 31st, from 1995 and onward this observation is made in November. For each of these attachments we get to observe the start-date of employment and, if applicable, the termination date. So for choices, the calendar year is not the only possible choice of yearly time-period. For any study that includes education it is more convenient to measure the year according to the academic year. I have defined the choice period of a year as beginning July 1.

Any model of career choices will come up against the problem that in the data we

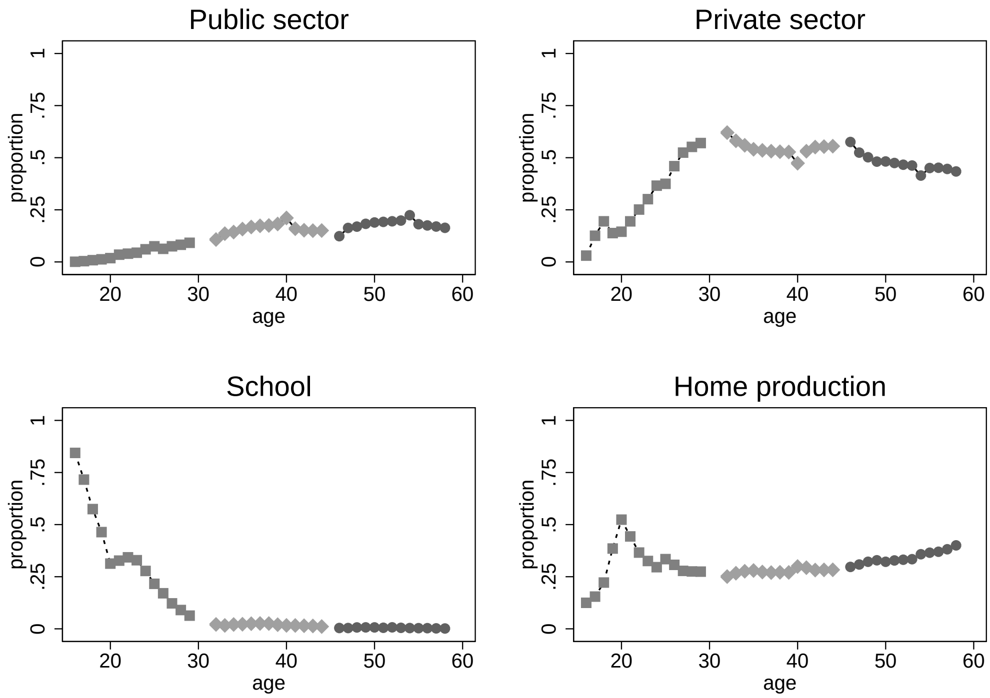

Figure 1: Choice distributions. Classifications are defined in the text. Three cohorts, the 1941, then 1955 and the 1970 cohorts are shown, with choices recorded in the window 1986–2000.

see people combining several different choices within the defined time period. People may take on a part-time job while still in school, some people may quit school in the middle of the year to start work, some drop out without any evident cause, and others take on a succession of high-frequency jobs. Any rule that attempts to force data into year-long decision periods and exclusionary choices will be arbitrary to some extent. I follow Keane and Wolpin (1997) and Lee (2005) in hierarchically classifying the individual into categories, starting with education, following with the two working choices and at last let the home production category be defined residually.

2.1 Defining choices from data

Education For education, we get to observe the highest education level achieved by individuals, and the graduation date of this level. From 1986 and onward, we also get to observe the flow of education from administrative records; whether an individual was enrolled in part-time or full-time education as of November 1st (Vassenden 1995;

Statistisk sentralbyrå 2001). I use the "enrolled in full-time education" variable to classify the individual as enrolled in school. Possibly this could be refined by checking to see that the individual actually graduates, but this is complicated by the fact that some educational categories span several years, and a not insignificant fraction of the students seem to move laterally in the grade system in a way that is not captured well by the stock variable of "highest level achieved". Education is measured as the number of years above mandatory education, and is capped at 9 years.

Work As mentioned, we observe if the individual is attached to an establishment once a year. We also get to observe ownership-structure, industry, a categorical measure of weekly hours, and start and stop dates.[8] Using this history of work-spells, I calculate the hours worked in the year as defined from July 1 to June 30 the following year using the Labor Force Survey as guide to how many hours per week to impute within the categorical groups, and I use only the job spells of 20 or more hours per week.[9] Since we only have yearly snapshots of workplace attachment, this method may not work well when there are a lot of high-frequency job changes. I have therefore chosen a cut-off for whether an individual is classified as working as low as 1000 hours per year. In order to qualify as "working" the individual also needs a minimum level of earnings. Since the working / academic year spans two calendar year, I demand that in at least one of those two calendar years the individual earned $1G$ (with G defined as by the National Insurance Administration). If they in fact worked that year, this should easily be fulfilled, but some workplaces may wrongfully keep people in their records after short summer jobs, and if individual do not attach themselves to other workplaces or go to school, this could cause errors in the classification.

Sectoral categorization For those categorized as working following the above definition, I make the categorization into public and private sector based on the job-spell that contributed most hours. We get to observe a unique identifier of the establishment, and this identifier can be used to link to the central register of establishments and corporations (Olsen 1993), in which there is a variable indicating whether the establishment is owned by central or local government. When this link works, this is the sectoral categorization. I have only had access to the data from the central register of establishments and corporations for the period 1986–1995. For the remaining years

[8]Weekly hours is categorized in 3 levels, (1) less than 20 hours per week, (2) between 20 and 30 hours and (3) more than 30 hours per week.

[9]I have calculated averages of total hours for men over the period 88-95, these averages are 12.59, 25.65 and 36.84 hours per week.

1996–2000, I have used the public/private categorization as available in 1995 for all those establishments that existed then. For new establishments, I have used the 1995 data to create a mapping from industry code (ISIC Rev2 at the 4-digit level) to sector (using the within industry mode of the "public" indicator, weighting with the number of employees) and used this industry-based categorization on the new establishments 1996 to 2000. For the years 1999 and 2000 a new mapping from the ISIC to the NACE industry code is needed, I use the within NACE mode of the ISIC codes, again weighted by the number of employees and at the 4-digit level.

Residual home production and National Service Those not categorized under the above rules are classified as at home, an observational group that also includes the (observationally equivalent) individuals in the National Service.

The final distributions of choices as people get older for the 3 cohorts used in estimation is shown in Figure 1. There is slow but steady growth in public sector employment as people get older, whereas private sector employment peaks around the age of 30 and then slowly tapers off as people get older. There are few people in education after the age of 30, and the slow decline in private sector employment is (nearly) matched by a slow growth of people in home production. The early spike in home production reflects that this category also contain those in National Service.

In Figure 2 we see the average conditional choice probabilities by age. There is, of course, a lot of noise in the transition probabilities from the choices few make. We can, however, see that there is a quite a lot of persistence in choices and that there is more mobility among the young than among the old.

2.2 Wages and earnings

Because I define the decision year as starting at the beginning of the third quarter, calculating wages is difficult, since the data I have to work with is yearly earnings, reported for the calendar year. However, since I do not attempt to match individual level data, I will not have to match individual decisions to individual choices. It is therefore possible to take some liberties when calculating wages. I have in fact chosen to offset the measurement of wages by half a year, such that the calendar years wages corresponds to the decision period that starts in the third quarter of that year. Since the square of wages is used when matching the model to the data, measurement error in wages/earnings will have consequences for identification. I have therefore chosen to only use the earnings of full-time workers who can be matched to a workplace for at

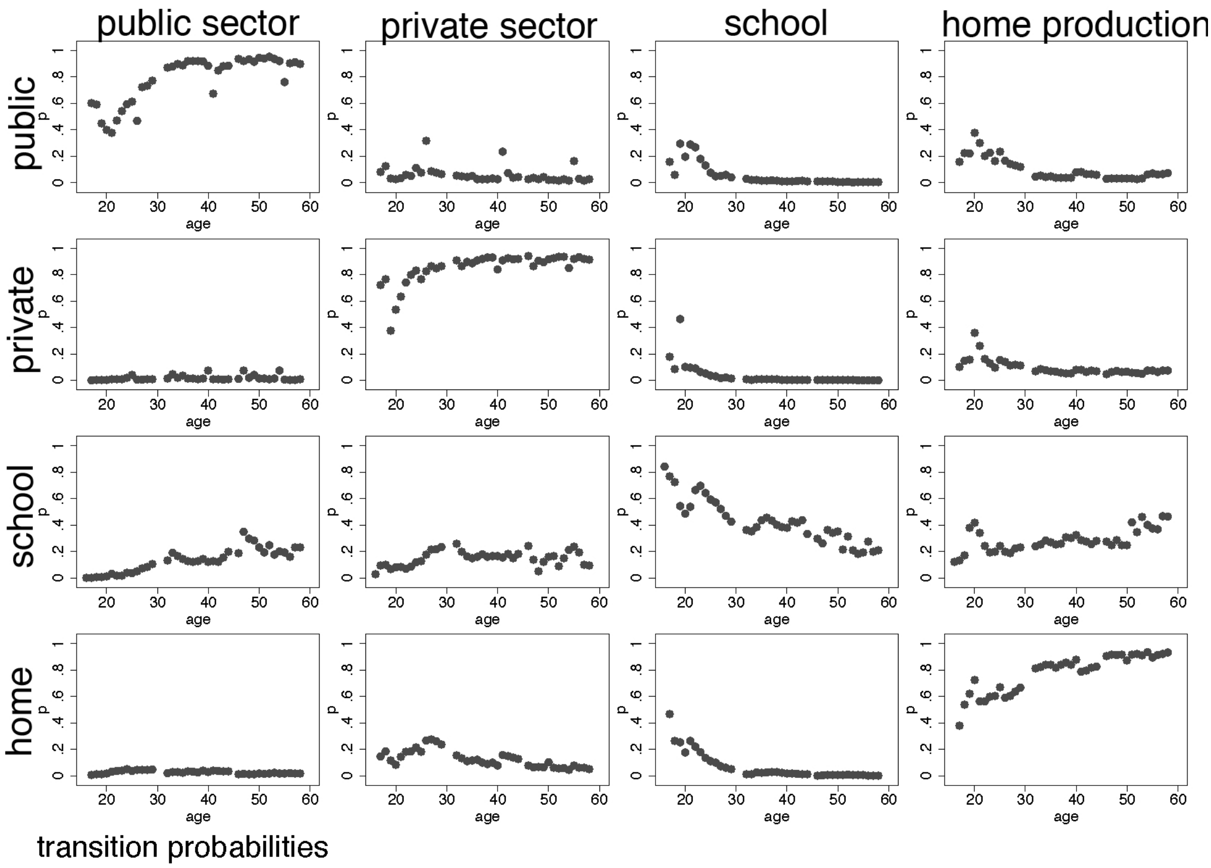

Figure 2: Choices conditional on previous choices.

Note: A $p_{ij}(t)$ entry is the fraction of people who choose i in the previous period who choose j at age t. Calculated on the 1941, 1955 and 1970 cohort in the observational window 1986-2000.

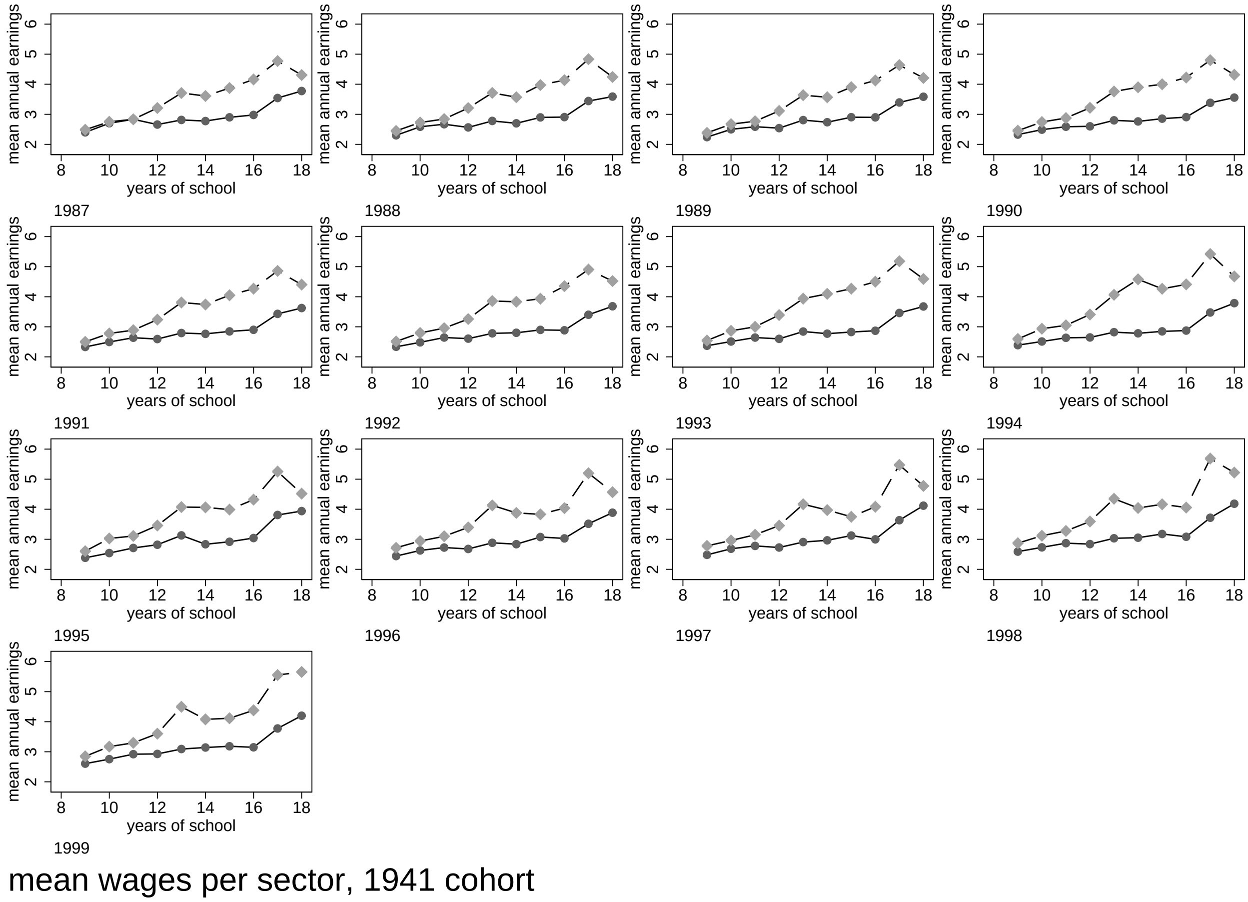

Figure 3: Annual earnings in units of 100 000 1998 kroner. The solid line is the public sector.

least 320 days of the relevant calendar year. I have then taken the means of the labor earnings at 1998 consumer prices as the measure of that period's full-time return to labor. All money measures throughout the paper are calculated in terms of 1998 NOK.

The figures 3–5 shows the earnings/education profiles in both the public and the private sector for the years the model is estimated on. Among the older cohorts there is a substantial gap between private and public sector wages, and a clear positive earnings profile with respect to education. In the younger 1970 cohort, we see neither as clearly. It seems as if experience can partially offset education in the first years of the career, and that the public sector has a flatter experience profile than the private sector.

Since the private sector supplementary pension benefits are transferrable, and the variation in types of private supplementary benefits is too large to take into account, I impute pension premiums to current wages in the private sector. I use the coverage

Figure 4: Annual earnings in units of 100 000 1998 kroner. The solid line is the public sector.

Figure 5: Annual earnings in units of 100 000 1998 kroner. The solid line is the public sector.

rates from Pedersen (2000) and the rates I have shown in Figure 10.

2.3 National Service

In Norway most young males serve a year of military services. This is different from the military career-choice modeled by Keane and Wolpin (1997) in that it results from a mandatory military service for all young males. In principle, everyone is called to serve for a year at age 19, but in fact there are a number of exceptions. First, not all are in fact called to serve, some are turned away because of weak health or because the full cohort is not needed. Second, it is possible to postpone military service if it would interfere with studies, and third; one can refuse military service as a conscientious objector, in which case there is mandatory civilian service.

The military service is also only partly documented in the data available to us (being only observable from 1993 and onward). This makes difficult to model the early career choices of young men, since military service will potentially corrupt the first 5 or 6 years of young men's career histories.

These fact that military service is only partially controlled by individuals also make it difficult to model the interruption to young men's careers as either completely exogenous or freely chosen by individuals. An intermediate solution, in which eligible young men who find themselves called to serve with probability $p^m_{tg}(\alpha_{t-1})$ is possible, by letting the National Service arrive at a rate that depends on the choice in the previous period it would be possible to avoid military service by staying in school hoping that one will not be called up. From Table 2 it is evident that it attending school is not a safeguard against being drafted. In fact, draft-probabilities are higher among those who attended school the previous period than among those who worked or stayed home, a somewhat surprising result. A full understanding of this will probably require a more detailed modeling of the selection into National Service, a task the model in this paper is not fit for. As a first approximation, I let the probability of being drafted depend only on age and cohort.

Table 2: Probability of being drafted for National Service, $p^m_{1975,t}$

year	Choice in previous period			
	work	school	home	total
1993	0.0520	0.0204	0.0564	0.0239
1994	0.1751	0.2476	0.1637	0.2351
1995	0.2770	0.2992	0.2899	0.2978
1996	0.1633	0.2347	0.1602	0.1993
1997	0.0609	0.1520	0.0628	0.0960
1998	0.0440	0.1234	0.0468	0.0560
1999	0.0104	0.0356	0.0154	0.0016

Note: Calculated on the 1975 cohort. Draft is defined as the first entry into the military or the social service that lasts more than 100 days.

Since individual level information on National Service is only available from 1993 and onward, most of those serving will wrongly be classified as engaged in home production. In order to extend the information available, I have collected aggregate historical information from the National Draft Board about how many of a given cohort they have in fact trained. I then propose a multiplicative model of draft intensity, where

$$p^m_{dt} = \zeta_d \cdot p^m_{1975,t}. \tag{14}$$

The empirical hazard rate for the 1975 cohort is available in Table 2, and I then scale these hazard rates by a cohort-specific intensity ζ_d that is calibrated so that the process fits the aggregate proportion of a cohort that has been trained.[10]

The draft into the National Service is mostly in the beginning of every quarter. The third quarter has the largest numbers of draftees, so that for most, conscription is more or less synchronized with the academic year. Since most draftees spend about 9 months in the service, I want to classify them as drafted the decision year they spend to the largest extent in the National Service, so I categorize those who enter between April 1 this year and March 31 the following year to be in the National Service in the decision year that starts July 1 that year. In order to define the entry date, I look at the first spell that lasts more than 100 days.[11]

[10]Numbers on cohort specific training data have been provided by the National Draft Board ("Verneplíktsverket").

[11]After first joining smaller sub-spells with less than a month in between them, as these are mostly

2.4 Taxes

Since individuals aim to maximize the present value of earnings and utility flows net of taxes, I need the tax-functions τ_{dt} for the years 1956–2040. I have constructed the income tax-functions for the case of an individually assessed individual with no wealth or capital income for the years 1956–2000. My income-tax measure includes local and central government taxation, temporary income taxes and reliefs and National Insurance premiums.[12] These functions are so complicated and non-smooth that they cannot be used directly in estimation. Instead I have fitted fourth-degree polynomials to the average tax-rate, using the distribution of yearly earnings of those 40 years old as weights. Since I do not have annual earnings for the 1956–1966 years, I have scaled the 1967 earnings back using the index of manufacturing wages (table 10.3, Statistisk sentralbyrå 1995). This sequence of polynomials fit within sample fairly well. Out of sample, particularly for high earnings, these polynomials do not fit well. I therefore use the distribution of earnings to calculate the 90th percentile, above which I restrict the average tax function to be flat.

Since the choices made by individuals in the 1986–2000 period depend on their projections about future taxation, this projections must be explicit. I assume that individuals project the 2000 tax function into the future. In an attempt to construct the future taxes that keeps the current rate of redistribution more or less constant, I scale the parameters from this polynomial while estimating, using the estimated rate of labor-augmenting technical progress to stretch the tax function.

Pension benefits are treated favorably by the tax system: a lower rate is used to calculate National Insurance premiums, there is a special allowance for old age and there are rules limiting the total taxes of pensioners. These rules partially offset each other, and Rønningen and Fredriksen (2002) show that the total effect lies in a fairly narrow band, it varies from 9900 to 15700 (both in NOK98) in 1995 (Table 4 Rønningen and Fredriksen 2002). I use the midpoint of this, NOK 12800, as an estimate of the net tax benefit of retirement in 1995, and adjust this value at the same rate as the basic pension unit – at the rate of technological progress estimated.

transfers from one army base to another.

[12] I have drawn on Statistisk sentralbyrå (1958, 1968, 1975, 1988, 1994) and current numbers from `http://www.odin.dep.no/`.

2.5 Accumulated average of past pension points

I use the actual distribution of final pension points of those aged 60-70 years old in 2000 to calculate a 4-point grid on the final pension points, at the 12.5%, 37.5%, 62.5% and 88.5% percentiles, (3.03, 4.44, 5.50, 7.06). See Figure 6 for the full distribution of the final pension point score. Note that the 37.5th percentile is just enough to reach above the minimum pension for a single man (for which $4.26G$ in accumulated average is needed). It is then straightforward to classify individuals as being in the corresponding intervals defined by the 0-25%, 25%-50%, 50%-75% or 75%-100% percentiles (with cut-offs (3.90, 4.96, 6.18). For this truncation of the average accumulated pension points, I estimate reduced form logit-equations of the probability of jumping up or down on this grid, with the current pension points (a function of earnings, see discussion in Section A.1) affecting the probability of switching up or down. This way, the effect of current earnings on accumulated average pension points is smooth and continuous.

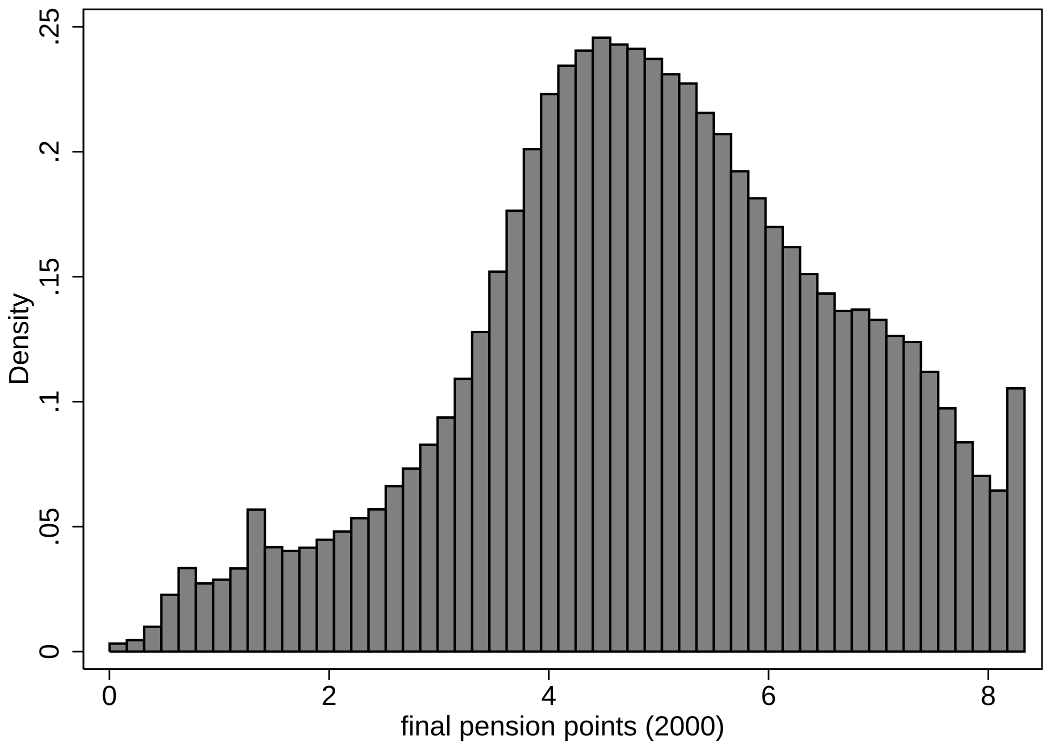

Figure 6: Final pension point score of potential retirees (60–70 years old in 2000)

Table 3: Average empirical transition rates on the pension capital grid grouped by experience. Calculated on the 1940, 1955 and 1970 cohorts.

	pp^f_{t+1}								
	0-9 yrs of experience				10-19 yrs of experience				
pp^f_t	1	2	3	4		1	2	3	4
1	0.972	0.027	0.000	0.000	1	0.951	0.049	0.000	0.000
2	0.061	0.840	0.092	0.000	2	0.013	0.930	0.056	0.000
3	0.005	0.070	0.851	0.075	3	0.000	0.015	0.952	0.034
4	0.002	0.006	0.072	0.920	4	0.000	0.000	0.017	0.984
	20-29 yrs of experience				30+ yrs of experience				
	1	2	3	4		1	2	3	4
1	0.912	0.084	0.000	0.000	1	0.968	0.032	0.000	0.000
2	0.000	0.916	0.084	0.000	2	0.000	0.979	0.021	0.000
3	0.000	0.000	0.945	0.055	3	0.000	0.000	0.992	0.008
4	0.000	0.000	0.000	1.000	4	0.000	0.000	0.000	1.000

Whereas one would not need more than a two-point discretization of the accumulated pension points to capture the main incentives, a finer grid is needed in order to robustly estimate the reduced form process on the grid. Separate regressions are made for each experience category, with experience calculated as number of years with earnings of at least one G (this corresponds to National Insurance practise). The average transition rates by experience group is shown in Table 3. We see that from year to year, there is not excessive volatility on this grid. In Table 4 are the estimated logit-equations where the probability of a jump up (or down) on the grid is estimated as a function of current earnings (in pension points), with different parameters by experience group. These estimates are then applied as parameters in the structural model.

3 Estimation

In order to facilitate estimation, I introduce some further notation. Remembering that $\boldsymbol{D}$ is an index set of observable (demographical) groups, one such group could be all males born in cohort 1970. Within each of these observable groups, let $k \in \{1, \dots, K\}$ index a set of unobservable types in the sense of Heckman and Singer (1984). Each of these K are associated with a weight $\lambda_{dk} \in [0, 1]$ (such that $\sum_k \lambda_{dk} = 1$) and a type-

Table 4: Reduced form (Logit) estimates of final pension points process on the 4-point grid. Standard errors in parentheses. Estimated on data from the 1940, 1955 and 1975 cohorts.

	0-9 years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up	down	up	down	up	down
pension points	1.62	−1.31	1.97	−1.28	1.94	−1.21
	(0.01)	(0.02)	(0.02)	(0.02)	(0.04)	(0.03)
constant	−11.98	3.70	−15.03	4.93	−16.58	5.65
	(0.06)	(0.07)	(0.15)	(0.11)	(0.30)	(0.19)
	10-19 years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up	down	up	down	up	down
pension points	1.31	−0.05	1.54	−0.75	2.00	−0.63
	(0.01)	(0.01)	(0.02)	(0.01)	(0.04)	(0.02)
constant	−10.51	−0.11	−13.26	0.19	−18.29	0.06
	(0.05)	(0.04)	(0.12)	(0.06)	(0.31)	(0.09)
	20-29 years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up		up		up	
pension points	1.21		1.23		1.37	
	(0.01)		(0.01)		(0.02)	
constant	−8.25		−9.43		−11.72	
	(0.07)		(0.08)		(0.13)	
	30+ years of experience					
Pension capital $(t-1)$	1	2		3		4
jump up/down	up		up		up	
pension points	2.45		2.76		3.02	
	(0.13)		(0.12)		(0.14)	
constant	−14.50		−18.03		−22.43	
	(0.68)		(0.63)		(0.90)	

specific parameter-vector γ_k, the structural parameters. Now the data will be composed of DK types each solving their own problem as defined in the model described in Section 1, and from now on I will on occasion subscript the terms defined earlier by their dk group.

3.1 Measurements and moments

The distribution μ in equation (13) is defined as the distribution of individuals over the state space $\boldsymbol{\Theta}$. This distribution is not directly useful for estimation since there are several unobservable components of this space. Let us therefore determine a measurement function $y : \boldsymbol{\Theta} \times \mathcal{A} \rightarrow \boldsymbol{Y}$, where $\boldsymbol{Y}$ is a space of measurements: yearly earnings, schooling, pension capital and choices in the previous period. Note that there are also endogenous state-variable that are not measured, such as sector-specific experience (for the older generations) and sector-specific skills, and choices can also be obscured by forced National Service.

The distribution over states induced by the optimal choices defines a distribution function $F(\mu) : \boldsymbol{Y} \rightarrow [0, 1]$ on the measurements. With this distribution F it is easy to define a series of observable moments predicted by the model, this distribution $F_d(\boldsymbol{\gamma}, \boldsymbol{\lambda}_g)$ will depend on all the type-specific parameters ($\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_K)$) and the population shares of the types in the d group ($\boldsymbol{\lambda}_d = (\lambda_{d1}, \ldots, \lambda_{dK})$). It is now straight-forward to define a moment-generating function $\Delta : \boldsymbol{\Gamma} \rightarrow \mathbb{R}^Q$, with Q being the number of moments.

This approach is similar to the algorithm of Keane and Wolpin (2001). Keane and Wolpin also model the distribution of measurements from the unconditional distribution over the full state-space. The approach taken in this paper is different on three counts. (1) It does not depend on an approximation to the solution to the individual choice problem. (2) It does not depend on a simulation to find the unconditional distribution over the states space. (3) I propose a GMM-approach instead of the maximum likelihood approach taken by Keane and Wolpin (2001). These deviations make it possible to construct a GMM-objective that is smooth in the parameters, making it feasible to use gradient-based methods for minimizing the GMM objective.

The empirical moments I match are outlined in Table 5. This list is a variation of those used by Lee (2005), but with more conditional choice probabilities than Lee was able to calculate from his data. The reason for choosing to fit the mean and the mean squared wages instead of the mean and the variance such as Lee does, is that only one pass through the states is necessary. This makes an difference for computational efficiency since the problem for each of the K types can be calculated separately by parallel processors, and fewer numbers need to be passed along. In order to calculate the variance of wages one would need to pass the full distribution across states for all types. This reduction in the information that has to be passed between processes reduces the time needed for communication on inexpensive clusters with distributed and limited memory, and makes it possible to apply the efficient methods for finite mixture models developed by Ferrall (2005).

The model is estimated with GMM, with the objective function

$$\sum_{d \in \mathcal{D}} \left[\left((\widehat{\boldsymbol{\Delta}}_d - \Psi_{k=1}^{K} (\boldsymbol{\Delta}_d(\boldsymbol{\gamma}_k), \lambda_{dk})) \right)' \boldsymbol{\Omega}_d \left((\widehat{\boldsymbol{\Delta}}_d - \Psi_{k=1}^{K} (\boldsymbol{\Delta}_d(\boldsymbol{\gamma}_k), \lambda_{dk})) \right) \right]. \quad (15)$$

The Ψ operator aggregates the moments from the K unobserved types using the λ_{dk} weights and ancillary information about conditional masses. The standard formula for the variance of the estimates, is now

$$\text{var}(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}_0) = (\boldsymbol{J}'\boldsymbol{\Omega}\boldsymbol{J})^{-1}\boldsymbol{J}'\boldsymbol{\Omega}\left(\text{var}(\widehat{\boldsymbol{\Delta}})\right)\boldsymbol{\Omega}\boldsymbol{J}(\boldsymbol{J}'\boldsymbol{\Omega}\boldsymbol{J})^{-1}, \quad (16)$$

where $\boldsymbol{J}$ is the Jacobian of the moment vector with respect to the parameter vector.

Since I use a large number of moments, I do not attempt to use an optimal weighting matrix. For computational tractability I use a diagonal weighting matrix. The weights are shown in Table 5. The estimation, like that of Lee (2005) and Lee and Wolpin (2006), is not fully standard in that conditional moments are matched (such as sector and education-level specific wages).[13] This means that the full covariance matrix of $\boldsymbol{\Delta}$ is not defined. As a practical approximation I use only a diagonal estimate of $\text{var}(\widehat{\boldsymbol{\Delta}})$, calculated with the size of the relevant sub-populations.

3.2 Estimated parameters

Estimation was performed on the five-node cluster "paulus" at Norwegian School of Economics and Business Administration using the FmOpt library (Ferrall 2005).[14] Tables 6, 7 and 8 contain the estimates of the structural parameters where only base

[13]There is a brief discussion about this in footnote 21 of Lee (2005) related to Lee's weighting strategy, but no information is provided about the calculation of standard errors.

[14]Running on 9 processors, estimation took about a week, with a sequence of simplex and gradient-based methods.

Table 5: Empirical moments

aggregate moment	#	units	weight
mean earnings by sector, age/cohort, education and experience	$2 \times (2 \times 13 + 14) \times 10 \times 5$	10^5NOK98	0.5
(mean earnings)^2 by sector, age/cohort, education and experience	$2 \times (2 \times 13 + 14) \times 10 \times 5$	10^{10} NOK^{2}98	0.25
choice distribution by age/cohort and education	$4 \times (2 \times 13 + 14) \times 10$		5
choice transitions by age/cohort	$4 \times 4 \times (2 \times 13 + 14)$		2
employment transitions by age/cohort, experience and pension point score	$2 \times 2 \times (2 \times 13 + 14) \times 5 \times 4$		5
total number of moments before restrictions	13440		

Note: There are some further restrictions: no empirical moments calculated with experience or education larger than age. Individuals only have 25 years to go to school, the 10 last years before legal retirement age sector change is restricted to zero. No-one can transfer from illegal states. In all, 10508 moments are given positive weight in estimation.

levels and the sector specific skills are allowed to differ by type. The discount factor is fixed at 0.98, the same level as in Rust and Phelan (1997). This value is higher than that fixed by van der Klaauw and Wolpin (2005) who fix it at 0.95, but lower than the rates estimated by French (2005), who find 0.985–1.04. State-dependent transition probabilities for the value of home production turned out to be hard to identify, I have restricted the value of home production to be constant.

The two types are distinctly different. Type 1 has a fairly flat earnings-profile with small returns to experience and small levels of sector specific skills. Type two starts with a much lower base level of skills, but has the possibility of gaining large amounts of sector specific skills – these almost triples the skill-level. These large difference between the types, together with the fact that the type-shares for the 1941 cohort is very different from those of the 1955 and 1970 cohort may indicate that there is in fact room for more heterogeneity in the model, and that I should attempt to estimate model with three types.

The probability of gaining the sector specific skills are slightly lower in school (at 0.059, an expected length to gain skills of 17 years) than at work (at 0.066, or an expected length to gain skills of 15 years). But this difference is not large, and those who gain skills at school accumulate "general skills", in the sense that they are free to choose what sector to apply them to, whereas the skills gained at work are sector specific. There is quite a high probability of losing skills that are not used (at 0.42, or an expected length to loss of skills of 2.4 years). This creates a significant skill-based lock-in for type 2 individuals who have reached the high-skill level.

Most of the parameters that are not close to being so small as to have no economic impact on choices are estimated at a moderately comfortable level of precision. We should not be surprised that parameters that estimated to have little or no effect on choices or observed earnings (such as the correlation between the shock to public and private sector wages) are imprecisely estimated.

4 Policy experiment

The "Johnsen-commission", with a mandate to examine the National Insurance scheme has suggested several changes to the current scheme, to be implemented from 2010, starting with the 1951 cohort. Many of the suggested changes are much to fine-grained to be captured well by the stylized model in this paper, but some of the larger feature

Table 6: Structural estimates, common parameters

parameter	value
smoothing factor, ρ	8.24e-6
	(0.67e-6)
discount factor, δ	0.98
	(fixed)
technological progress, ξ	13.91e-3
	(1.14e-3)
utility of school, u_3	160e3
	(15e3)
cost of returning to school, u_{3r}	115e3
	(30e3)
prob. of losing non-used skill	0.416
	(0.068)
prob. of gaining skill by work	0.066
	(0.017)
prob. of gaining skill at school	0.059
	(0.011)
base productivity in home production	148e3
	(9e3)

Table 7: Structural estimates, type shares

Demographic group	Share of type 1
1941 cohort	0.698
	(0.046)
1955 cohort	0.341
	(0.047)
1970 cohort	0.279
	(0.053)

Table 8: Structural estimates, public and private sector skill equation

	public sector		private sector	
parameter	type 1	type 2	type 1	type 2
constant	125e3	83.7e3	115e3	75.3e3
	(9.3e3)	(8.1e3)	(13.2e3)	(12.9e3)
sector-specific skill	0.032	0.958	0.184	1.088
	(0.685)	(1.267)	(0.089)	(0.176)
own experience	1.73e-3		16.0e-3	
	(2.40e-3)		(3.24e-3)	
own experience squared	-0.10e-3		-0.40e-3	
	(0.04e-3)		(0.11e-3)	
experience in other sector	1.72e-3		0.10e-3	
	(1.96e-3)		(1.7e-3)	
σ_ϵ	0.269		0.019	
	(0.046)		(0.820)	
$\text{corr}(\epsilon_1, \epsilon_2)$	0.860			
	(31.9)			

can be captured here: The commission proposes to abolish the 20-best-years averaging of the accumulated pension points, and a replacement by a flat function of past income. It also proposes to replace the early retirement scheme (AFP) with a flexible choice of retirement age, where individuals are less favorably compensated when they retire early. One would think it would be of value to know to what extent this would affect retirement decisions. The analysis in Norges offentlige utredninger (2004) is based on analysis using "MOSART" (Fredriksen 1998). It is based on the assumption that future choices under the new system is a simple average of current conditional choices and the choice from before the introduction of AFP.[15] As of today, the agreement between the political parties (Arnstad, Nilsen, Steensnes, Stoltenberg, and Vihovde 2005) does not indicate an abolishment of AFP (although adjustments may be made).

The details of the agreement among the political parties is hard to interpret in terms of the primitives of my model, e.g. it includes a lower rate of indexation, in which the average of wage growth and inflation is used for adjusting the value of benefits, and I have not attempted to model a change in the life expectancy as forecasted by the public commissions. As a first approximation, I model only a simple "core" reform. I

[15]Personal communication with Dennis Fredriksen, Statistics Norway.

model the removal of the "twenty-best-years" rule, and I adjust the National Insurance down slightly. I re-estimate $P_{\text{pp}^f}(\text{pp}^f(\theta_{t+1})|\alpha_t, w_{\alpha_t}(\theta_t), \text{pp}^f(\theta_t), t, d)$, the transition function for the final pension point score. I use the same cut-offs as earlier, but with a flat average of past pension points instead of the "twenty-best-years" rule (allowing for more "down" movements on the fixed pension point grid). I also reduce the National Insurance benefits by a modest 5%. In Figure 7 one can see the effect of this on the out-of sample predictions for the 1970 birth cohort.

First, we can note that the overall employment levels seem rather low. However, we can note that the model does not allow for part-time workers (who are classified as primarily doing home production). The jaggedness of the labor supply profile is an artifact of the Markov transition matrix only being updated every fifth year. The overall pattern under the baseline is in line with the historical experience: A more or less flat participation in the public sector, with a decrease in the private sector over time. This decrease in the private sector is reasonable since there are weaker pension capital benefits in the private sector. There is, however, a peculiar influx into the public sector at the end of the life cycle: These are people who are transferring to the public sector to be guaranteed AFP – early retirement benefits in the last year such a transfer is possible (by a restriction in the model). Since I have no firms in the model, I let private sector workers face a gamble on whether they could take AFP retirement, and it seems as if the value of AFP is to great for people to take that risk. This is not in itself alarming, since a number of other papers have shown that there are extremely strong incentives for people to take up the AFP if possible, and people who put any value on leisure will be predicted to retire early unless there are outside restrictions on the possibility of gaining a AFP-qualifying job.

The effect of the policy experiment is to reduce the labor supply of the middle aged group. Because of the relatively high rate of technological progress estimated, the retirement benefits seem to have early effects on labor supply. This is in line with what Rust (2001) has found. In my model it might be further suppressed by the fact that people realize that the most likely retirement option is with AFP from the public sector, and hence weights life-cycle contributions less. The policy experiment has a further effect at the end of the life cycle. There is a marked increase in labor supply in the private sector among people in their mid-sixties. These people are the ones who did not win out in the AFP gamble – either they stayed on in the private sector, or they did not get the high public-sector wage draw they had hoped for, and which would have made them retire early with a high level of benefits.

5 Concluding remarks

I have estimated a basic two-sector life-cycle labor supply model with only two initial types of workers. However, as time goes on and individuals make different choices, a large amount of ex-post heterogeneity is created. The estimates indicate that there is a type of worker for which comparative advantage is generated by their accumulated life-history, and for which the endogenous costs of switching sectors can be large. The model extends the human capital literature that takes education as a deterministic determinant of wages into a more realistic environment where education is one of the factors (together with work experience) that goes into producing skills that are valued by workplaces – without guaranteeing a monetary payoff of education.

The policy experiment emphasizes the fact that the labor supply effects of a National Insurance reform cannot be analyzed entirely independent of the benefits in the public sector, and that the early retirement interacting with public sector benefits can have large effects on incentives. We cannot, of course, expect to see the extreme influx of workers to the public sector predicted by the model in real life: The model supports the conclusion of earlier studies that the incentives to retire using the AFP scheme are strong, and the fact that people do continue to work after qualifying for early retirement is hard to reconcile with a positive value on home production.

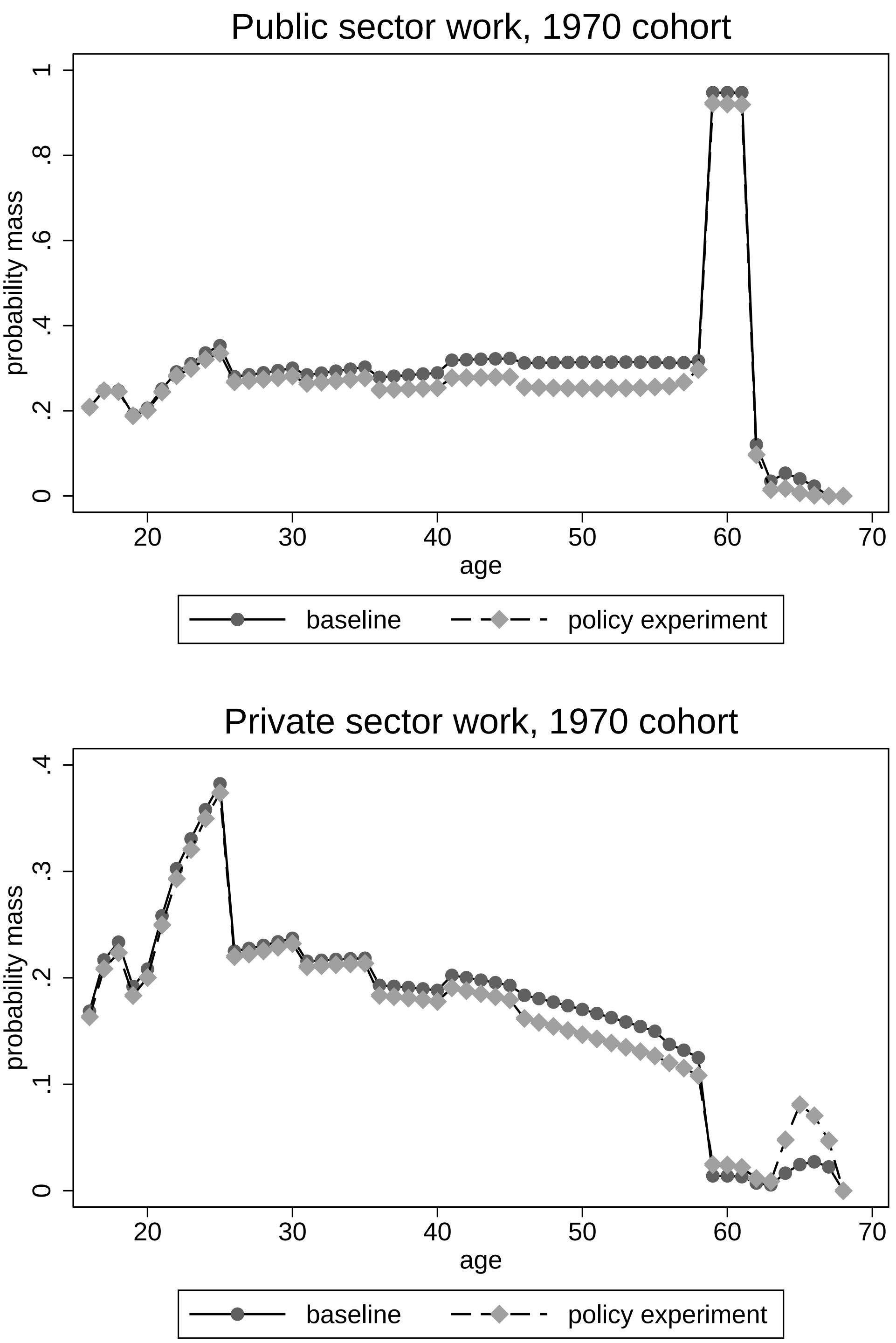

Figure 7: **Policy experiment.** A removal of the "twenty-best-years" rule and a 5% reduction in National Insurance benefits. The "baseline" is the current policy, under which the model is estimated.

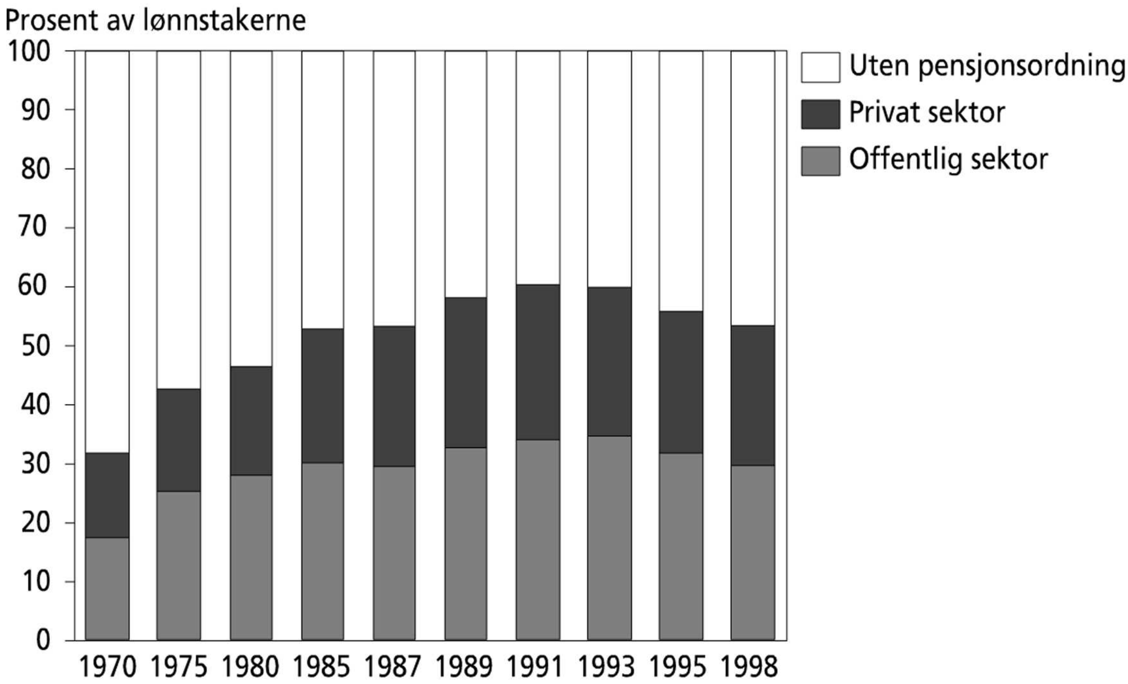

Figure 8: Share of workers covered by supplementary pension plans.
Public sector at bottom, private sector in the middle and no coverage on top. The figure is taken from Pedersen (2000) and is based on Labor Force Survey data (AKU).

Appendix A Institutional setting: The Norwegian pension benefits

In this section I describe the Norwegian pension benefits at a level of some excess detail, to make it possible to evaluate the representativeness of the incentives that are embedded in my empirical model.

There are many different private sector schemes. The great majority are also of the defined benefit type, but employers typically pay one of the large insurance houses to provide actuarial annuities, and firms contribute based on the characteristics of individual workers. These terms are much like what individuals can obtain in individual retirement savings schemes. If a worker leaves a firm before retirement, the insurance firm providing the retirement issues a bond guaranteeing an annuity at retirement corresponding to the contributions of employers. Not everyone are covered by the private sector supplementary pension schemes, but if a corporation choose to provide a pension scheme, it must do so to all employees in order to expense contributions before taxes.

A.1 National insurance

The National Insurance pension benefits consists of a basic pension provided to all with more than three years of residence.[16] The legal retirement age is 67. There are three elements in the National Insurance old age pension. There is the **basic pension**, the **special supplement** and the **supplementary pension**. The basic pension and the special supplement together makes up the **minimum pension**, and is not dependent on past contributions to the national insurance system, but is available to all residents. If they have lived less then 40 years in Norway, the minimum pension is reduced accordingly. For national insurance and public sector pension benefits, all earnings (no capital income, no pension benefits) are calculated in units of **basic amount** ("grunnbeløp") G, the unit of account. Inflation is corrected for by adjustments of G, sometimes several times a year, but G is also meant to follow the average wage growth, although it has grown slower than wages historically, and only since the last revision in 1992 of the National Insurance in 1992 has there been any significant real growth in the value of a basic amount.

The basic pension is $1G$, except when the pensioner has a spouse who works and earn more than $2G$ or when the spouse is retired, in which case the basic pension is $0.75G$ for each.

For single persons or for people with non-retired spouses, the full special supplement is $0.7933G$. If the pensioner is providing for a spouse over the age of 60, the couple together gets $1.5866G$. For a couple where both spouses are retired, the supplement is also $0.7933G$ unless one them has a higher supplementary pension, in which case the special supplement to the spouse is $0.74G$.

The supplementary pension is based on earned incomes in the past. Earnings from a given year is first mapped into a number of **pension points**. Over time, there have been some changes in how pension points have been calculated from earnings, and Figure 9 shows how the progressiveness has changed over time.

The yearly pension points are summarized by a **final points score**, pp^f. This is the average of the 20 highest yearly pension points earned. If a pensioner has earned pension points in $x < 20$ years, it is the average of these x years. The final supplementary pension in a year t is then $G_t \cdot \text{pp}^f \cdot \sum_j \pi_j / 40$, where the summation j is over the working years but limited to 40 years. Work experience before 1992 is awarded a replacement rate $\pi_t = 0.45$, while experience from 1992 and onward is given a replacement rate $\pi_t = 0.42$. The national insurance old age pension is then the basic pension plus the maximum of the special supplement and the supplementary pension.

[16]For a more detailed description, see *The Norwegian Social Insurance Scheme 2002*, `http://www.dep.no/sos/engelsk/publ/handbooks/044051-120003/index-dok000-b-n-a.html`.

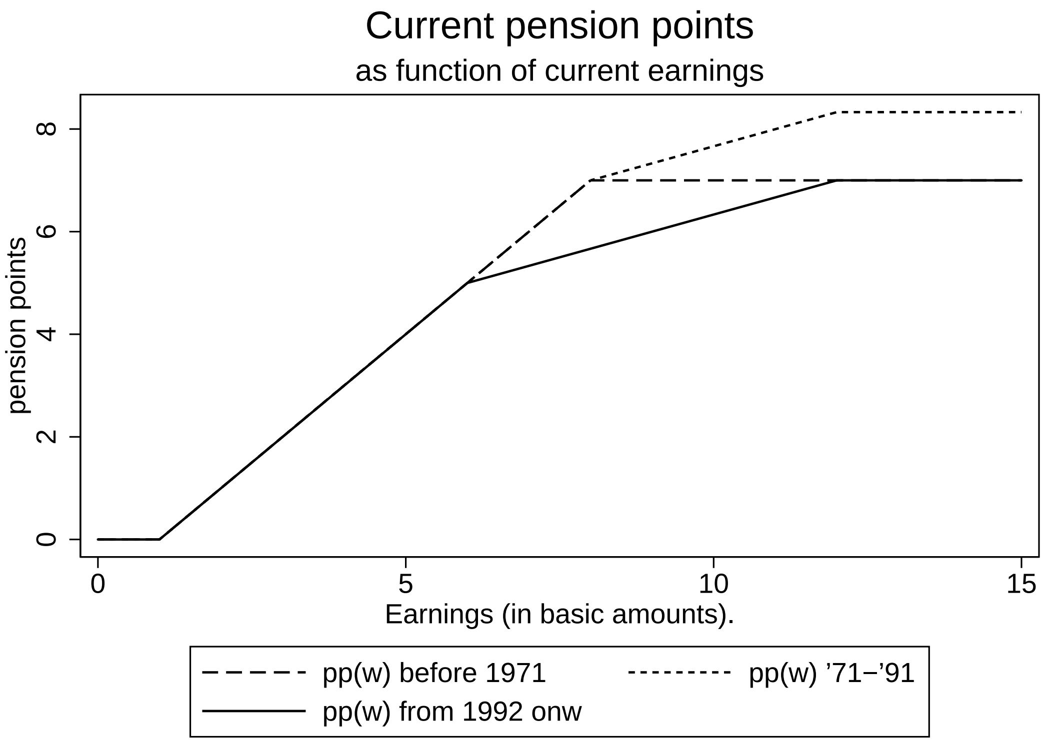

Figure 9: Pension points as functions of current earnings. Earnings in units of basic amount (G). In the year 20000, one basic amount was NOK 48400, approximately USD 7500.

A.2 adding public sector benefits

The public sector benefits are different. The system is complex, but the main characteristics for people who stay in the public sector is possible to explain. The various public sector insurance funds have agreements that the fund responsible for the individual when retirement takes place pays all benefits. For employees of the central administration, this system also applies for mobility between the Scandinavian countries.

There is a guaranteed benefit level of $0.66 \cdot \min(w^f, 12G_t)\lambda$, where w^f is the wage at the end of the career and $\lambda \in [0, 1]$ is the share of a "full career" worked in the public sector. For a "full career" in the public sector, only 30 years are needed – but if a person does not retire from the public sector, the "full career" career is calculated using between 30 and 40 years, depending on how long the actual career could have been if the retiree had stayed on in the public sector. If the ending wage is low in comparison with the wage over the rest of the career, adjustments will be made. For the years before the record wage, the record wage is applied. If a person reaches 30 years of experience at age 67, but there was a decrease in wage from 100 to 70 at age 60, the "final wage" will be calculated as $23/30 \cdot 100 + 7/30 \cdot 70 = 93$. But if the individual in fact has 37 years of experience in the public sector at retirement age and the record wage of 100 occurred at the end of a 30-year spell, the ending wage will be applied as 100.

The public sector benefits are only an addition to the national insurance, these two benefit levels are consolidated in ways that most Norwegians find somewhat cryptic.[17] An individual is guaranteed the maximum of national insurance benefits and public sector pensions, and the specifics is such that usually people will get more than this. The pension paid by the public sector pension fund is first reduced by $0.75 \cdot \lambda \cdot G_t$, corresponding vaguely to the basic pension. The supplementary pension from the national insurance is in principle deducted by a λ factor, unless parts of it was earned in the private sector. In this case, the part earned outside is not included in the deduction from the public sector. The calculation of the deduction is "simplified" by using the final wage to calculate a "fictitious" number of pension points, and the pensioner is allowed to keep an amount $(w^f - G^f) \cdot y^p/40 \cdot \pi \cdot G_t$ of the national insurance benefit.

To a first approximation, the pensioner gets the maximum of his guaranteed level from the public sector fund and what he is earned by the national insurance.

A.3 private sector benefits

Less than half of all those employed in the private sector are members of corporate sponsored pension plans, see Figure 8 taken from Pedersen (2000). In most cases, the firms buy these pension services from private insurance companies. While there

[17] It is regulated in *Lov om samordning av pensjons- og trygdeytelser*, 1957-07-06 no. 26, `http://www.lovdata.no/all/hl-19570706-026.html`.

are peculiarities with some of these programs, these programs are not consolidated with the national insurance. If an employer leaves a firm to work at another, the value of past pension contributions is calculated and a special retirement bond is issued. Individuals are allowed to continue paying into the programme on the same rate as those the corporation paid, but the next firm the individual applies to can also take on the previous pension contributions and respect those as if they were contributions into the new firm's programme. And lastly, the individual can hold on to the bond as such as cash in the stream of payments which will start at the official retirement age of 67.

Most private pension funds provide pension programmes which mimic the public funds (a fixed proportion of final wage), but at a corporation level. Also the guarantee is not absolute: One is in fact provided an annuity of the difference between the fixed percentage of the final wage and the national insurance level at the time of retirement, while the public programmes will compensate should the national insurance benefits decline in the future. If one should leave early in these programmes, ones future benefits are converted into a transferable promise of an annuity, but some corporations (e.g. Storebrand) advertises that a benefit to the corporation of these are more loyal employees. This has to do with how contributions and benefits were calculated in the past and there have in fact been clauses that firms can refuse to give older newly hired workers access to the corporation's pension programme. They have had their reasons for this: The contributions firms have made on behalf of their employees have traditionally been non-linear and concentrated in the latest years of a worker's career. In Figure 10 I have predicted typical age-profiles of contributions using the "annuity-method" (Clausen and Hersoug 2003).[18] This method has been the most popular one, since this method was required for tax exemption, and we see that it reaches 5% of earnings only as people are in their fifties.

A.4 early retirement

There are some local forms of early retirement possible. But in later years, a major policy innovation has been the AFP program. The "AFP" (Avtalefestet Pensjon) was introduced as an agreement between the parties in centralized wage-bargaining. Today, the programme allows retirement from the age of 62, but is not extended to everyone. It is extended to all of the public sector, and to the firms affected by private sector agreements between NHO and LO (the two main employer and employee organizations involved in wage bargaining). When it was introduced around 1990, it only allowed for retirement one year early, and has then been gradually extended. The latest extension (from 64 to 62 years) was in March 1998. Unfortunately, it not clear from our data which firms in the private sector is covered by the central wage-bargaining agreement. A further complication with this programme is that it depends on your history with the latest *firm* you are attached to: you need at least 3 years with the latest

[18] These are calculated by estimating a 4th degree polynomials of earnings on age on those of the 1937 cohort employed in private firms in 1986, then using this polynomial to calculate benefits.

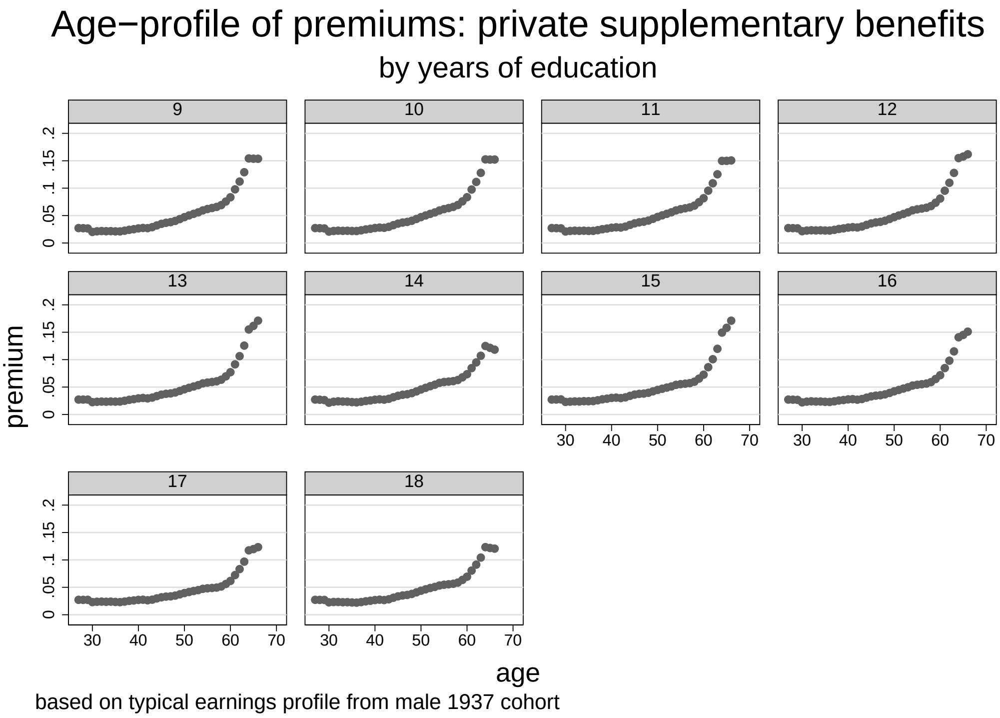

Figure 10: Typical private supplementary benefit premium profile

firm or 5 years with firms covered by the AFP agreement.

References

Arnstad, Marit, Oddvard Nilsen, Einar Steensnes, Jens Stoltenberg, and May Britt Vihovde, 2005. "Pensjonsavtale mellom regjeringspartiene og Ap/Sp." Oslo, 19. mai.

Bratberg, Espen, Tor Helge Holmås, and Øystein Thøgersen, 2001. "Assessing the effects of early retirement programs." mimeo, Norwegian School of Economics and Business Administration.

Brinch, Christian, Erik Hernaes, and Steinar Strøm, 2001. "Labour Supply Effects of an Early Retirement Programme." CESifo Working Paper No. 463.

Clausen, Jan R. and Tor Hersoug, 2003. *Bedriftens pensjonshåndbok*. Gyldendal.

Domeij, David and Paul Klein, 2002. "Public Pensions: To What Extent Do They Account for Swedish Wealth Inequality." *Review of Economic Dynamics*, 5(3): 503–534.

Ferrall, Christopher, 2004. "Estimation and Inference in Social Experiments." mimeo, Queen's University.

Ferrall, Christopher, 2005. "Solving Finite Mixture Models: Efficient Computation in Economics Under Serial and Parallel Execution." *Computational Economics*, 25(4): 343–379.

Fredriksen, Dennis, 1998. "Projections of Population, Education, Labour Supply and Public Pension Benefits: Analyses with the Dynamic Microsimulation Model MOSART." SØS 101, Statistisk sentralbyrå.

French, Eric, 2005. "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour." *Review of Economic Studies*, 72(2): 395–427.

Hægeland, Torbjørn, Tor Jakob Klette, and Kjell G. Salvanes, 1999. "Declining Returns to Education in Norway? Comparing Estimates across Cohorts, Sectors and Time." *Scandinavian Journal of Economics*, 101(4): 555–576.

Heckman, James and Burton Singer, 1984. "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data." *Econometrica*, 52(2): 271–320.

Hernæs, Erik, Marte Sollie, and Steinar Strøm, 2000. "Early Retirement and Economic Incentives." *Scandinavian Journal of Economics*, 102(3): 481–502.

Jia, Zhiyang, 2005. "Retirement Behavior of Working Couples in Norway. A Dynamic Programming Approach." Discussion Paper 405/2005, Statistics Norway.

Keane, Michael P. and Kenneth I. Wolpin, 1997. "The Career Decisions of Young Men." *Journal of Political Economy*, 105(3): 473–522.

Keane, Michael P. and Kenneth I. Wolpin, 2001. "The Effect of Parental Transfers and Borrowing Constraints on Educational Attainment." *International Economic Review*, 42(4): 1051–1103.

van der Klaauw, Wilbert and Kenneth I. Wolpin, 2005. "Social Security, Pensions and the Savings and Retirement Behavior of Low Income Households." Penn Institute for Economic Research 05-020.

Lee, Donghoon, 2005. "An Estimable Dynamic General Equilibrium Model of Work, Schooling, and Occupational Choice." *International Economic Review*, 46(1): 1–34.

Lee, Donghoon and Kenneth I. Wolpin, 2006. "Intersectoral Labor Mobility and the Growth of the Service Sector." *Econometrica*, 74(1): 1–46.

Midtsundstad, Tove, 2004. "Hvor mange har rett til AFP?" FAFO, rapport 728.

Møen, Jarle, Kjell G. Salvanes, and Erik Ø. Sørensen, 2003. "Documentation of the linked employer-employee dataset at Norwegian School of Economics." mimeo, Norwegian School of Economics.

Norges offentlige utredninger, 2000. "Tjenestepensjonen etter skifte av arbeidsgiver og strukturendringer i offentlig sektor." Arbeids- og administrasjonsdepartementet, NOU 2000:28.

Norges offentlige utredninger, 2004. "Modernisert folketrygd: Bærekraftig pensjon for fremtida." Finans- og Sosialdepartementet, NOU 2004:1.

Olsen, Grete, 1993. "Rutiner for ajourhold av det sentrale bedrifts- og foretaksregister." Statistisk Sentralbyrå, Notater, 93/13.

Pedersen, Axel West, 2000. "Konkurranseutsetting og pensjon." FAFO rapport 338.

Raaum, Oddbjørn, Kjell G. Salvanes, and Erik Ø. Sørensen, 2003. "The Impact of a Primary School Reform on Educational Stratification: A Norwegian Study of Neighbour and School Mate Correlations." *Swedish Economic Policy Review*, 10(2): 143–170.

Raaum, Oddbjørn, Kjell G. Salvanes, and Erik Ø. Sørensen, 2006. "The Neighbourhood Is Not What It Used to Be." *Economic Journal*, 116(508): 200–222.

Rønningen, Dag and Dennis Fredriksen, 2002. "Special tax rules for old-age pensioners." *Economic Survey from Statistics Norway*, X(2): 49–55.

Rust, John, 2001. "Modeling Behaviorial Responses to Changes in Social Security: A Life-Cycle Framework." mimeo, University of Maryland.

Rust, John and Christopher Phelan, 1997. "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets." *Econometrica*, 65(4): 781–831.

Statistics Norway, 2001. *Statistical Yearbook.* Statistics Norway.

Statistisk sentralbyrå, 1958. "Det norske skattesystemet 1958 - The Norwegian System of Taxation 1958." Samfunnsøkonomiske studier 7.

Statistisk sentralbyrå, 1968. "Det norske skattesystemet 1967 - The Norwegian System of Taxation 1967." Samfunnsøkonomiske studier 20.

Statistisk sentralbyrå, 1975. "Historisk oversikt over skattesatser m.v.:Del I, årene til og med 1969." Arbeidsnotater IO 75/5.

Statistisk sentralbyrå, 1988. "Skatter og overføringer til private: Historisk oversikt over satser mv. Årene 1970-1988." Revidert utgave, rapporter 88/20.

Statistisk sentralbyrå, 1994. "Skatter og overføringer til private: Historisk oversikt over satser mv. Årene 1975-1994." Rapporter, 94/21.

Statistisk sentralbyrå, 1995. "Historisk statistikk 1994 - Historical Statistics 1994." Norges offisielle statistikk, C188.

Statistisk sentralbyrå, 2001. "Utdanningsstatistikk: Dokumentasjon 2000 av den individbaserte utdanningsstatistikken." Norges offisielle statistikk, C645.

Vassenden, Elisabetta, 1995. "The register of the level of education: a product of the integration of other registers, census and surveys." Statistisk sentralbyrå, Notater 95/4.