



Three Essays in Applied Economics

Social Insurance

Junge, Martin

Publication date:
2007

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Junge, M. (2007). *Three Essays in Applied Economics: Social Insurance*. Department of Economics, University of Copenhagen. Rød Serie, No. 122

PhD Thesis No. 153

Three Essays in Applied Economics: Social Insurance

by

Martin Junge

September 2007
Department of Economics
University of Copenhagen
StuDiestraede 6, DK-1455 Copenhagen
www.econ.ku.dk

Three Essays in Applied Economics: Social Insurance

By Martin Junge

Table of Contents

Preface

Acknowledgement

Chapter 1..... An Econometric Model of Early or Ordinary Retirement in Denmark

Chapter 2..... The Interaction of Long-term Unemployment and Retirement

Chapter 3..... Labour supply in an economy with a generous UI system

Preface

This thesis contains three chapters. They are interrelated; but they can be read independently. Chapter 1 and 2 have in common the focus of old age insurance and chapter 2 and 3 are interrelated by the focus on unemployment insurance. All three chapters are connected to the analysis of social insurance. Here we briefly discuss the topic of social insurance and the methodology applied in the chapters before turning to a brief summary of each chapter.

Social insurance is concerned with government provision of unemployed, disabled or old-age. It is financed by employers, employees and government. The institutions that provide social insurance vary a great deal across countries in the world. All this underlines that as a research field social insurance is very broad and complex. Social insurance influences and distorts decisions of those who receive and pay for it. Social insurance affects people and especially those at risk of losing their basic income resource, human capital. Securing the income base, however, is identical with an efficiency cost in the form of 'potential' moral hazard problems as social insurance crowd out private insurance or reduces labour supply. This makes it a very fascinating and essential research topic. This thesis only deals with certain aspects of social insurance. Chapter 1 and 2 are devoted to old age insurance, while chapter 3 is devoted to unemployment insurance.

The analytical framework in this paper is that of neoclassical theory and empirical measurement. We embed individual behaviour in computational models that are abstractions of reality. In particular, we will assume that individuals are rational and pursue individual interests i.e. maximizing consumption and leisure bundles under a budget constraint. The question that is frequently asked in this thesis is how large is the distortion induced by social insurance on leisure/labour supply? Further can existing policy proposals alleviate some of the distortions? In each chapter we devote a section to evaluations of the various trade-offs involved in setting social insurance policy.

Social insurance involves a complex rule set. Oftentimes multiple welfare programs exist and individuals can choose between the programs (e.g. unemployment insurance or social assistance). Moreover ‘grandfathering’ effects are frequently in effect, which makes the choice set very different across cohorts. The resulting budget set has kinks, holes and corners, which makes empirical models very complex. The thesis gets around these obstacles in a ‘standard’ way by imposing an additional restriction that choices are discrete. The standard first order conditions from neoclassical theory are therefore replaced by a series of inequality restrictions.

In Chapter 1 we estimate the effect of social security and early retirement eligibility and benefit on retirement in Denmark in a reduced form discrete choice model. Due to the endogeneity of early retirement eligibility, we estimate the eligibility condition simultaneously with the discrete choice of retirement by a bivariate probit model. The retirement equation allows the error term to pick up missing variables, unobserved heterogeneity and possible problems with the time-stamping of retirement and age. We use variation from major social security and early retirement reforms as well as minor year-to-year changes in actual social security and early retirement rules from 1980 to 1996 in a micro simulation model to compute the present value of social security wealth. This incentive measure enters into the retirement decision and also affects eligibility. Exogenous changes in the retirement programs are necessary for identification of incentives. We use a data set compiled from administrative registers in Denmark. First, we find that eligibility is endogenous. Second, public provided support in old age is an important predictor of retirement, but is not influenced much by eligibility to early retirement or serial correlation in the error terms. However, other variables, e.g. age and health, are more sensitive to these. Policy simulation reveals an important effect of retirement route choice on average retirement age, which is not considered in other studies.

In Chapter 2 we are modelling and estimating the joint process of exiting from the labour market through unemployment and retirement. Unemployment is a common pre-retirement state and around 25 per cent of transitions to retirement at age 60 are from (long-term) unemployment. Despite the large amount of unemployment prior to

retirement it is surprising that unemployment is not more widespread as the unemployment insurance scheme is more generous than the early retirement scheme. The model is a dynamic discrete choice model with imperfect control. It is estimated using maximum likelihood on a sample of workers, whose main retirement income is public provided support. Individuals maximize remaining lifetime utility. In each period, they control their labour market state by choosing between unemployment, retirement or full time work. The control is imperfect in the sense that the decision to stay in a given job can be overruled by a layoff, which is modelled as a function of creative destruction in plants. Utility is derived from income and leisure and includes taste shifters. The value of leisure varies across unemployment and retirement due to the absence of search requirements in retirement. The model incorporates the social insurance system in detail, which induces dynamic incentives, e.g., duration dependent benefits and an interaction between the unemployment benefit level and early retirement benefits. Moreover wages depend on the history of labour supply, general human capital (experience) and firm specific human capital (tenure). The results of the paper points at very large search cost for the unemployed, though search also reduces cost of getting a new job. Simulation with the model shows that unemployment related retirement schemes can be extremely popular even at low economic incentives. Moreover, we compute individual's willingness to pay for an unemployment scheme without search cost. For the unemployed, we find, they are willing to live close to subsistence level.

In Chapter 3 the emphasis is on the need to allow for considerable heterogeneity in the distaste for work that is correlated with productivity when considering the design of a UI system. The Danish UI system is – for low wage workers – very generous. Six months of work entitles low wage individuals to 30 months of UI benefit at 90% of their wage. By any standards this is likely to distort incentives. First, workers who would normally work all the time have strong incentives to cut back work substantially. Second, workers, who would not normally work, have strong incentives to work for six month to entitle themselves to UI benefit. Despite the generosity, we find that a very large proportion of our sample of low wage single male workers works most of the time and that a substantial fraction works 6 months. In a simple labour supply model we can

only rationalize this by posting substantial heterogeneity on preferences. The analysis suggest that preferences are bimodal.

Acknowledgement

A number people must be thanked for making this thesis possible. First, my supervisor Martin Browning has been a very inspiring mentor. My studies in economics did take a turn when I by coincidence joined a course on consumption taught by Martin. Until then most of my interest in economics was in microeconomics and economic policy issues. However, bridging the gap between theory and empirical economics as was done in this course and studying Angus Deaton and John Muellbauer “Economics and Consumer Behavior”, Cambridge University Press, 1980, put economics in a completely different perspective for me.

I will also thank Department of Economics, University of Copenhagen. First, ‘pizza’ seminars were a nice forum to present and discuss own and other PhD-students papers. Second, lunch and coffee breaks with Henning and Jacob was a good opportunity to share good and bad side of life. Third, Department of Economics provides excellent research facilities.

During the PhD I stayed at CEMFI in Madrid from November 1, 2001 to February 1, 2002. I would like to thank professors and PhD-students for great hospitality. It was a very stimulating environment as Department of Economics, University of Copenhagen only recently had started to expand the area of micro econometrics, which was very well advanced at CEMFI at this time.

To do empirical economics we need data. Thanks to DREAM for providing data for this thesis and Statistics Denmark to allow office space to access the data.

Finally, I need to thank Svend E. Hougaard Jensen, Ulrich Kaiser, Hans Christian Kongsted and Nikolaj Malchow-Møller for excellent help in the last stage of the thesis. Without their support I would not have been where I am today.

An Econometric Model of Early or Ordinary Retirement in Denmark

Version 1

Martin Junge¹

Centre for Economic Business Research, CBS

Department of Economics, University of Copenhagen

Abstract: This paper we model the discrete choice of retirement using a fully parametric model that accounts for unobserved heterogeneity, serial correlation in the error terms and endogeneity of the eligibility to early retirement. We use a data set compiled from administrative registers in Denmark. First, we find that eligibility is endogenous. Second, public provided support in old age is an important predictor of retirement, but is not influenced much by eligibility to early retirement or serial correlation in the error terms. However, other variables, e.g. age and health, are more sensitive to these. Policy simulation reveals an important effect of retirement route choice on average retirement age, which is often not considered in other studies.

Keywords: Retirement, social security, panel data, bivariate probit.

¹ The Author would like to thank Martin Browning, Ulrich Kaiser, Hans Christian Konsted and Nikolaj Malchow-Møller for valuable comments.

1 Introduction

This paper estimates the effect of social security and early retirement eligibility and benefit on retirement in Denmark in a reduced form discrete choice model. Due to the endogeneity of early retirement eligibility, we estimate the eligibility condition simultaneously with the discrete choice of retirement by a bivariate probit model. The retirement equation allows the error term to picking up missing variables, unobserved heterogeneity and possible problems with the time-stamping of retirement and age. We use variation from major social security and early retirement reforms as well as minor year-to-year changes in actual social security and early retirement rules from 1980 to 1996 in a micro simulation model to compute the present value of social security wealth. This incentive measure enters into the retirement decision and also affects eligibility. Exogenous changes in the retirement programs are necessary for identification of incentives.

The Danish labour force participation rate for elderly workers is high by international standards both today, in the 1980's and in the beginning of the 1990's, the period considered in this paper. Two main reasons for this exist. First, Danish women have had a high participation rate in the labour market for many years, starting in the fifties. Second, the normal retirement age was 67 years in the period under investigation, which is high compared to other countries. Policy makers and economists today are, however, concerned about the future fiscal balance (The Economic Council ("Det Økonomiske Råd") 2005; The Danish Welfare Commission ("Velfærdskommissionen", 2006). The average retirement age, about 61 years of age, is much lower than the official normal retirement age, which is mainly due to a very generous early retirement scheme enacted in 1979.

Retirement has been studied extensively in the literature for the last two decades. An important result of existing studies is that publicly provided old age insurance is an important predictor of retirement. Other important factors that determine retirement have also been identified, such as wealth, health and Medicare² (in countries where this is an issue), and labour market opportunities. It is of utmost importance for

² Medicare is not relevant in Denmark as hospitals and doctors are free and prescription drugs and other treatments are either public available or largely subsidized.

economic policy to obtaining reliable estimates of the empirical magnitude of these effects in order to design an optimal retirement policy.

The literature relevant for this study mostly uses reduced form analysis of the “pull effects” of social security and early retirement benefit. Stock et al., 1990, find strong and significant effects of variables measuring the incentives of firms’ pension plans, e.g. the option value of postponing retirement. In the option value model an individual continues to work if the expected value of retirement in the future is worth more than the value of retirement today. The option value approach has been widely used in many papers for different countries and for various retirement routes (see i.e. NBER volumes by Gruber et al., 1998 and Gruber et al., 2005). In most, if not all of these studies, the incentive measure and declining labour force participation rates are strongly correlated. A very closely related incentive measure, the “peak value”, was introduced by Coile et al., 2000. The peak value measures the difference between maximum expected social security wealth of the future and social security wealth today. It is hence the financial equivalent of the option value and can easily be computed by a micro simulation model with pension system rules. Coile et al. suggest entering this into a probit model with retirement decision as the dependent variable.

Having established some of its successes the literature has not matured yet. First, it is often the case that retirement routes are multiple. Thus, there can be endogeneity involved in the retirement route. The response to this issue in the reduced form literature has been to estimate the take up rate and use these as weights in estimating expected social security wealth (Börsch-Supan, 1999). The leading example is disability pensions, which are restricted to the disabled. First, the screening process is seldom observed, which makes this route partially observed. Second, individuals might influence the outcome in this process as discussed by Börsch-Supan. As Börsch-Supan notes, the estimation of take up rates and applying the weights to aggregation across retirement routes can be seen as an instrumental variable procedure. But, the literature does not discuss identifying assumptions or provide a test for whether choice of route is endogenous. Exceptions do exist. For example Chan and Stevens, 2004, estimate a model with forward looking incentives measures included as discussed above. Since their dependent variable is expected retirement age, they use a

fixed effect estimator. This can under certain simplifying assumption account for the endogeneity of choice of retirement route.

Moreover, failure of accounting for unobserved heterogeneity can bias the incentive estimates. The pooled probit estimator in panel data is consistent; but inefficient then errors are dependent across time. The problem here, however, is more severe since retirement is often modelled as an absorbing state. While lagged dependent variables do not enter the equation directly, the absorbing state changes the model into a sequential choice model. It is known that parameters are biased in this model, if the errors are dependent (Maddala, 1983). Börsch-Supan, 1999, shows that the consequences can be severe.

This paper contributes to the literature on reduced form models of retirement by in two ways. First, we account for possible endogeneity of eligibility to early retirement. Eligibility requires long-term UI fund membership, which is voluntary in Denmark. Evidence of individuals timing their UI fund membership to become eligible for early retirement benefit is given in Ejrnæs et al., 2006. Long-term membership condition was tightened from 10 to 20 years in 1992. Some cohorts were unable to meet the new requirements. Instead they could become eligible to early retirement, if they joined the UI funds immediately. The result was a large increase in UI fund members. We model eligibility to early retirement in a separate equation, which leads to a bivariate probit model, where it is possible to test for endogeneity of eligibility of early retirement.

Second, we account for unobserved heterogeneity by estimation of a probit model with a flexible error structure. Since consistent estimates are obtained in the pooled probit model, this is rarely done in practice. However, the optimal stopping assumption in the retirement literature changes the problem into a sequential choice model. It is well known that this model give biased estimates if we do not control for unobserved heterogeneity (Maddala, 1983). The same problem is recognized by Börsch-Supan, 1999, who estimates the model with an AR(1) error and an random effect. He finds that the parameters changes dramatically, if the more sophisticated error structure is used, which has not had an influence on the literature following this paper. Therefore, it is of some interest to repeat Börsch-Supan on flexible error structure. We have generalized the error to an ARMA(1,1) model. The MA(1) error

term allows correlation between two consecutive periods to pick up problems with i.e. time-stamping.

The outline of the paper is as follows. First, we give an overview of the institutional setting of the Danish social security system in Section 2. We also discuss the recent reform and suggest a way to looking at these from a simulation perspective. In Section 3, the econometric model is derived. Section 4 discusses sample selection issues and we define eligibility to early retirement and retirement. In Section 5 we lay out the micro simulation model we use. Description of retirement and eligibility is done in Section 6. Results are presented in Section 7, and before wrapping up in Section 9, we simulate two fictive policy proposals in Section 8.

2 Institutional setting

This section focuses on normal retirement, early retirement and eligibility conditions for early retirement,³ which are relevant for this paper. The description of the institutional settings relates to the period 1980 to 1996 which also is the period we have data for. Since the 1999 reform of early retirement is part of our policy analysis, we will touch upon this in the end of this section.

Normal retirement or old-age-pension (“folkepension”) is provided from the age of sixty-seven. It is a two-tier universal provided scheme, consisting of a base amount which is means tested against earnings and a supplement which is means tested against household income. Before 1993, the base amount was untested. Obtaining the full amount of old-age pension requires at least forty years of residence in Denmark. Shorter residence periods reduce the amount proportionally. The system does not depend on previous work history and is only residence and age based. Historically, couples and singles have received different amounts. Delaying retirement does not lead to an adjustment of future benefits. Therefore, the old age schemes induce large incentives for retirement at age 67. The system is financed through general taxes.

Disability pension (“førtidspension”) is based upon a rather complex set of rules. The main criterion for eligibility is reduced work ability. A reduction of at least fifty

percent is needed to qualify for that scheme. However, individuals aged 50 years and above may receive disability pension for social reasons. Delaying retirement does not increase future benefits and therefore no actuarial fair adjustment takes place. Benefits are more or less comparable with old age pensions except that work reduces benefit.

The main route for early retirement (“*efterløn*”) was enacted in 1979. Eligibility is gained through membership of unemployment insurance funds (UI). It is worth noting UI membership is voluntary in Denmark. It is organized by trade unions; but membership in a trade union is, however, not required. A fund exists for all workers in the private sector.

The main rules for eligibility to early retirement are first entitlement to UI benefit, which requires six months of work within the last 36 months. Second, a long period of membership is needed. This eligibility criterion has been tightened throughout the years. In the 1980’s, ten years of membership within the previous 15 years were a condition for eligibility. For individuals born before May 1, 1930, only five years of membership within the previous ten years were required. This requirement was tightened to 20 of the previous 25 years in 1992. From the beginning of the 1990’s, the work requirement (six months within last 36 months) was abolished for individuals older than 50 years. These eligibility criteria were meant to target a group with a long service life on the labour market

The early retirement benefit level has also changed in the sample period. In the early period it was 90 percent of previous earnings, up to some maximum level of approximately 100,000 DKK in 1980 per year. The maximum is reduced to 80 percent after 30 months. After an additional 24 months, the maximum is reduced to 70 percent of the initial maximum level. The latter reduction was abolished in 1989. It is possible to maintain the maximum level from age 63 to 67 if the application for early retirement is delayed to age 63 since 1993. Finally, the amount of work allowed for individuals in this scheme has been rather strict, allowing for 200 hours per year. Furthermore, leaving the scheme meant that it was impossible to rejoin at a later date. Some of these restrictions were lifted in the beginning of the 1990’s. Part time early

³ The social security system and other retirement routes in Denmark are reviewed in detail by Bingley et al. (2002).

retirement was introduced and it was possible to re-enter once after a period of employment. The benefit is individual-based and does not depend on other household income.

The three routes to retirement are modelled in this paper. These constituted the most important pension options for two thirds of the labour force (Bingley et al., 2002) in the 1980's and the beginning of the 1990's. Some public employees could draw a civil servants pension after a long period in civil service. Some individuals with an academic degree additionally had savings in labour market pensions. For details about these groups, see Bingley et al., 2002. These groups will be excluded from our analysis.

In the future, the labour market pensions are expected to gain importance, as they have spread to larger groups in the labour market. Though projections from the Danish Welfare Commission ("Velfærdskommissionen") show that labour market pensions will only be the dominant source of old-age income in the highest income deciles in 2030 (The Danish Welfare Commission, 2006). Hence, publicly provided pension programmes will remain constituting the most important retirement scheme. Finally, note that private firm pension schemes are almost non-existent in Denmark.

In the policy scenario in Section 8 we will look at the early retirement reform in 1999. It is interesting, as at several points it changed the dynamic incentive structure. First, the period of UI fund membership was extended from 20 years to 25 years. In addition a contribution must be paid of approximately 4500 DKK in 1999 in each year. The ordinary retirement age was decreased from age 67 to 65. If retirement is postponed to after age 62 a tax credit is allowed at age 65. This credit is incremented for each 3 months of postponed retirement. Finally, the benefit level is 90 percent of maximum UI for those who retire before age 62. It is 100 per cent of maximum UI for those who retire at age 62 or later.

3 Econometric model

This section explains the econometric model. The left hand side variable is the discrete choice of retirement between age 59 and 71, where we assume that retirement

is an absorbing state. On the right hand side we include incentive measures calculated from actual rules from 1980 to 1996. As incentive measures we use social security wealth and peak value. We also add other explanatory variables like earnings, age, education, experience, gender, a proxy for health, cohabitation⁴, potential labour market experience and some measures of growth of employment within the plant individuals are employed as labour market shocks. Adding an unobserved iid normal distributed error generates the model of Coile et al., 2000.

We extend the model in two directions. The first direction is to extend the error term and account for unobserved heterogeneity, missing variables and problems with the dating of events. The second direction is to model eligibility to early retirement. This formulation leads to a discrete choice model with a dummy endogenous variable. The eligibility to early retirement is decided before retirement and we model eligibility to early retirement at age 59. The left hand side variable is eligibility at age 59 (see “Data” section for how eligibility is computed), and on the right hand side we include potential social security wealth, earnings, gender, occupation, industry, education, experience, a proxy for health, cohabitation, and measures of growth of employment within the plant individuals are employed.

Serial correlation and random effects in the panel probit selection model is often ignored. The main reason is that consistent, though inefficient, parameter estimates can be obtained without modeling the correlation in errors (Bertschek et al., 1998). Since we treat retirement as absorbing, the model essentially is a sequential choice model. Ignoring serial correlation hence leads to biased estimates which are one reason for including the flexible errors in this model (Maddala, 1983). To pick up unobserved heterogeneity and missing variables we include an AR(1) error and a random effect.

Another issue is that eligibility is dependent on age and the correct measurement of the timing of retirement becomes crucial. One problem here is that retirement is decided on the basis of income source type and labor supply in November each year whereas age is calculated as an individual’s age at the end of each year. It is therefore

⁴ The variable “cohabitation” will be used throughout the paper. It covers both married and cohabiting couples.

likely that we find a correlation in the error terms across two periods. It might also be the case that it is difficult for some individuals to explicitly time their year of retirement due to employment contracts. We therefore include an MA(1) term as well.

This final error structure is an ARMA(1,1) with time invariant random effects that leads to a multidimensional integral as we have 13 (age) periods, that needs to be integrated out. Fortunately, recent developments in estimation by simulation can solve these problems (Hajivassiliou et al., 1994).

Discrete models with dummy endogenous variable are quite common in economics.⁵ The common reference for estimation of this model is Heckman, 1978. Due to the presence of a dummy endogenous variables full information maximum likelihood are usually applied. Other methods exist, however. Angrist, 2000, for example suggests using instrumental variables in limited dependent variable models with dummy endogenous variables. Another approach uses natural experiments, which constitutes another option (see Krueger et al., 1992, on retirement). We use full information maximum likelihood. Identification of the model is achieved through functional form assumptions regarding the error terms, though it is common practice to have some exclusion restrictions to make it more robust towards misspecification of the error term.

More specific we formulate the retirement equation as:

$$y_{it} = 1 \text{ if } y_{it}^* = \beta^{ssw} (e_{it} ssw_{it}^1 + (1 - e_{it}) ssw_{it}^0) + \beta^{peak} (e_{it} peak_{it}^1 + (1 - e_{it}) peak_{it}^0) + x_{it}' \beta^1 + \varepsilon_{it} > 0$$

$$y_{it} = 0 \text{ otherwise}$$

Where y^* is a latent variable measuring the monetary value of retiring. We only observe whether this is positive or negative and hence y_{it} is one or zero. e_{it} is the endogenous dummy variable of eligibility, and ssw^j ($peak^j$) is potential social security wealth ($peak$ value) for ineligible and eligible to early retirement for $j = 0, 1$. X is a vector of exogenous variables. The error term has the following structure:

⁵ A leading example is the influence of fertility on labour supply of women (see for example Carrasco, 2002).

$$\begin{aligned}\varepsilon_{it} &= \lambda_{it} + \nu_i \\ \lambda_{it} &= \rho\lambda_{it-1} + \eta_{it} + \theta\eta_{it-1}\end{aligned}$$

Where η and ν are iid normal distributed error terms and independent of X . From this and an assumption of stationarity, we can derive the covariance matrix as:

$$\begin{aligned}E[\varepsilon_{it}^2] &= \sigma_\lambda^2 + \sigma_\nu^2 \\ E[\varepsilon_{it}\varepsilon_{it-1}] &= \rho\sigma_\lambda^2 + \theta + \sigma_\nu^2 \\ E[\varepsilon_{it}\varepsilon_{it-k}] &= \rho^k\sigma_\lambda^2 + \sigma_\nu^2, k > 1\end{aligned}$$

$$\text{where } \sigma_\lambda^2 = \frac{1 + \theta^2 + 2\rho\theta}{1 - \rho^2} \sigma_\eta^2.$$

In a discrete choice model it is necessary to set the scale for identification. We impose $\sigma_\lambda^2 + \sigma_\nu^2 = 1$.

Eligibility for early retirement has the following form

$$\begin{aligned}e_i &= 1 \text{ if } e_i^* = z_i'\delta + \zeta_i > 0 \\ e_i &= 0 \text{ otherwise}\end{aligned}$$

Again e^* reflects the latent value of being eligible to early retirement. We only observe whether e^* is positive or negative through e . Z is a vector of exogenous variables. Note that eligibility is formulated as a cross section estimation at age fifty-nine. In this respect, it is closely related to modelling initial conditions as we only consider retirement from age 59 to age 71. The error term is normalized to have variance 1, but is possibly correlated with the time-invariant error in the retirement equation, $E[\varepsilon_{it}\zeta_i] = \rho_{\varepsilon\zeta}$.

With the formulation above, a fixed effects estimator of the retirement equation would wipe out all selection (due to time-invariance). However, probit models with fixed

effects are only in their infancy (Arrellano, 2005).⁶ Instead the random effects specification is only robust to selection if the random effect is uncorrelated with the participation error (Verbeek et al., 1992). By joint estimation of retirement and eligibility we can test for whether this correlation is significant.

Even though the model is identified by functional form of the error terms, exclusion restrictions are often imposed to help identification. We propose the possibility of using individuals' industry affiliation and occupation to explain retirement only through eligibility. The main idea is derived from Parson et al., 2003, where social pressure on UI fund membership differs across industries and occupation. However, in many studies industry affiliation and occupation has been an important predictor of retirement due to correlation with other determinants of retirement, i.e. the demand for labor and physical requirements of jobs. In the model we include measures of health and labor market shocks to capture some of these effects. Another exclusion restriction is potential social security wealth for eligible (ssw^1). Since it only enters in the retirement equation for those with $e_{it} = 1$. Another argument for the exclusion restrictions is that the eligibility condition is modeled at age 59 which is well below the average retirement age at 61.7 in the sample.

The econometric model is estimated using simulation methods to solve the multidimensional integral in the likelihood function. We use the GHK simulator (Hajivasiliou et al., 1994). The procedure produces biased estimates unless the number of simulations goes to infinity. In practice, however, most researchers set the number of draws to 10. We also use 10 draws and uses antithetic variates, which is a variance reduction technique, implying a total of 20 draws.

We estimate the simultaneous model and a number of simpler models nested within the simultaneous model. This allows for standard likelihood ratio test, also then we use simulation-based estimation (Hajivasiliou et al., 1994). The correlation between the error terms of eligibility, early retirement and ordinary retirement are of particular interest.

⁶ We experimented with using a conditional logit model. But the estimation was very unstable. If we excluded age and year variables and some of the other explanatory variables the model converges, but the results on incentive measures were very hard to understand.

A limitation of this model is first order serial correlation and random effects indicate omitted variables. Second, availability and test of exclusion restrictions are another potential problem. Moreover, we have excluded a number of important predictors of retirement, i.e. labour supply of spouses and lagged unemployment, since these are endogenous variables. This essentially means that they are partly absorbed by the error term. When we conduct policy simulations, these are performed conditional on policies that do not affect the error term, which leads to biased policy predictions.

4 Description of retirement/eligibility and sample selection

The data is compiled from various administrative registers in Denmark.⁷ Every person with a residence in Denmark has a personal ID number. This ID number is used for many types of transactions: paying taxes, receiving various government benefits, unemployment insurance, job centres, bank accounts, etc. Statistics Denmark has compiled a labour market database from some of these transactions (named “IDA”) and access for researchers and others can be granted. The database essentially covers the adult population in Denmark, but our data consists of a 10 percent random sample for 1980 (the first year of “IDA” data), and individuals are then followed to until today.⁸ These data also allow us to track spouses and children living at home. This data has various attractions: it covers the population of Danish residents. It does not bear any survey component and therefore problems associated with non-response and attrition do not arise. The complete data set contains 400,000 adults and tracks them for 18 years, making it a total of more than seven million person year observations.

Definition of retirement and eligibility

We identify retirement based on primary attachment to the labour market in the last week of November. Both actual retirement and other withdrawals from the labour force are considered “retirement”, while employment and unemployment are considered “work”. However, it only requires very little work to be classified as in

⁷ The compilation we use was kindly provided to us by Ministry of Finance’s group DREAM.

⁸ The final year is 2004 in “IDA” at the moment. The estimation period, 1980-1996 is chosen so minimize the impact of labour market pension. .

“work” state for individuals that actually have their main income from pensions.⁹ We therefore impose the extra requirement that wage earnings need be higher than pensions in order to be classified as employed/unemployed, otherwise individuals are classified as “retired”.

Eligibility to early retirement is not observed in the data and must be computed by looking at retrospective membership of UI funds. Note that for many individuals eligibility can be decided on membership in UI funds in the 1970’s, which is unobserved. We split the sample in three parts. The “eligibles” at age 60, defined as belonging to cohort 1925-1930 (1931-1936) and *observed* membership of an UI fund for at least five (ten) years before 60 years of age. The “ineligibles” are defined as belonging to cohort 1925 to 1936 who have not been members of an UI fund for at least six years after 1980 or for at least six years within last ten (15) years if available for cohorts 1925-1930 (1931-1936). The “remaining”, where either eligibility cannot be determined or the exact age at which eligibility is gained cannot be determined, is not used as timing uncertainty can provide the model poor policy simulation policies.

Figure 1 presents the retirement hazard (transition from “work” to “retirement”) by eligibility type. Those eligible (at age 59) for early retirement have very low retirement probabilities before age 60. For ineligible the retirement is much higher before age 60. It reflects the fact that to become eligible a very stable employment pattern must be observed. It is likely that some of those retiring at age 55 because of for example disability would have been eligible if they were not disabled. For the same reason we avoid modelling disability and the decision of eligibility before age 60 by focusing on the 59+ group in the estimations. Furthermore, we remove all individuals retired before age 59. After age 60 a large increase in the hazard is seen for the eligible. It also quite high for age 61 and especially age 68, which motivates our MA(1) error in the retirement equation. For ineligible there is a slight increase in retirement after age 60, perhaps reflecting declining health conditions or easier access to disability pension, which can be given for social reason at that age. At the age of old age pension (67 years) the retirement pattern looks similar across eligibility groups.

⁹ In each year individuals can be both working and retired. Then Statistics Denmark classify primary attachment work has much higher priority than pension (see Statistics Denmark, 2000).

Sample selection

As discussed under the institutional setting above, a large group of individuals is covered by public pensions only (and perhaps also by savings on bank personal accounts). This is the group we are interested in this paper. The data contains net taxable wealth, which includes real estate and portfolios on bank accounts and deposits, but does not contain either labour market pensions or private pension savings accounts. We do also lack information on whether public employees are eligible for civil servants pensions. For this reason we exclude all public employees and employees in transport, post and telecommunication, where the latter industries are - to a large extent - characterized by government involvement. Self-employed individuals for whom it is often quite difficult to obtain income measures and hence to compute a reliable social security measure are removed. The final group of agents left out of the sample are academics (18+ years of schooling), who are more likely to have a labour market pension. The final sample consists of private sector employees without academic education. The sample is further restricted to individuals working at age 59 and who are observed for at least two periods. The estimation sample consists of 15,580 person-year observations¹⁰.

Table 1 provides insights into the sources of income for the selected sample. Retirement income sources are early retirement benefits (if eligible), disability- and old age pensions and “Other income” consists of earnings, capital income, pensions (taxable), and ATP¹¹. The median retirement income for the retired under the age of 67 is ten times larger than median other incomes. For age group at 67 and older it is only 1½ times larger. The main reason “other income” is important for 67+ is that the benefit level is much lower and ATP and other pension plans are received.

5 Incentive measures

Coile et al., 2000, suggest using the peak value to incorporate the incentive effects of retirement. The peak value is the difference between the maximum present value

¹⁰ In addition to the selection mentioned in the text we also remove individuals with very low (<8,000 (1980 DKK) or very high earnings (> 1,000,000 (1980 DKK) and we require individuals to be in work at age 59.

social security wealth in the future from continued work and the present value of social security wealth today. If the peak value is positive additional years of work will generate higher social security wealth. The approach is closely related to the option value of postponing retirement. In the option value the maximum expected value of continued work is compared to today's social security wealth under a payoff function.

In Appendix A we develop the detailed computation of retirement incentives¹². Here we just note that individuals take the system as it is and adjust future benefits, with a growth factor of one percent (in 1983-1986 the government did not change the maximum early retirement benefit and in this case it can be discussed whether the one percent increase should be used) for the future. The present value calculation uses a real interest rate of three percent and a gender and age specific survival probabilities from publicly available life tables. If individuals are part time insured we calculate the early retirement benefit as for full time insured but adjust the maximum with $2/3$ of a full time insured. Old age pension is modelled for couples and singles separately because of the tax treatment. Indexation of old age pension is again with a real growth rate of one percent. For disability pension we take a very simple approach and assume all ineligible to early retirement can draw this pension for social reasons. It is in contrast to the literature, where the take up rate is modelled and applied to construct a measure of expected social security wealth. When we assume everyone can draw disability pension, this induces a downward bias in the incentives of retiring. The micro simulation model includes major reforms in early retirement and normal retirement system as well as minor year-to-year adjustments, which we have collected through annual publication in the area of social insurance (see Forlaget Idag).

The changes in retirement benefit rules are important for identification. Early retirement benefits had some major changes in the incentive structure in the estimation period and depend on previous earnings, which provide enough variation for identification for eligible.¹³ Less adjustment have taken place for old age pension. Since we assume that retirement is absorbing and associated with complete

¹¹ ATP is a small supplementary public regulated labour market pension.

¹² From the social security wealth formula it is also clear how eligibility to early retirement influence the social security wealth measure.

withdrawal from the labour force, old age pensions are not dependent on individual earnings. As we include year trend, gender and cohabitation status in the estimation, this assumption made it very difficult to identify the incentive effects from old age pension per se. Therefore we tried another avenue where we have computed old age pensions taking into account labour supply of spouse.¹⁴ The old age supplement is depending on household income and this provided enough variation to get a better fit of old age pension. The computation of social security wealth therefore includes spousal income.

In Table 2 the median social security wealth and peak value are shown by age group and eligibility to early retirement. For both “eligible” and “ineligible” social security wealth is falling with age, which reflects forgone early retirement benefit and disability pension and later old age pension. The level is quite different, which reflects that early retirement benefit paid at the maximum level is more generous than disability pensions. The same can be seen from the peak value. It is negative and incentives are stronger in early retirement. Note that the peak value is rising in absolute terms up to age 64-65 for “eligible”. Two things can explain this. First, the reduction in early retirement benefit after 2½ year and later after five years does not take place for those, who postpone retirement. Second, low wage earners retire earlier than high wage earners. For “ineligible” the peak value drops in absolute terms. This is because singles retire earlier than couples. Singles have a larger tax deduction and spousal income reduces old age pensions for couples. Moreover, it can be seen from the spread between tenth and 90th percentile that the variation in early retirement and disability/old age pension are not that different.

With the micro simulation model we can simulate the social security wealth of those ineligible to early retirement assuming they could draw early retirement benefit. This potential social security wealth can be used as and control variable in eligibility equation and therefore provide an additional exclusion restriction in the econometric model.

¹³ It is possible that earnings are correlated with preferences for leisure in that case we will overstate retirement incentives. We follow the literature and include as additional controls a polynomial in earnings to adjust for this.

¹⁴ We are aware that spouse labour supply is endogenous and should be modelled simultaneously with individuals. Here we assume it is exogenous and that spouse work until the age of 71.

6 Descriptive statistics

We look at figures combining retirement and eligibility with some of our exogenous explanatory variables. Previous analyses into the determinants of UI fund membership is very sparse (Parson et al.,2002). First, voluntary UI funds only exist in a few countries, for example also in Sweden and Finland where there are voluntary element to UI. By contrast, retirement is described in detail in many studies (for Denmark Bingley et al. 2002 is good reference).

Table 3 shows the descriptive statistics for variables other than the incentive measures across eligibility to early retirement status at age 59. First, 85 per cent of the sample is eligible to early retirement benefit, which is slightly more than the population. The reason is that we exclude academics and public employees. Second, we comment on the time invariant variables or those who changes slowly across time. The “eligibles” to “early retirement” constitute a more homogenous group than the “ineligibles”. This becomes apparent from comparing the two columns with the corresponding standard error. Mean real income (earnings) are approximately the same for both groups; but median real income for “ineligible” are 105,870 DKK and 76,100 for “eligible”. Apart from that, the sample is dominated by males with a spouse. Education is low in the sample. Only 6% (2%) have an education of 16 years for “ineligibles” (“eligibles”), while 57% have twelve to 14 years of education. Most individuals hold a full time job, in particular the “eligibles”. On one hand the “ineligible” are mainly living in the area of Copenhagen and are white collar workers and on the other hand the “eligible” are blue collar workers living out side Copenhagen. The majority of the sample is owner-occupiers without (taxable) debt. Moreover manufacturing and trade and hotels are the most import industry affiliations.

Industry and occupation are supposed to pick up union pressure of joining a UI fund, which are managed by the unions (Parson et al., 2002). Since industry and occupation are indicators of unemployment and health condition, which are known to be good predictors of retirement, it is clear that industry and occupation also directly influence retirement.

Overall retirement patterns by gender, as shown in Figure 2, are similar to the patterns in Figure 1: “eligibles” to early retirement retire much earlier than the “ineligible”. Women retire earlier than men. They also tend to have lower earnings, implying that their replacement rate is higher.

Figure 3 repeats the general pattern that “eligible” retire earlier than “ineligible”. For “eligible” retirement differs little across single and couples. Generally, “ineligible” singles retire earlier than couples.

Figure 4 shows that earnings are important for retirement. Individuals are classified as low wage earners if their earnings are less than the 1. quartile at age 59 and high wage earner if earnings are higher than the 3. quartile. Initially, at age 60, more than 50 per cent of the “eligible” with a low wage retires. First, low wage earners have a higher replacement rate and therefore higher incentives to retire earlier. Second, earnings are possibly correlated with preferences for leisure. The latter means that high wage earners have lower preferences for leisure (Coile et al., 2000).

We construct individual health shocks from public provided sickness benefit. This is only an indicator of health conditions and for certain employment groups the employer might pay full income in the period of sickness.¹⁵ The latter problem would be pronounced for public employees and academics, which are excluded from the analysis. We will expect that sickness benefit might be very differently distributed across eligibility. The reason is that disability pension requires a period of screening before disability pensions are paid. On one hand those eligible for early retirement can retire immediately in case of poor health and on the other hand those ineligible to early retirement benefit will have to go through the screening period. Table 4 shows that sickness benefits provide evidence of increasing retirement. Especially, for “ineligible” – as expected – receiving sickness benefits increase changes of getting a pension.

¹⁵ If the employer does not pay full income, he/she is typically required to pay two weeks of sickness benefit.

Table 5 indicates the role of individual labour demand shock. We provide the probability of retiring as a function of three types of shocks. The first is plant closure. The second, plant expansion is a dummy equal to one if numbers of new employees as a percent of ultimo employees in the firm multiplied with number of employees initial exceeds 20. This definition balance size of firm and growth rate. The third, plant downsizing is a dummy equal to one ‘if number of employees leaving as a percentage of initial number of employees multiplied with initial plant size’ exceeds 20. Again correction for plant size put more weights on larger firms. For closing plants and downsizing we see the expected rise in retirement rates as compared with other. Surprisingly expansion in a firm induces retirement for “ineligible” and only a minor change for “eligible”, but this is an artefact of the construction of the variable, which is seen as a percentage of ultimo size. If we control for plant downsizing is zero, the probabilities changes to 5.7 (18.3) for “ineligible” (“eligible”). The same can be said for downsizing if we control for plant expansion is zero.

7 Estimation results

In this section we take a look at the bivariate models result. Three models are estimated. First, a model is estimated without serial correlation in retirement equation and correlation between eligibility to early retirement and retirement, which is the same as eligibility is treated exogenous. This model is the Coile et al., 2000. The second model adds correlation between retirement equation and eligibility and a random effect. The third model extends model two with an ARMA(1,1) error. All three models are estimated with the same covariates.

The three models predict reasonably well the retirement outcome. The coefficients on the incentive measures are estimated with the correct sign in all three models and all control variables have their expected sign, which can be seen in Table 6. A Wald test for exogeneity of eligibility is provided by the square of the t-ratio on the coefficient for correlation between eligibility and retirement, $\rho_{\epsilon\zeta}$ in the bottom of Table 6. This is value is 51.4 (47.6) in model II (III) and we therefore reject exogeneity. It seems as if incentive measures are only slightly affected by accounting for unobserved

heterogeneity and serial correlation in the error term. However the age terms seem to have picked up most of the unobserved heterogeneity.

Before we plunge into a detailed discussion of coefficients and incentive effects, we shall discuss the model fit. We compare model outcomes with the observed retirement pattern. Prediction from model I to III is done by simulation. First, we draw an error for the eligibility equation and for the retirement sequence, $t = 1, \dots, T_i$ for individual i jointly.¹⁶ Second, for each individual we use the simulated errors to calculate retirement age. Third, we do one and two 15 times and average across simulations.

Figures 5a and 5b show the fit of model I to III for “eligible” and “ineligible”, respectively, where we use the simulation strategy above. The models predicts retirement pattern for the “eligible” reasonably good in Figure 5a. All three models are hard to distinguish. In general, retirement is slightly understated in the models. All models seem to have troubles at age 67. Figure 5b shows the fit for “ineligible”. The three models seem to follow each other quite closely. Model I overstates retirement at age 60, and model II and III understates retirement in the mid 60'es.

While coefficients in a probit model are hard to interpret except for their sign, Table 7 give the effect of a change in social security wealth of DKK 10,000 (1980) on retirement, where we have evaluated the probability at mean sample values for other control variables. We shall only comment on results from model III as only small differences occur across model I - III. The effect of social security wealth, DKK 10,000 (1980), is an increase of 0.43 percentage point in the retirement hazard. On the other hand an increase in peak value of DKK 10,000 reduces retirement with 7.3 percentage points a quite large effect. Another way to evaluate this measure is by the elasticity of non – participation, i.e. for social security wealth $\left[\frac{\partial \log P}{\partial \log ssw} \right]$, which are estimated to be 1.30 and 0.54 for social security wealth and peak value respectively. These estimates show a quite high response to social security wealth and the peakvalue. We have not been able to recover the elasticity of non-participation for other Danish studies; but can compare with US. In Coile and Gruber (2000) the

¹⁶ For model II and III we use rejection sampling to draw an error structure consistent with individual's eligibility status.

elasticities are 0.30 and 0.15 respectively. The main reason for the very large elasticities in our study is that we are analyzing a sample of very old¹⁷. To control for the effect of age on the elasticities we have computed the elasticity at age 59 (and all other controls at their mean values), which changes the elasticities to 0.75 and 0.31 respectively.

It is hard to think about the effects of social security wealth and peak value separately. Therefore, in the next section, we dwell more into the effects of pension incentive by simulation of some changes to the social security system.

Age is an important variable as we saw above. It influences the elasticity of non-participation. Moreover, it also reflects health, which is only included with the proxy variable sickness benefits. Accounting for intertemporal linkages can be quite important for interpretation of effects. From estimates of the three models in Table 6, we can see that age picks up unobserved heterogeneity. In Table 8 the increment in the probability of retiring from age t to age $t+1$ is given. As we can see the probability is positive for all age groups and reflects that the probability of retiring increases with age. The increase in model II and III is much stronger than in model I. Therefore accounting for these linkages can be quite important in simulations with the model.

A few other variables are also affected by the error structure in the model. First, health, which is interacted with eligibility, has clearly become more important. For “ineligible” health has the correct sign and is highly significant. While for “eligible” poor health has an insignificant effect. This of course reflects the option “eligible” has of retiring in case of poor health without going through a period on sickness pay. Labour market shocks are not affected by the error structure as expected. Perhaps a little surprising we find that males are much more likely to retire than females. But we condition on individuals are working at age 59. It is therefore difficult to compare this with other studies as all females retired before age 60 are excluded from the analysis. The remaining controls in the regression show more or less the expected sign.

¹⁷ The mean age in the Coile et al., 2000, sample is 58.5 compared with our sample with an average age of 60.9 years. Comparing coefficients across discrete choice models can be difficult as the marginal effects depend on other covariates.

These are the key results from our retirement equation. The main results from the eligibility to early retirement equation concern the exclusion restrictions. Industry, occupation and potential social security wealth are all highly predictive of eligibility (see Table 6, lower panel). The sign of coefficients are all of the expected sign. In Table 7 the effect of an 10,000 DKK increase in potential social security wealth is associated with a 0.88 percentage points change in probability of being “eligible” and translated to an elasticity it is as high as 0.74. The other covariates show the expected sign.

For investigating sensitivity of the model to some of our assumptions, we have also estimated models, which include industry and occupation in the retirement equation and this did not have an effect on the results. Moreover, we have estimated a version, which includes lagged unemployment for individuals and spouse retirement status and this did not change results either.

8 Policy simulation

In this section we simulate the outcome for the bivariate probit model with random effect and serial correlation from two different policy scenarios, model III. The first is similar to the reform in Bingley et al. (2002). We change the window of early retirement and old age pension by three years. The second policy scenario has to do with the 1999 reform of early retirement. The policy simulation looks at both the effect on retirement and on eligibility. The usual disclaimer applies that all variables and the error term are invariant to the policy simulation.

3 year reform

We have implemented the three year reform by abolishing early retirement and disability pensions for age group 60 to 62. Otherwise the benefit and tax rules applies as before. Duration of early retirement benefit is seven years and old age pension is available from age 70.

In Table 9 below the fictive reform increases retirement age by 2.2 years for “eligible” and 1.1 years for “ineligible”, the same reform in Bingley et al., 2002, only increased retirement by 1.3 years. The big difference can be attributed to an analysis of the very old in our paper. The reform shows that there is a potential large effect from abolishing early retirement benefit. The effect on “ineligible” will be smaller, as they did already retire later in the baseline.

We can access the effect of policy on eligibility status, which is not included in the estimates on retirement age. We see that eligibility drops with 7 percentage points due to the lower level of social security wealth. This would increase the average retirement age further. Note, in the discussion on policy reform it is sometimes argued that removing early retirement benefit just shifts the burden to a different retirement scheme. This is not possible here as disability pension is also delayed.

1999 reform

The reform in 1999 changed normal retirement from age 67 years to 65 years. Moreover, working to age 65 means a tax allowance –always given at age 65 - of 6 percent of maximum UI benefit for each of 3 months worked after age 62 for “eligible”. Moreover the reform introduced an annual contribution, which must be paid for at least 25 years to be eligible to early retirement. We have estimated the contribution to be DKK 40,000 (1980) and subtracted it from social security wealth. Retiring before age 62 the maximum early retirement benefit is 91% of maximum unemployment insurance and for individuals retiring after age 62 it is 100% of maximum unemployment insurance.¹⁸

The results of the simulation are given in Table 8 under policy II. As we can see that average retirement age is falling for “eligible” with 0.2 years and status quo for “ineligible”.

¹⁸ The last major element in the reform is a reduction in early retirement benefit for those who retire before age 62 and have private pension deposits. Even if the private pension plan is delayed the reduction on benefit will take place. We have excluded this.

First, the incentives to postpone retirement to age 62 has increased retirement age; but as the calculations show the income effect from early retirement at age 62 is quite high and actually induces more to retire at age 62. Moreover the tax allowance also has a substitution effect and income effect, which means that retirement at age 63 and 64 is quite high.

Second, for “ineligible” the status quo is because shifting the retirement age from age 67 to age 65, does not change the old age pension retirement incentives as we have assumed that individuals had access to disability pensions.

However, the largest effect of the 1999 reform we find on “eligibility”. The magnitude is almost as removing the early retirement benefit and disability pension for three years. This will in the long run rise average retirement age as well.

9 Conclusion

The key findings of the paper can be summarized as: We found that eligibility is endogenous. In the policy simulation some of the largest effects come from the eligibility equation. Unobserved heterogeneity had minor influence on the incentive measures and not as strong effects as in Börsch-Supan, 1999. But unobserved heterogeneity changed the importance of some of the other included variables, mainly those relating to health, i.e. age and the health proxy. A finding Börsch-Supan also did.

Literature

Angrist, Joshua (2000): “Estimation of limited dependent variable models with dummy endogenous regressors: simple strategies for empirical practice”, NBER technical working paper 248.

Arellano, Manuel and Jinyong Hahn (2005): “Understanding Bias in Nonlinear Panel Models: Some Recent developments”, Invited lecture, Econometric Society World Meeting, London

Bertschek and Lechner (1998): “Convenient Estimators for the Panel Probit Model”, *Journal of Econometrics*, Vol 87. pp. 329-371

Bingley, Paul; et al. (2002): “The Effects of Pension Programme Incentives on Retirement Behaviour in Denmark”, Centre for Labour Market and Social Research, Working Paper 01-08, March 2001, Revised March 2002

Börsch-Supan, Axel (1999): “Incentive effects of social security under an uncertain disability option”, in David Wise “Features in the economics of ageing”, University of Chicago Press, 2001

Carrasco, Raquel (2002): “Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the effect of Fertility on Female Labor Participation”, *Journal of Business and Economic Statistics*, Oct. 2001, 19, 4. pp. 385-394

Chan and Stevens (2004): “Do changes in pension incentives affect retirement? A longitudinal study of subjective retirement expectations”, *Journal of Public Economics* 88, pp. 1307-1333

Ejrnæs, Mette and Hochguertel, Stefan (2006): “Early retirement incentives, Selfemployment, and Voluntary Unemployment Insurance: Evidence from Administrative Data”, Draft May 2006.

Forlaget Idag: Various issues of “arbejdsløshedsforsikring” and “Social Pension”.

Gruber, Jonathan and David Wise (1998): “Social Security and Retirement around the world”, University of Chicago Press

Gruber, Jonathan and David Wise (2005): “Social Security programs and retirement around the world: Micro estimation”, University of Chicago Press

Hajivassiliou and Ruud (1994): “Classical Estimation Methods for LDV models using simulation”, *Handbook of Econometrics*, Vol IV.

Heckman, James (1978): “Dummy Endogenous Variables in a Simultaneous equation system”, *Econometrica* Vol. 46, no. 4, pp. 931-959

Krueger and Pischke (1992): “The effects of Social security on Labour supply: A cohort analysis of the notch generation”, *Journal of labor economics*, vol 10. issue 4. pp. 412-437

Maddala (1983): “Limited-dependent and qualitative variables in Econometrics”, *Econometric Society monographs* (NO. 3), Cambridge

Parsons, Donald; Tranæs, Torben and Lilleør, Helene Bie (2003): “Voluntary Public Unemployment Insurance”, Working Paper 03:25, Socialforskningsinstituttet.

Statistics Denmark (2000): *Statistiske efterretninger, Arbejdsmarked 2000:7.*

Velfærdskommissionen (2006) “Fremtidens velfærd – vores valg”;

Verbeek and Nijman (1992): “Testing for selectivity Bias in panel data models”, *International Economic Review*, 33, pp. 681-703

Økonomiske Råd, Det (2005) “Dansk Økonomi, forår 2005”

Appendix A: Incentive measures

We compute the peak value of retiring from the early retirement benefit scheme and disability and old age pension. A calculation of early retirement benefits requires the knowledge of gross earnings.¹⁹ Therefore, we need to estimate earnings to project future earnings. This is done in two steps.

In the first step we correct any earnings in a given year for periods of unemployment. Correction for unemployment is done in the following way. If we observe less than 6 months of unemployment, we simply divide earnings with one minus the unemployment rate. For individuals with more than six months of unemployment we use lagged earnings from a period with less than 6 months of unemployment. Second, we project future annual earnings in a very simple way assuming a real growth rate of 1 percent. Coile et al., 2000, use this assumption.

In calculating social security wealth we use actual rules from the social security system and the tax system from 1980 to 1996. This data is collected from the Danish tax authorities and a series of publication from “Forlaget Idag”. In very rough terms the tax system can be described as a flat local government tax of around 30% with minor variation across municipalities and a progressive state tax system with a major reform in 1987 and again in 1994. Early retirement benefit depends on previous earnings up to a maximum limit. This limit was approximately 84,000 in 1980. The median worker has a replacement rate of 0.82 in our sample at age 59. The level of old age insurance was 30,000 DKK in 1980. The disability system is modelled as old age pension, which is comparable if disability pensions are provided for social reasons. Different supplements exist depending on the degree of lost work capacity in case of a disability.

We impute the present value of social security wealth by the following formula for an individual of age t retiring at age R . Then we have myopic expectations with respect to policy changes; the formula changes from year to year based on the social security system in a given year.

¹⁹ In calculation of disability and old age pensions we use spouse income if a working spouse is present. The same projection of earnings is used up to age 71 for the spouse.

$$SSW_t(R) = \sum_{s=R}^T (1+D)^{-(s-t)} \{ \rho_{s,t} I(s \geq 60) I(e=1) Erb_{R,s} + \rho_{s,t} I(s < 67) I(e=0) Oap_{s,t} + \rho_{s,t} I(s \geq 67) Oap_s \}$$

where we have suppressed dependence on i individual and τ year. T is time at certainty of death (age 100). It is discounted by discount rate D and a gender and age specific survival probability $\rho_{s,t}$ of being alive at age s conditional on survival to age t . e denotes eligibility and Erb and Oap are the after tax benefit levels.

For the regression we use as explanatory variables the current social security wealth $SSW_t(R=t)$ and the peak value:

$$Peak_t = \max_j (SSW_t(j)) - SSW_t(t), j > t$$

Appendix B: Tables

Table 1: Median annual income in retirement, 1980 DKK

	Retirement income	Other income
Age < 67	66,232.3	6,640.1
Age >=67	29,861.2	17,632.2
Number of observations: 4702		

Note: Retirement in year t and income in year t+1.

Source: IDA database

Table 2: Incentive measure, social security wealth and peak value, (1980 DKK)

Age		Social security wealth		Peak value	
		Ineligible	Eligible	Ineligible	Eligible
59	Median	622543	693462	-12798	-3567
	1. decile	502524	625175	-14821	-11085
	9. decile	833157	881544	-11159	4645
60	Median	605231	670143	-13376	-24427
	1. decile	476873	608945	-19046	-27209
	9. decile	791029	849676	-11686	-20629
61	Median	557645	625989	-13254	-24120
	1. decile	451301	567436	-18430	-26776
	9. decile	750734	795836	-11309	-20592
62	Median	515430	583515	-13109	-24279
	1. decile	427382	525901	-19188	-26906
	9. decile	839358	742498	-11153	-21575
63	Median	485873	547201	-13014	-24091
	1. decile	406308	488487	-17877	-26310
	9. decile	674362	693516	-10877	-21447
64	Median	455218	507601	-12709	-27018
	1. decile	383884	452332	-16684	-28663
	9. decile	604988	654897	-10544	-24218
65	Median	429163	466573	-12599	-30643
	1. decile	362568	415334	-16715	-32074
	9. decile	583202	722920	-10445	-25757
66	Median	405995	424152	-12058	-29660
	1. decile	346246	374867	-15906	-32521
	9. decile	571438	590734	-10459	-23220
67	Median	406262	382221	-12187	-11897
	1. decile	329728	329728	-15023	-15340
	9. decile	539710	539710	-10157	-10216
68	Median	367987	359982	-11649	-11620
	1. decile	309927	313711	-15020	-14989
	9. decile	491311	509061	-9883	-9903
69	Median	345676	345676	-11215	-11215
	1. decile	291023	294713	-14176	-17820
	9. decile	479829	584124	-9191	-9585
70	Median	330186	324427	-11170	-10893
	1. decile	275750	316558	-17835	-17614
	9. decile	548967	548967	-9091	-10436
71	Median	300762	304128	-10245	-10257
	1. decile	259169	297395	-16888	-17424
	9. decile	515764	528780	-8917	-10232

Table 3: Descriptive statistics at age 59 by eligibility status

	Ineligible				Eligible			
	Mean	std. Err.	min	max	Mean	std. Err.	min	max
Log(income)	11,45	0,03	9,02	13,67	11,50	0,01	9,07	13,46
Male	0,62	0,02	0	1	0,74	0,01	0	1
Single	0,18	0,01	0	1	0,16	0,01	0	1
School: 12-14 years	0,57	0,02	0	1	0,57	0,01	0	1
School: 16 years	0,06	0,01	0	1	0,02	0,00	0	1
Poor Health	0,01	0,00	0	1	0,06	0,00	0	1
Bad Health	0,02	0,01	0	1	0,01	0,00	0	1
Part time job	0,24	0,02	0	1	0,08	0,00	0	1
Living in Copenhagen area	0,50	0,02	0	1	0,35	0,01	0	1
White collar	0,60	0,02	0	1	0,41	0,01	0	1
Year	89,62	0,15	84	96	90,42	0,06	84	96
Neagtive taxable wealth	0,21	0,02	0	1	0,16	0,01	0	1
Homeowner	0,51	0,02	0	1	0,61	0,01	0	1
Downsizing plant	0,81	0,02	0	1	0,81	0,01	0	1
Closure	0,03	0,01	0	1	0,04	0,00	0	1
Expanding plant	0,80	0,02	0	1	0,79	0,01	0	1
Potential experience	0,74	0,01	0	1	0,81	0,00	0,04	1
Unemployment degree	0,00	0,00	-0,13	0,12	0,00	0,00	-0,12	0,31
Manufacturing	0,30	0,02	0	1	0,51	0,01	0	1
Utilities	0,03	0,01	0	1	0,02	0,00	0	1
Construction	0,07	0,01	0	1	0,11	0,01	0	1
Trade and Hotel	0,28	0,02	0	1	0,23	0,01	0	1
Financial intermediation	0,25	0,02	0	1	0,10	0,00	0	1
Private services	0,01	0,00	0	1	0,01	0,00	0	1
Proportion retiring	0,45				0,75			
Mean retirement age	62,74	0,14	60	71	61,53	0,04	60	71
Number of observations	669				3817			

Table 4: Health and retirement probability

	Ineligible	Eligible
No sickness benefit	0,08	0,23
Sickness benefit < 20,000 DKK	0,32	0,23
Sickness benefit > 20,000 DKK	0,64	0,32

Source: IDA database and own calculations

Table 5: Labour market shocks and retirement probability

	Ineligible	Eligible
Plant closes	0,27	0,35
Plant expansion	0,15	0,33
Plant downsizing	0,16	0,34
Other	0,10	0,30

Source: IDA database and own calculations

Table 6: Results from bivariate probit model

Retirement equation	Model I		Model II		Model III	
	Coefficient	T-stat	Coefficient	T-stat	Coefficient	T-stat
Peak value	-5,91	-25,65	-5,70	-23,27	-5,67	-22,60
SSW	2,42	7,27	2,16	5,68	2,28	5,55
Age	-0,02	-0,05	1,68	3,64	2,07	3,05
Age2	2,50	5,92	1,65	3,53	1,25	1,74
Log income	-0,51	-13,81	-0,54	-12,35	-0,54	-12,16
Male	0,65	9,09	0,62	7,31	0,64	7,01
Single	-0,31	-4,91	-0,23	-3,10	-0,25	-3,18
12 to 14 years of education	-0,02	-0,53	-0,04	-1,08	-0,04	-1,06
16 years of education	-0,16	-1,84	-0,26	-2,43	-0,28	-2,50
0 < sickness benefit < 20,000	0,51	3,06	0,71	3,93	0,71	3,86
sickness benefit > 20,000	1,45	7,32	1,65	7,91	1,65	7,89
(0 < sickness benefit < 20,000)*eligibility	-0,45	-2,47	-0,73	-3,74	-0,73	-3,66
(sickness benefit > 20,000)*eligibility	-1,17	-4,81	-1,53	-5,99	-1,53	-5,94
Part time job	-0,21	-3,90	-0,23	-3,80	-0,23	-3,73
Living in area of Copenhagen	-0,10	-3,41	-0,16	-4,52	-0,17	-4,52
Year	-0,11	-1,64	-0,05	-0,68	-0,05	-0,67
Negative wealth	-0,20	-4,56	-0,24	-4,76	-0,24	-4,74
Home owner	-0,08	-2,65	-0,11	-2,84	-0,11	-2,85
Firm contraction	0,25	4,53	0,27	4,53	0,27	4,48
Firm close	0,38	5,79	0,39	5,76	0,39	5,68
Firm expansion	-0,37	-6,91	-0,39	-7,03	-0,40	-7,02
Experience	-0,63	-8,17	-0,88	-8,46	-0,90	-7,78
Unemployment by industry	0,96	1,72	0,79	1,30	0,75	1,21
Constant	2,73	6,55	3,39	7,10	3,29	6,65
Number of observations	15.580		15.580		15.580	

Table 6 (continued)

Eligibility equation	Model I		Model II		Model III	
	Coefficient	T-stat	Coefficient	T-stat	Coefficient	T-stat
SSW	4,68	7,29	4,62	7,23	4,60	7,20
log(income)	-0,43	-5,40	-0,41	-5,15	-0,41	-5,14
Male	0,88	5,64	0,92	5,90	0,92	5,88
Single	-0,84	-6,79	-0,82	-6,65	-0,82	-6,63
12 to 14 years of education	-0,09	-1,60	-0,10	-1,70	-0,10	-1,70
years of education	-0,47	-3,57	-0,47	-3,58	-0,47	-3,58
0 < sickness benefit <= 20,000	0,41	2,60	0,30	1,92	0,30	1,92
sickness benefit > 20,000	-0,55	-3,58	-0,83	-4,02	-0,83	-4,02
Part time job	-0,60	-6,02	-0,60	-6,04	-0,60	-6,05
Living in area of Copenhagen	-0,21	-3,93	-0,22	-4,17	-0,22	-4,19
Manufacturing	1,06	7,92	0,92	6,84	0,92	6,82
Utilities	0,58	3,02	0,49	2,60	0,49	2,58
Construction	0,96	6,42	0,84	5,57	0,83	5,54
Trade and hotels	0,90	6,55	0,80	5,84	0,80	5,82
Financial intermediation	0,50	3,47	0,38	2,63	0,38	2,63
Private services	0,82	2,96	0,69	2,51	0,70	2,52
White collar	-0,23	-3,89	-0,24	-4,10	-0,24	-4,10
Year	0,50	4,67	0,49	4,62	0,49	4,63
Negative wealth	-0,17	-2,57	-0,18	-2,67	-0,18	-2,67
Unemployment by industry	0,40	0,38	0,61	0,59	0,61	0,59
Home owner	0,00	-0,01	-0,01	-0,23	-0,01	-0,23
Firm contraction	0,03	0,33	0,01	0,08	0,01	0,07
Firm close	0,22	1,51	0,18	1,24	0,18	1,25
Firm expansion	-0,16	-1,55	-0,13	-1,30	-0,13	-1,38
Experience	0,42	3,05	0,23	1,61	0,23	1,60
Constant	1,24	1,64	1,33	1,76	1,34	1,77
Number of observations	4.486		4.486		4.486	
Loglikelihood	-10.007		-9.952		-9.932	
Covariance parameters						
σ_v			0,33	6,19	0,11	0,61
$\rho_{\varepsilon\zeta}$			0,24	7,17	0,24	6,90
ρ					0,61	4,52
θ					-0,38	-3,88

Note: Model I: covariance matrix is identity matrix,

Model II: a random effect and correlation of error in eligibility and retirement equation

Model III: random effect, correlation of error in eligibility and retirement equation and ARMA(1,1) in retirement equation

Table 7: Incentive effects

Retirement equation		Model I	Model II	Model III
Peak value	Dkk 10,000	6,81%	7,31%	7,27%
	Elasticity	0,52	0,54	0,54
Social security wealth	Dkk 10,000	0,43%	0,41%	0,43%
	Elasticity	1,28	1,23	1,30
Eligibility equation				
Potential social security wealth				
	Dkk 10,000	0,90%	0,88%	0,88%
	Elasticity	0,75	0,74	0,74

Table 8: Age effects

	Model I	Model II	Model III
60	0	0,02	0,02
61	0,01	0,03	0,03
62	0,01	0,04	0,05
63	0,02	0,06	0,06
64	0,03	0,08	0,08
65	0,05	0,10	0,10
66	0,07	0,11	0,11
67	0,09	0,12	0,12
68	0,12	0,12	0,11
69	0,13	0,10	0,10
70	0,13	0,08	0,07
71	0,11	0,05	0,05

Table 9: Policy scenarios

	Average retirement age	
	Eligible	Ineligible
Actual	61,1	63,6
Baseline	61,1	63,9
Policy I	63,3	65,0
Policy II	60,9	63,9

	Proportion eligible
Actual	0,85
Baseline	0,86
Policy I	0,79
Policy II	0,80

Note: Average retirement age is simulated conditional on eligibility

Appendix C: Figures

Figure 1: Retirement hazard by eligibility

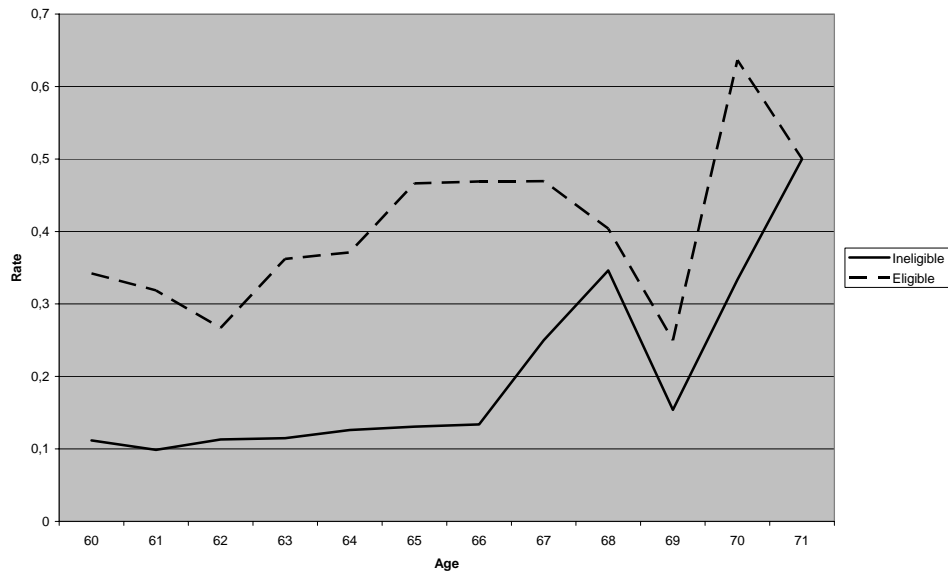


Figure 2: Retirement and gender

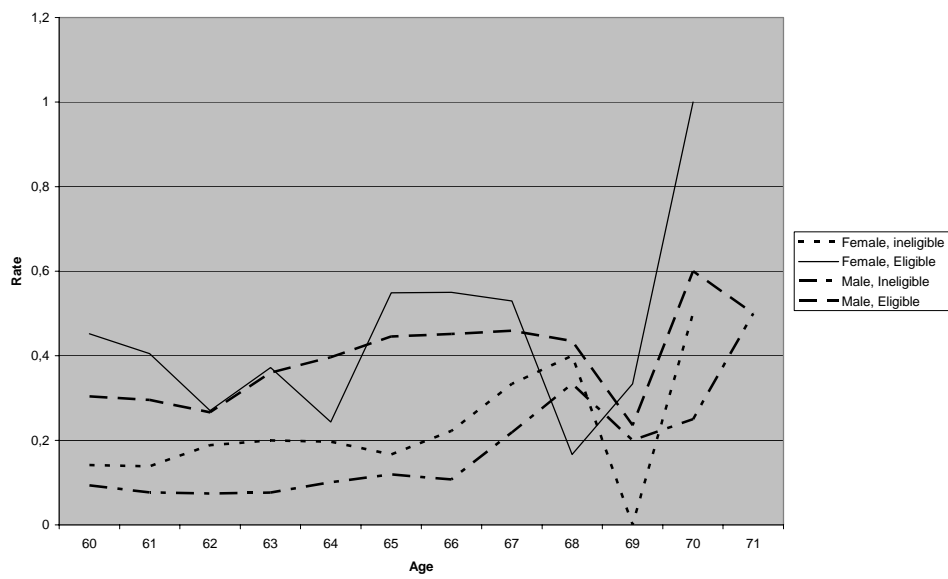


Figure 3: Retirement and Single status

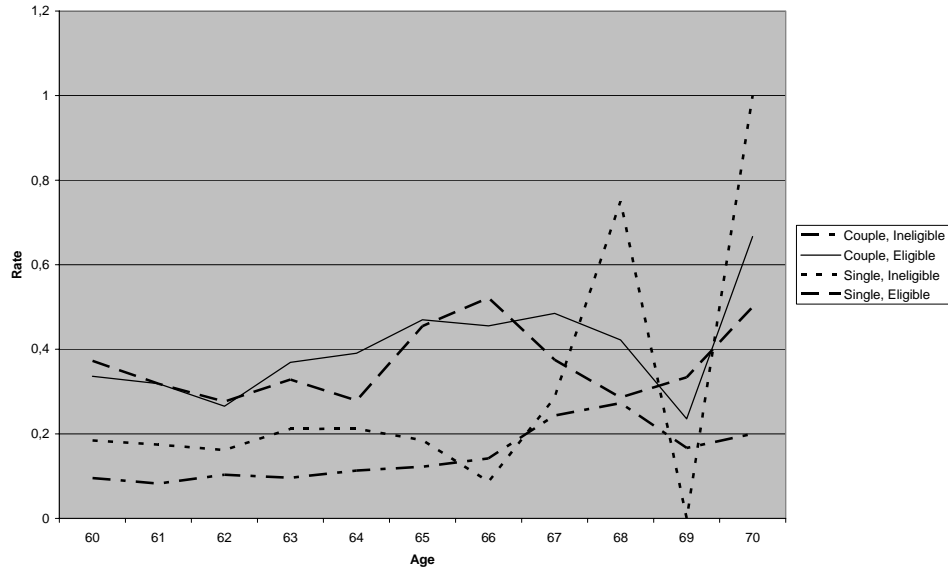
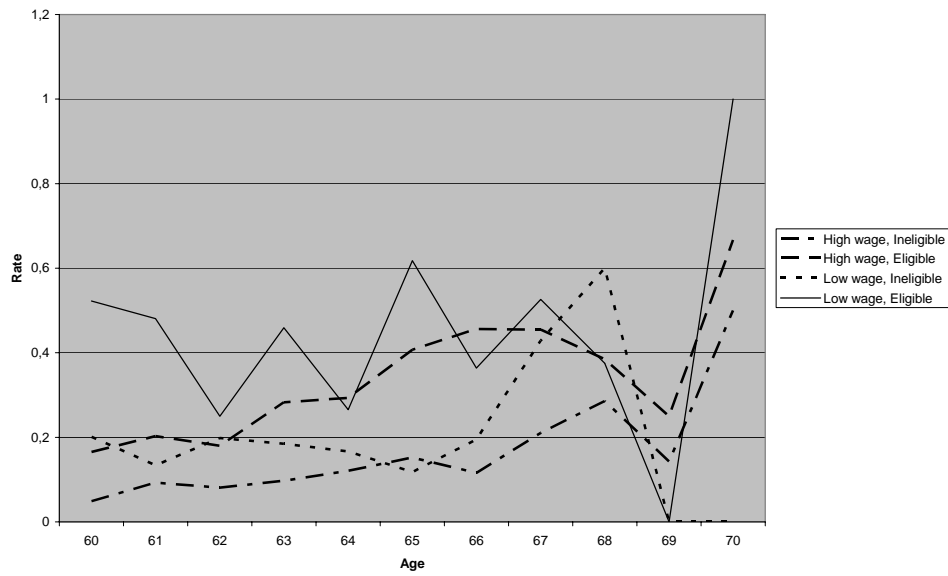
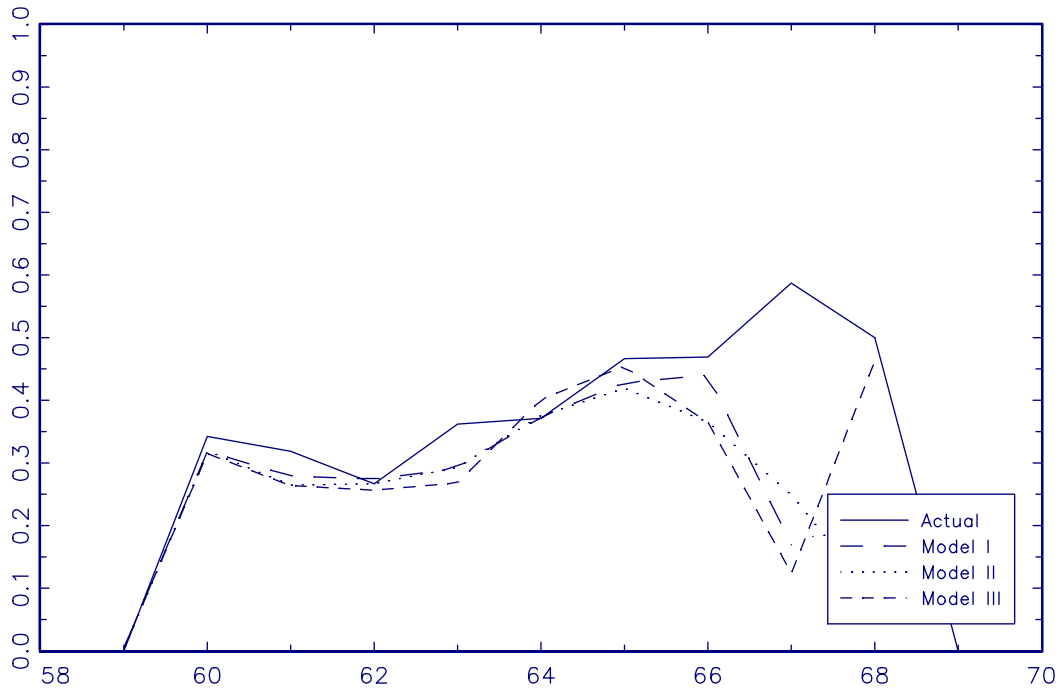


Figure 4: Retirement and wages



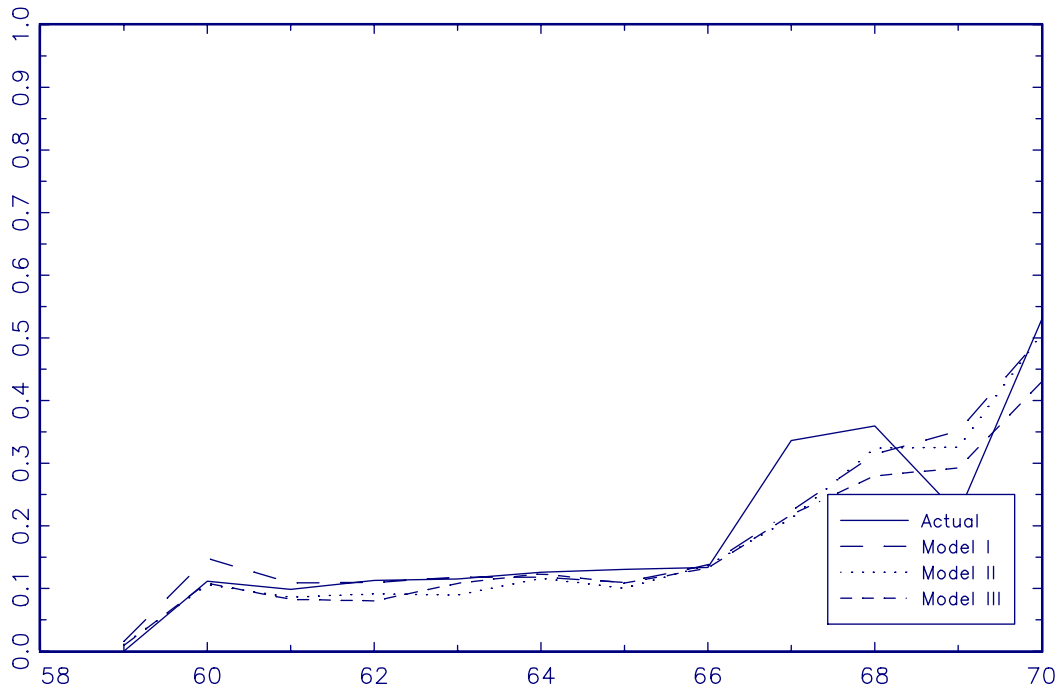
GAUSS Mon Mar 05 08:29:46 2007

Figure 5a: Fit to retirement pattern
Eligible



GAUSS Mon Mar 05 08:30:48 2007

Figure 5b: Fit to retirement pattern
Ineligible



The Interaction of Long-term Unemployment and Retirement

Martin Junge¹

Centre for Economic and Business Research, Copenhagen Business School

Department of Economics, University of Copenhagen

1. Version

Abstract: In this paper we model and estimate a discrete choice model of retirement and long-term unemployment with a sample of older workers in Denmark. Unemployment is a common pre-retirement state and around 25 per cent of transitions to retirement at age 60 are from (long-term) unemployment. Despite the large amount of unemployment prior to retirement it is surprising that unemployment is not more widespread as the unemployment insurance scheme is more generous than the early retirement scheme. We argue here that only if the unemployed are imposed large search cost we can explain this fact. The results of the paper points at very large search cost for the unemployed, though search also reduces cost of getting a new job. Simulation with the model shows that unemployment related retirement schemes can be extremely popular even at low economic incentives. Moreover, we compute individual's willingness to pay for an unemployment scheme without search cost. For the unemployed, we find, they are willing to live close to subsistence level.

Keywords: Unemployment, retirement, social insurance, dynamic programming, incomplete control.

¹ The Author would like to thank Martin Browning, Ulrich Kaiser, Hans Christian Kongsted and Nikolaj Malchow-Møller for valuable comments.

1. Introduction

In this paper, we are modeling and estimating the joint process of exiting from the labour market through unemployment and retirement. The model is a dynamic discrete choice model with imperfect control. It is estimated using maximum likelihood on a sample of workers, whose main retirement income is public provided support. Individuals maximize remaining lifetime utility. In each period, they control their labour market state by choosing between unemployment, retirement or full time work. The control is imperfect in the sense that the decision to stay in a given job can be overruled by a layoff, which is modeled as a function of creative destruction in plants. Utility is derived from income and leisure and includes taste shifters. The value of leisure varies across unemployment and retirement due to the absence of search requirements in retirement. The model incorporates the social insurance system in detail, which induces dynamic incentives, e.g., duration dependent benefits and an interaction between the unemployment benefit level and early retirement benefits. Moreover wages depend on the history of labour supply, general human capital (experience) and firm specific human capital (tenure).

A state variable – unobserved to econometrician – enters the model, which allows us to derive conditional choice probabilities. A layer of complication exists, as we have to integrate out the unobserved layoff outcome in the likelihood function. This leads to a mixing distribution for those who separate from their job. Either they were quitting or were laid off. Estimates of parameters are within plausible range of other studies; though they appear to be somewhat imprecisely estimated. Policy simulation reveals that early retirement distorts incentives heavily.

Public support for early retirement is widespread in many countries. It takes the form of either disability pensions, unemployment related pensions or, in some countries, specific early retirement schemes. These schemes were quite common from 1980's to the 1990's (see Blöndal and Scarpetta, 1999). As the pool of unemployed has started to decrease and projections of the dependency ratio well into the 21st century shows a decline, many governments have started to roll back these schemes (Zaidi, 2006). The focus of this paper is on the interaction between unemployment and retirement and the incentives

created from public support. This is highly relevant for unemployment related pension schemes. As many early retirement schemes are made less generous, the future role of unemployment insurance as an alternative exit route from the labour market can become much more important than we see today.

Denmark is no exception in this history of Europe. Three points illustrate this. First, a generous early retirement scheme from age 60 (“*efterløn*”) was introduced in 1979. The scheme was intended for workers with a life-long service in the labour market. Eligibility was gained through long-term membership of an unemployment insurance fund^{2,3}. Though the benefits in this scheme are less generous than unemployment insurance benefits, it was right from the start extremely popular. Figure 1 presents the transition from work to retirement and from work to unemployment for private sector employees, who are eligible for early retirement benefits and who have not retired on disability pensions, (aged 50 to 55 in 1980, and with low tenure). Up to age 60 modest unemployment transitions exist. At age 58 and 59 somewhat larger transitions occur up to 5 per cent. Retirement is almost non-existing before age 60. At age 60 early retirements benefit can be drawn which can be seen by the dramatic increase in the retirement rate.

Second, in 1992 a very early retirement programme was introduced for long-term unemployed (“*overgangsydelsen*”). It was introduced for the age group 55 – 59 years and later extended to the age group 50 to 59 years. It quickly became very popular, and in 1996 the government abolished the programme as it became too expensive and because the Danish economy was experiencing a boom. The model in this paper will be useful for measuring the incentive effects of this type of unemployment related pension programmes.

Finally, today the government is making it less generous to exit through early retirement in order to reduce the inflow. A reform in 1999 thus made it more expensive to join the

² UI funds are private in Denmark and membership is voluntary. It is organized in trades. A fund exists for all private sector workers.

³ Membership was required five years of the previous ten years to be eligible in 1980. This has been strengthen many times since then and today 25 years of membership out of the previous 30 years is required.

early retirement scheme and introduced tax related subsidies for individuals postponing retirement. Moreover, today the government has secured support for a new reform making early retirement even less generous.

The model in this paper will try to predict the observed pattern in Figure 1 with a structural model of (long-term) unemployment and retirement. The problem is that unemployment insurance is more generous than early retirement benefit. And from Figure 1 it is clear that early retirement benefits are having much more distorting effects on retirement behaviour than UI benefits. Explanations for the spike at age 60 like mandatory retirement age, health related explanations and alternative income source (i.e. private pension plans) are unlikely to be important here. Mandatory retirement is 'officially' not very common in Denmark for these workers. Moreover, the sample consists of low tenure workers, who are not likely to have special retirement contracts with their employer. Health is likely to change gradually over time and not as abrupt as in Figure 1.

Instead, we argue that a model with differences in the valuation of leisure could be the explanation. Leisure will be influenced by job search requirements or whether sanctions and activation in the UI system are effective.

The previous literature on retirement has made it clear that unemployment is an important pre-retirement state. The literature describing the transitions to retirement thus finds unemployment to retirement transitions accounts for as much as 40 percent of transitions into retirements in Japan (Oishi et al., 2005) and more than 20 percent in Germany and France (Börsch-Supan et al., 2005 and Mahieu et al., 2005). These studies rely on income definitions to define the unemployment and retirement states. For comparison 25 percent of those retiring at age 60 were long-term unemployed, who would probably be classified as unemployed by income, in our estimation sample.

Structural models that account for both unemployment and retirement decisions are very rare in the literature. We are aware of one structural retirement model where unemployment and the incentives of UI benefit are modeled (Heyma, 2004) jointly with other exit routes from the labour market. But the model treats unemployment as an absorbing state and some of the issues like search cost and loss of human capital cannot be studied. The focus in his paper is not on unemployment as a pre-retirement state but as a retirement state, which is very different from what we do. In the mainly US dominated structural retirement literature long-term unemployment is a non-work state and equivalent to retirement. Some examples are Rust et al., 1997, who define retirement as unemployment/out of labour force. They consider applying for social security and Medicare, but not UI benefit. In Van der Klaauw et al., 2006, no unemployment exists. Unemployment hides in no-work state and part time work. Berkovec et al., 1991, convert short unemployment states where individuals return to previous employer to work and long-term unemployment are considered as retirement. French et al., 2004, model work as more than 300 hours a year and long-term unemployed gets classified as retired.

The focus of these papers has been on retirement and not unemployment. However, we have been convinced by this work that health, wealth and labor market opportunities are important for retirement. Moreover social security, private pension plans and Medicare (US related literature) are also extremely important predictors of retirement.

Another difference between the present paper and the existing literature is that the structural retirement models focus on the case of complete control. Exogenous shocks from e.g. labor demand are not modeled. Again, Heyma, 2004, makes an exception. He allows for involuntary quits by modeling layoff probabilities. Individuals in his model take the layoff probability into account when they decide on their optimal retirement path and can involuntarily retire if they are laid off. In Heyma layoffs are observed from a questionnaire, if individuals are subsequently unemployed or retired. We model this incomplete control case quite differently. In this paper we model layoffs under the perhaps more common case that it is only partially observed. Moreover, individuals can choose to continue in a new job in this paper.

Another literature relevant for this paper is that of search models. Note our model contains a search element as individuals can change job and be laid off. The description of search for work either from unemployment, work or somewhere else has received a very thorough treatment in the literature. On the other hand retirement receives relatively little attention, mainly assuming that retirement takes place in a distant future (Eckstein et al., 2006). In their survey Eckstein et al. point at Berkovec et al., 1991, as a search model including retirement and social security. As already discussed this model does not include unemployment but allows for transition between fulltime work, part time work and retirement (non-employment).

This paper contributes to the literature by considering unemployment and retirement decision jointly. We incorporate some of the search elements into a standard structural retirement model. First, job changes are allowed. A job change can either be a layoff or it can be voluntary. As job changes are associated with a loss of firm specific human capital, which leads to a drop in wages, individuals will change job because of non-pecuniary shocks to their existing job or the new job offer. Second, unemployment will be chosen either voluntarily or involuntarily in case of a layoff. Third, job offers are given every period; but there is a cost of leaving unemployment, current job or retirement, which can be interpreted as the quality of job offers. Fourth, retirement and unemployment benefit, duration and eligibility rules are modeled in detail and provide dynamic incentives, which are influencing choices.

The paper is structured as follows: First Section 2 updates the reader on the institutional setting in Denmark from 1980-1996, which is our sample period. In Section 3, we formulate a dynamic discrete choice model of retirement and unemployment. The econometric specification and estimation is outlined in Section 4. Section 5 describes the Danish administrative data set that we use. Results and measures of fit are provided in Section 6. And finally before we conclude in Section 8, we use the estimated model for two experiments in Section 7.

2. Institutional setting

We have already touched on the Danish system in the introduction. This section gives further details. The discussion relates to cohorts born in 1925 to 1930.

We first take up public support for old-age-pension and early retirement. Old-age-pension is provided universally and is only based on long-term residence in Denmark. Benefits can be drawn from age 67 for individuals born 1925-30 and consist of two tiers: a base level tested against labor market income and a supplement tested against household income.

Early retirement was introduced in 1979⁴ for workers with a long service life. Eligibility is obtained through long-term membership of UI funds, which is voluntary in Denmark. Initial requirements, which are the rules applying to our sample, required five years membership out of the previous ten years. This requirement has been strengthened several times since then. Individuals should also be eligible for UI benefit at the time of retirement (see below). The benefit level is 90 percent of the previous wage up to a maximum. After 2½ years on maximum benefits the level is reduced to 80 percent and after another period of 2½ years to 70%. Re-entry to early retirement after exit was not possible. This was changed in 1993 where individuals could leave for work once and then come back. Moreover, 200 hours of work per year were allowed without reduction in benefits. Benefits do not depend on other income sources in the household.

UI benefits are available for members of UI fund. 6 months of work within the previous 36 months was necessary to be eligible to UI benefits. The benefit level is 90 percent of previous wages up to a maximum.

Unemployed are supposed to be available for work. In many aspects the unemployment state reminds of the work state. They have to document they are searching actively for a job. They have to note the job centre if they take vacation, which is of the same length as

⁴ Before 1979 alternative retirement routes existed, including widow pension and disability pension.

workers. Job offers have to be accepted otherwise they can be drawn in their benefits. It is important to emphasize that these are imposed on older workers, which is often not the case in other countries. Though anecdotal evidence points at some loose administration of rules in the estimation period because of high unemployment rates.

Active labor market policy was mainly concentrated on maintaining another UI period for individuals. Thus a job offer was given before individuals exhausted their UI benefits (see The Economic Council (“Det Økonomiske Råd”), 2005).

3. Model

The model is cast in the standard consumption-leisure framework. The model represents the decision of an individual and maximizes expected remaining lifetime utility from middle age over consumption and leisure bundles. While this framework is very general we need to specify a utility function, a state space and uncertainty. We discuss the properties of our specification and transform the problem of the individual into a form that can be solved by backward recursion inside a likelihood function. We return to the latter aspect in the next section.

The standard consumption-leisure framework

The standard framework assumes that individuals maximize lifetime utility, which is obtained from consumption and leisure in each period, given initial endowments. We assume that utility is additively separable over time and that individuals discount the future. The model focuses on older workers, initially of age 50-55, and the problem of maximizing remaining lifetime utility conditional on history. Mainly due to data limitations we focus on the case where consumption equals income⁵ or, alternatively,

⁵ The data contain taxable wealth, which account for the main wealth holdings of individuals. Saving can be backed out from changes in wealth, and a correction to disposable income, which needs to be computed from taxable income, would then give us consumption. However, computing disposable income from taxable income requires much more information than is available in our data. See Browning et al. (2003) for a successful attempt and some of the associated problems, where they use a much more comprehensive data set on taxable income components than ours. Rust et al. (1997) rationalize the consumption-equal-

where utility is derived from income. Moreover, only discrete decisions can be taken with respect to labor supply. Individuals can work full time, be (long-term) unemployed or retire.

So we are looking for a sequence of discrete labor supply choices to maximize remaining lifetime utility. The value function V , which is total utility at the optimal sequence of discrete choices, can be written:

$$(1) \quad V_a(x_a) = \max_{\{h_t\}_{t=a}^{T^*}} E_a \left[\sum_{t=a}^{T^*} u_t(y(h_t, x_t), h_t, x_t) \mid x_a \right]$$

$$s.t. \quad x_{t+1} = f(x_t, h_t)$$

where a is initial age, u_t is instantaneous utility, h_t is the discrete decision of choosing unemployment, retirement, or full time work. T is age 100 and T^* is age 74, the last period where labour market decisions take place. The decision set varies with age. From age 50 to 66 all three states can be chosen. And since no unemployment program is available after age 66, the choice stands between retirement and full time work. Individuals can get laid off in the beginning of a period, and hence labor supply is an incomplete control variable. After a layoff, individuals can choose between a new full time job, unemployment or retirement. y is income derived from their choice, x contains state variables, influencing utility and uncertainty. The state variables are discussed below but include: Tenure in current job, labor market experience, gender, single status, age, year⁶, a risk of being fired in the current job, length of current unemployment spell, length of current early retirement spell, eligibility criteria for early retirement, and an unobserved state variable, which is known to the individual in the current period and unknown in future periods. Individuals maximize expected utility, where they have to account for the uncertain events of death or - if working - getting fired. Expectations at

income assumption by focusing on low income individuals and using an argument of incomplete markets. In Deaton, 1991, individuals with non-stationary income processes and liquidity constraint find it optimal to set consumption equal to income.

⁶ Year is taken into account here because it represents the parameters of the tax and social insurance system in the given year. Individuals are assumed to be myopic with respect to these parameters, and they therefore think current policy parameters are unchanged in the future.

age a are conditional on initial state x_a . The state variables evolve according to a function f , the transition equation, which is specified below.

We can write the problem in a recursive form where we explicitly solve for death and layoff uncertainty. Let $x_t = (\bar{x}_t, m_t)$ where m_t is tenure and \bar{x}_t contains the remaining state variables.

(2)

$$\max_{h_a} \left\{ \begin{aligned} & u(y(h_a, x_a), h_a, x_a) + \beta \eta_{a+1|a} \pi(x_a, h_a) E_a \max_{\{h_t\}_{a+1}^T} \left(E_{a+1} \left[\sum_{t=a+1}^T u_t(y(h_t, x_t), h_t, x_t) \mid \bar{x}_{a+1}, m_{a+1} = 0 \right] \right) + \\ & \beta \eta_{a+1|a} (1 - \pi(x_a, h_a)) E_a \max_{\{h_t\}_{a+1}^T} \left(E_{a+1} \left[\sum_{t=a+1}^T u_t(y(h_t, x_t), h_t, x_t) \mid x_{a+1} \right] \right) \end{aligned} \right\}$$

where β is the discount factor, which is positive and less than 1. $\eta_{a+1|a}$ is the survival rate to age $a+1$ conditional on survival to age a . π is the probability of getting fired, which is a function of creative destruction at the workplace for the individual if he/she is working or the potential workplace if they are not working. Below we discuss the precise details of these functions. In case of a death ($1 - \eta_{a+1|a}$) future utility become zero. The maximization problem is broken down into two parts: First, the solution to the next period problem and second, maximization of current utility taking into account the effect on future utility. The layoff probability, π , resets the tenure state variable to zero. Note that we have explicitly written the expectation operator as of age a . It is easy to see from (1) and (2) that we can replace the second and third maximization operator with a value function as of age $a+1$, and write the model in the dynamic programming form:

(3)

$$V_a(x_a) = \max_{h_a} \left\{ u(y(h_a), h_a, x_a) + \beta \eta_{a+1|a} \pi(x_a, h_a) E_t [V_{a+1}(\bar{x}_{a+1}, m_{a+1} = 0)] + \beta \eta_{a+1|a} (1 - \pi(x_a, h_a)) E_t [V_{a+1}(x_{a+1})] \right\}$$

This is a system of equations, one for each age from a to T . This can be solved by backwards recursion subject to the transition equations. In period T no choice takes so the value function is simply

$$(3a) V_T(X_T) = U(X_T, h = R)$$

where R denotes the retirement state. The solution at age t between T^*+1 and T is,

$$(3b) V_t(X_t) = U(X_t, h = R) + \beta \eta_{t+1|t} E_t [V(X_{t+1})]$$

though there is no observable behavior in these equations, the utility at earlier life stages will vary with the old-age-pension system as discussed below and gender specific survival rates, which are present in the above equations.

Before age T^*+1 individuals have to decide on their labor supply. Therefore the value function at this age is specified as,

(3c)

$$V_t(x_t) = \max_{h_t} \left\{ u(y(h_t), h_t, x_t) + \beta \eta_{t+1|t} \pi(x_t, h_t) E_t [V_{t+1}(\bar{x}_{t+1}, m_{t+1} = 0)] + \beta \eta_{t+1|t} (1 - \pi(x_t, h_t)) E_t [V_{t+1}(x_{t+1})] \right\}$$

Empirical implementation

Empirical implementation requires specification of functional forms for u , π and f to be able to solve the dynamic programming problem⁷. Moreover, implementation requires that we can identify individuals with a state in the state space. Therefore we have to estimate missing state variables, i.e. potential wages for non-workers or single status at age 95. First we discuss the functional form of u , f , and π and then estimation of the state

⁷ Recently semi- and nonparametric dynamic discrete choice models have occurred. See Heckman et al., 2005 for a semiparametric model and Aguirregabiria, 2005 for a nonparametric model. Magnac et al., 2002 also discuss the possibility of relaxing parametric specification of some of the functions.

space. Finally, empirical implementation requires a functional form for the unobserved state variable, which is last topic of this section.

The instantaneous utility function is given by,

(4)

$$U(y(h_t), h_t, X_t) = (1 + e^{\nu I(h_t=w)})g(y(h_t)) + e^{\sum_{j=u,r} \lambda^j X_t} + e^{\phi age} + \sum_{j=R,U,W} \alpha^j I(h_t = N, h_{t-1} = j) + \sum_{j=r,u,n,w} \varepsilon_{jt}$$

$$g(y(h_t)) = \frac{y(h_t)^{1-\rho} - 1}{1-\rho}, \rho \geq 0, \rho \neq 1$$

$$g(y(h_t)) = \log y(h_t), \rho = 1$$

The first term says that utility is increasing and concave in income. The marginal utility of consumption (income) is a complement ($\nu < 0$) for or a substitute ($\nu > 0$) for leisure.

The second term reflects valuation of leisure, which is strictly positive. X_t is a vector of variables, which shifts the valuation of leisure, including a constant. The variables are dummies for sex, for being single and for the interaction of female and being single. As these variables do not vary across choices they are only identified relative to full time work. But we let the value of leisure differ across unemployment and retirement. We think of these differences as the search cost imposed on unemployed. During unemployment search efforts, meetings at job centers, acceptance of job offers (sometimes in a different occupation) and other requirements are necessary to retain UI benefits.

The taste shifters, sex and single, are catching observed heterogeneity. Across the sexes, the valuation of leisure can differ because of difference in non-market production. And even within sexes valuations could differ across individuals with different marital status.

The third term, age enters exponentially and reflects omitted health variables. While for the working population we do have an indirect measure of health, in the form of sickness

benefits, we lack the information for unemployed and retired. As with wages we could try to predict health for these, but from a previous study (Junge, 2007), this health proxy is insignificant in a reduced form retirement equation in a sample very similar to the current.

The age term is quite important in our specification. The reason is that early retirement is only available from age 60 to 66 and we need to remove changes in health as a function of age to be able to identify the differences in valuation of the unemployment and early retirement states.

In the fourth term, state dependence is introduced as a cost of moving from either retirement or unemployment or work to a new job, where we have introduced the notation R for retirement, U for unemployment, W for work and N for a new full time job. This ‘transitory’ utility cost can be interpreted e.g. as the difficulty of obtaining an acceptable wage offer, which is not explicitly modeled. Since we expect that cost of leaving unemployment is much lower than retirement, these are the positive side of search cost.

Finally, part of the state variables are unobserved, which is reflected in the fifth term. We assume that individuals observe this preference shock in the current period but do not themselves have information about future shocks.

Transition equations

The transition equations, f , are functions of historical labor force participation. Experience, e , on the labor market is given by,

$$(5) \quad e_{t+1} = e_t + I(h_t = W)$$

i.e. experience is incremented by 1 if the individual is working in the current period.

Tenure, m , is almost as simple,

$$(6) m_{t+1} = (m_t + I(h_t = W))I(j_t = j_{t-1})$$

Here tenure is incremented by 1 if the individual works an additional year in the same plant ($j_t=j_{t-1}$). From the Mincer type wage equation discussed below, we can predict potential wages with information on e and m .

We also keep track of duration of unemployment.

$$(7) Du_{t+1} = \min(4, Du_t + I(h_t = U))I(h_t = U) + 4I(h_t = R)$$

If Du_t takes the value 4 individuals have exhausted their UI benefits and must live on a subsistence level discussed below. Note that it is not possible to go from retirement to obtaining UI benefits, that is, we let Du_t equal 4 in this case. In the same way we keep track of number of years in early retirement,

$$(8) Dr_{t+1} = \min(7, Dr_t + I(h_t = R) + Du_t I(h_t = U))I(el_t = 1) + 8I(el_t = 0)$$

Where el_t is an indicator for eligibility. If Dr_{t+1} take the value 8 then individuals are no longer eligible. Suppose individuals enter retirement from unemployment. Then the period with high benefit is reduced accordingly, which is why Du_t enters the equation. The eligibility condition is,

$$(9) el_{t+1} = el_t I(h_t = r, h_{t-1} = r) + I(el_t = 1)(I(h_t = w, u) - I(Du_t = 4)), el_a = 1$$

The first term on the right hand side says that eligibility status do not change as long they stay retired. For example if individuals leave early retirement this term turns to zero. The second terms says that if individuals are currently working or unemployed and did not exhaust their UI benefits, he/she retains eligibility if they were eligible. Finally, initially all individuals are eligible through sample selection.

The remaining state variables, age and year, are straightforward. Below we discuss being single and creative destruction (the risk of being fired) at ‘potential’ workplaces.

Layoffs

We assume that before decisions on labor force participation take place there is a probability of getting fired, which is a function of the change in employment within the plant where the individual is employed. We do not observe layoffs. Instead it is assumed - to help identification - that when a plant shuts down we observe a layoff and we also assume that individuals cannot be laid off from a startup. Layoff probabilities in between are modeled as a function of the following index describing firm job creation and destruction,

$$(10) \text{ DHS}_t = 2 \frac{n_{t+1} - n_t}{n_{t+1} + n_t}$$

Where DHS_t stands for Davis, Haltiwanger and Schuhs index (Davis et al., 1996), and n is the number of employees in the firm. The index is the change in employment divided by average employment over two periods. Note that for a firm closing it equals -2 and for a startup it equals 2. Let π denote the layoff probability, which equals 1 if a plant closes and 0 if a plant opens. In between we let the layoff probability be given by the following function,

$$(11) \pi_t = 1 - \frac{\theta_1 \left(1 + \frac{1}{2} \text{DHS}_t\right)}{\theta_1 \left(1 + \frac{1}{2} \text{DHS}_t\right) + \theta_2 \left(1 - \frac{1}{2} \text{DHS}_t\right)}, \quad \frac{\theta_1}{\theta_1 + \theta_2} > k$$

where DHS_t equals -2, π_t is one and when DHS_t equal to 2, π_t is zero. The parameterization can be seen as modeling the layoff probability as a beta distribution with support $[-2, 2]$ and where $\theta_1(1+\frac{1}{2}\text{DHS}_t)$ and $\theta_2(1-\frac{1}{2}\text{DHS}_t)$ are the parameters. The inequality constraint in (?) assumes that the layoff probability for plants with status quo

in the number of employees ($DHS_t = 0$) is smaller than $1-k$. Where k is estimated as the proportion of older workers (age 51 to 59), which stay in a plant with $DHS_t = 0$ ($\hat{k}=0.963$).

Individuals currently in work are assumed to know the number of employees in the plant in the next period and can then back out their layoff probability. This assumption can at least on some grounds be justified by the fact that mass layoffs must be reported in advance. For future periods they expect that no changes will occur within the plant.⁸

Estimates of missing state variables

After the specification of functional forms we will discuss how we obtain values of the relevant variables for the entire state space. First, we need a projection of the future tax and pension system. In a given year we assume that individuals have myopic expectations and take the current system as given for all future periods. This is a standard assumption in many papers and makes it relatively easy to obtain tax and benefit values for years beyond the sample.

Experience, tenure and the duration and eligibility state variables are all controlled processes and we do not need to calculate future values for these as they will depend on the history of employment decisions. But we obtain real wages for individuals by a Mincer type wage equation, where human capital, plant specific capital and ability enter. We use this very simple specification from the literature because it does not introduce additional state variables. Specifically, we model the log real wage rate as a function of tenure, experience and experience squared. The log real wage can then be predicted for each individual through their employment history. In addition to the human capital variables, tenure and experience, agents are assumed to be endowed with a given ability, which we describe by a fixed effect. The results are given in Table 1. The regression is based on a sample of older workers (50+) which change plant in the first year we observe

⁸ Expectations on future layoff probabilities could also be estimated by introducing a parameter, which estimates DHS for future periods. Alternatively, agents could use plant size and age to predict job turnover in plants. But a regression of DHS on these did only produce very weak correlations.

them. This avoids problems with left censoring of tenure (prior to 1980 we have no knowledge of tenure). Therefore all individuals have initially no tenure in the estimation sample. Experience is measured from 1964 through a measure of contribution to a public related pension program, where contributions are based on annual hours worked. We also include a dummy for years after 1992 because of a break in the data series for wages. The effect of experience is found to very small whereas tenure is quite important. As an example: Let us assume an individual receives DKK 60 per hour of work with initial experience of 16 years. After 10 years of tenure (+experience) the wage has increased to DKK 71 per hour. After a job change the hourly pay will be DKK 61 per hour a drop around 14.0 percentage point. Changing or losing a job is therefore associated with a large drop in annual income.

We also need a prediction of creative destruction inside plants. For individuals employed we use the observed index for the current period and for others and for out of sample predictions, we assume $DHS = 0$. Individuals in employment at a specific plant therefore know the number of employees in the next period, which is not a bad approximation given the rules with advantage notice in case of a mass layoff. In all other situations individuals assume the firm will have no growth in employment.

The final state variables are “being single” and “mortality”. We apply the observed values from the sample, and for future periods we use the last observed value. One possibility would be to use the age-gender specific survival rates and apply these to the spouse. For mortality we use age-gender specific survival rates, which are obtained from public available life tables (www.statistikbanken.dk).

The unobserved state variable

We assume that the unobserved state variables, the fifth term in (4), are iid extreme value type I distributed. This has two implications: First we can solve for the future expected utility in the value functions (3). Second, the conditional choice probabilities have closed form solutions, which we return to below. This assumption is frequently invoked in

dynamic programming models (Rust, 1994, Rust et al., 1997, Heyma, 2004). Let us introduce the choice specific value function $V_t^j(X_t)$, $j = R, U, N, W$. We can then write the expected value of future utility in closed form,

$$(12) E \max_{k=W,R,U,N} \{V_{t+1}^k(X_{t+1})\} = (\gamma + \ln(\sum_{k=W,R,U,N} \exp(V_{t+1}^k)))$$

where $\gamma = 0.5772$ is Euler's constant. This is for individuals working. While for individuals not working we have,

$$(13) E \max_{k=R,U,N} \{V_{t+1}^k(X_{t+1})\} = (\gamma + \ln(\sum_{k=R,U,N} \exp(V_{t+1}^k)))$$

4. Estimation method

In this section we discuss estimation of the parameters $\Gamma = (\nu, \rho, \phi, \alpha^r, \alpha^u, \lambda^r, \lambda^u, \lambda^w, \theta_1, \theta_2)$ of the theoretical model. An additional layer of complication introduced by the unobserved layoff is addressed by introducing the EM-algorithm to maximize the log likelihood function.

The current period unobserved state variable is unobserved to the econometrician and needs to be integrated out of the value function. The multinomial logit is the closed form of the extreme value type I distribution (Rust, 1994). Hence, the conditional choice probability for alternative j in period t takes the following form,

$$(14) P_\delta(h_t = j | X_t, \Gamma) = \frac{\exp(V_t^j(X_t))}{\sum_{k|h_{t-1}} \exp(V_t^k(X_t))}$$

where we have introduced the notation $k|h_{t-1}$ to indicate that the sum is over the choices which are possible conditional on previous period's choice. For example the decision to stay in the current job depends on having a job in the previous period.

The maximization of the likelihood function over all conditional choice probabilities is straightforward except for the incomplete control assumption introduced by layoffs. Let $L_t = 1$ if a layoff occurs and 0 otherwise. If we observed data for layoffs, the choice probability would be influenced for choices j and layoff k in the following way for individuals working in the previous period,

$$(15) P(h_t = j, L_t = k | X_t, \Gamma) = [P_\delta(h_t = j | X_t, L_t = k, \Gamma) P_\pi(L_t = k | X_t, \Gamma)]^{I(h_t=j, L_t=k)}$$

Where P_δ is the conditional choice probability above and P_π is the probability distribution for L_t , which we have specified in Section 4. Since we do not observe L_t , we cannot compute this probability.

The formulation in (15) is a mixture of two components for the choice probabilities. The population is split in those experiencing a layoff and those who do not. This model can be estimated by maximizing the expected log likelihood function where we have integrated out the layoff probability conditional on the observed data.

For those who were out of employment during the previous period the layoff outcome is obviously zero and their choice probability is not influenced by layoffs at all.

For those who stay in the plant, the layoff outcome is also zero. Their choice probability is then as if the layoff was observed,

$$(16) P(h_t = j, L_t = 0 | X_t, \Gamma) = [P_\delta(h_t = j | X_t, \Gamma) P_\pi(L_t = 0 | X_t, \Gamma)]^{I(h_t=j)}$$

For those who experienced a separation from their plant, we need to compute the expected layoff probability conditional on data and the probability model. Let the unconditional layoff probability be π_t . We can then compute the conditional layoff probability as,

$$(17) E[\pi_t | X_t, h_t] = \frac{\pi_t \tilde{P}_\delta(h_t | X_t, \Gamma)}{\pi_t \tilde{P}_\delta(h_t | X_t, \Gamma) + (1 - \pi_t) P_\delta(h_t | X_t, \Gamma)}$$

where \tilde{P}_δ is the choice probability after a layoff, where the choices are restricted to “retirement”, “unemployment” or “a new job”. And P_δ is the choice probability without layoff, that is, choices can be “continue work”, “retirement”, “unemployment” or “a new job”. For a moment assume that $P_\delta(h_t = W) = 1$ then $P_\delta(h_t = j) = 0, j = R, U, N$. It can be seen that the conditional layoff probability is one. The intuition is quite clear. For individuals, who want to stay in the plant no matter what, a separation can only be associated with a layoff. For clarity assume that $P_\delta(h_t = R) = 1$ then $\tilde{P}_\delta(h_t = R) = 1$ and the conditional layoff probability is equal to the unconditional. First, this reveals that no information about layoffs can be drawn from this observation and second that layoffs occur before quits.

Moreover, suppose the plant shuts down. Then the conditional layoff probability is independent of choice probabilities and this is also the case for newly started firms.⁹ Perhaps a clearer interpretation of the conditional layoff is reached if we divide the numerator and the denominator by \tilde{P}_t . In this case we obtain the following expression,

$$(18) \pi_t | X_t, h_t = \frac{\pi_t}{\pi_t + (1 - \pi_t)(1 - P_\delta(h_t = W | X_t, \Gamma))}$$

The conditional layoff probability is the layoff probability divided by the probability of layoff plus the quit rate.

This formulation gives rise to the following likelihood contribution (taking logs of the choice probabilities), where we have split the contribution into,

a/ Those who were out of employment in the previous period

⁹ A more elegant solution would recognize that the exit probabilities and choice probabilities conditional on DHS should be an equilibrium condition.

- b/ Those who were employed at the same plant for two consecutive periods
 c/ Those who were separated from their plant either because of a layoff or because of a quit.

a/

$$(19a) \log(P_{\delta}(h_t = j | X_t, \Gamma)) = \log \left(\frac{\exp(V_t^j(X_t))}{\sum_{k=R,U,N} \exp(V_t^k(X_t))} \right), j = R, U, N$$

b/

$$(19b) \log(P(h_t = W | X_t, \Gamma)) = \log \left(\frac{\exp(V_t^W(X_t))}{\sum_{k=W,R,U,N} \exp(V_t^k(X_t))} (1 - \pi_t) \right)$$

c/

$$(19c) \log(P_{\delta}(h_t | X_t, \Gamma)) = E[\pi_t | X_t, \Gamma] \log \left(\frac{\exp(V_t^j)}{\sum_{k=R,U,N} \exp(V_t^k)} \pi_t \right) + \\ (1 - E[\pi_t | X_t, \Gamma]) \log \left(\frac{\exp(V_t^j)}{\sum_{k=R,U,N,W} \exp(V_t^k)} (1 - \pi_t) \right), j = R, U, N$$

The complete log likelihood function is the sum of the three contributions conditional on the appropriate employment history. This expression can then be directly maximized. An alternative procedure is the expectation maximization (EM) algorithm of Dempster et al., 1977, which is frequently applied in mixing models, where it is easy to maximize the likelihood function if the mixing components are observed. This procedure is very stable

but slow. The algorithm is straightforward to use in our case¹⁰ and only influence term c above. Instead of taking the expectations of layoff probabilities with respect to Γ , we do it with respect to a current guess of parameters Γ^t , iterate on the conditional choice probabilities a couple of times (note that the mixing components are fixed here) and obtain new parameter values Γ^{t+1} . These can be substituted into the computation of the mixing components. A new iteration increases the log likelihood function and so forth. This continues until convergence.¹¹

Estimates of covariance parameters for the parameters follow by a single iteration on the full maximum likelihood function after convergence.

5. Data

We use a 10 percent sample of Danish population from 1980-1997. The data has been kindly provided by the DREAM group¹² in Denmark. The data are derived from the fact that all residents in Denmark has a social security number which is used in all transactions recorded including the income tax register, job centers, bank accounts, participation in welfare programs, and other transactions with the welfare system. Moreover, the data are merged with plant information.

We focus on a sub-sample of private sector employees without academic degree, which are eligible for early retirement benefit. This is to minimize the potential influence of other pension schemes.¹³ The eligibility criteria for early retirement benefit is satisfied by focusing on individuals belonging to the cohorts 1925-1930, who need 5 years of UI fund membership to be eligible. We also removed self-employed as their disposable income is difficult to compute. The only retirement routes in this paper are early retirement or old

¹⁰ Note the EM-algorithm is a very general optimization procedure and we do only describe the way we have implemented it.

¹¹ We have chosen a fairly high convergence criterion (1e-3), which is due to the slow progress of the algorithm. Moreover we were often times able to calibrate the likelihood function much faster, which might indicate it would be a good idea with a sensitivity analysis of the likelihood function.

¹² The Dream group is a macroeconomic policy planning unit in the Ministry of Finance.

¹³ A large fraction of public employees had a civil servants pension. And labour market pensions were widespread among academics.

age pensions and we remove all disability pensioners. Finally, because of the need to obtain estimates of firm specific human capital effects (tenure) we rely on a sample of individuals, who changed job during 1980 or 1981 and therefore had their tenure reset to zero.

Descriptive statistics of the sample can be seen Table 2. The sample consists of 7,092 observations on 431 individuals, an average of 16.4 per individual. As noted above we have to solve for each individual the maximization problem inside a maximum likelihood procedure, which in general is quite expensive. However to solve for the 16 observations available on average we only need one backward recursion. This saves us a lot of computing time compared with shorter panels, where estimates of age effects must be obtained from different age groups. The sample is dominated by cohabiting or married males. Average tenure is low as we cannot observe tenure before 1980 (see also sample selection). Experience – on the other hand – can be tracked back to 1964. This is actual experience calculated from ATP contributions. ATP is a very small supplementary labor market pension paid on basis of supplied hours instead of income. The mean real wage is DKK 65 per hour. Income on a yearly basis is DKK 123,500 assuming 1,900 working hours per year. The UI (and early retirement) benefit maximum is approximately 84,000. This implies a mean replacement rate of 0.68 for the workers in our sample.

The possible choices are retirement, unemployment and full time work. The definition of retirement is based on the labor force status in the last week of November in a given year. Long-term unemployment requires a unemployment degree of at least 50% in a given year. Finally, work is the residual category. We re-classify a few workers as retired if their labor market income does not exceed pensions because the November status requires only very small amounts of labor market income to be classified as in work. The ten most typical employment patterns are given in Table 3. As expected the full time work with no intervening other jobs or unemployment is most dominant, accounting for 50 percent of the sample. The three following sequences are also straightforward. One sequence with a single job change and another sequence with unemployment before retirement, and a sequence with only fulltime work. After that the labor market histories

becomes more and more complex. A sequence which is very rare is the transition from retirement to work. Only 0.2 per cent of the individuals in our sample make a return to work from retirement. We allow for a transition from retirement to work, which allows us to estimate the cost associated with return conditional on the policy rules.

6. Results

In this section we discuss parameter results and fit of the model. The model was estimated with the EM-algorithm as discussed above. Depending on the starting values we frequently had difficulties with a degenerate point. The algorithm did end up in a point with no layoff at all. This was particular troublesome with the maximum likelihood algorithm, whereas the EM-algorithm in general was more stable. One possible solution is to estimate the wage equation jointly with the retirement model and let the quit rate influence the wage distribution directly.

The t-ratios on the parameter estimates are quite low, which is not uncommon in the dynamic programming literature, despite the large number of observations. The problem is that the value function is correlated highly with the instantaneous utility function (Rust et al., 1997). As we have only used 5 per cent of our sample to estimate the model, some of the problems could probably be solved by increasing sample size. However, this will slow down an already slow estimation procedure.

Finally, when we discuss fit of the model, we do it with a series of figures and tables. Alternatively, goodness of fit could be assessed through chi-square statistics.

Parameter estimates:

Table 4 provides parameter estimates and associated t-ratios for the estimated model. We find that leisure and consumption are substitutes ($\nu > 0$). The value might seem a little high. But we have not included a cost of going to work or a minimum subsistence level of living, which would have lowered the value. For example in Van der Klaauw et al., 2006,

the value is only slightly positive; but they also include a minimum subsistence level in their model. The CRRA, ρ , parameter estimate is fairly low but close to the findings of other studies, which impose the restriction that consumption equals income (Van der Klaauw et al., 2006).

Preferences for leisure differ across taste shifters and across unemployment and retirement. The reference person within each state is a cohabiting¹⁴ male, α^{constant} . First, sex, α^{female} , is the main difference for leisure preferences in retirement, where females enjoy a not very significant higher value than males. The other leisure parameters have very low significance. So the difference continues to exist independent of being single or not. Second, preferences for unemployment leisure (net of search cost) seem to vary across the taste shifters. Single males, α^{single} , cohabiting females, α^{female} , and single females, $\alpha^{\text{single}} + \alpha^{\text{female}} + \alpha^{\text{single*female}}$, have higher preferences for unemployment leisure compared with cohabiting males, though the difference is only slight significant for single males. The age effect, ϕ , is significant, which reflects the gradually deterioration of health.

We find that the value of leisure is significantly higher during early retirement benefit and old age pension than through unemployment insurance. As the model interprets the difference in valuation of leisure across unemployment and retirement as search cost of unemployed, we can compute the implied search cost for each group. One should be careful in evaluating these differences and remember taste shifters are in exponential form. Therefore the value of search cost to married males are $\exp(\alpha^{\text{constant,U}}) - \exp(\alpha^{\text{constant,R}})$, where R and U denote state. From these we see that single females have the largest search cost, males in couples, females in couples and finally single males.

State dependence reveals the expected pattern with retirement being associated with a deterioration of human capital and possibly a lack of search incentives, and therefore a higher cost of returning to work, though the effect is not very significant. Loss of human

¹⁴ With “cohabiting” we mean both individuals married and not married; but living together with another person.

capital is less for the unemployed compared to retired which implies a lower cost of returning to work for the unemployed. It could also be that search requirements raise the quality of wage offers. Interestingly there is no significant difference between the cost of changing job for the employed and for unemployment. Given that the employed do not suffer a loss in human capital; the interpretation is that search is more efficient while unemployed. We will have to remember these effects of search cost in accessing the difference in valuation of leisure (net of search cost).

Estimates of the layoff distribution can be seen in Figure 2 as a function of the DHS index. On one hand plants, which experience a considerable downsizing, the layoff probability is very high. On the other hand we find that it drops quickly as the DHS index approaches zero and that there is very small layoff probabilities in expanding plants.

Model fit

The estimated model can be simulated and we can compare the fit of the model with-in sample. For simulation of the model we use path simulations. We draw the unobserved state variable, the probability of death, and a probability of layoff. We use these values to simulate the decision for the initial age. Next periods state variables can be updated. For the DHS index we use the value zero in this updating. Therefore conditional on the simulated outcome, we simulate the decision for the next period and so on. For each individual we make ten simulations and average the results.

Figure 3a (3b) to 6a (6b) show the fit to unemployment, retirement, work in old job and work in a new job by single status for males (females). Retirement in Figure 3a and 3b is almost nonexistent until age 60 at that age a very large increase in retirement takes place for both genders and for singles as well for cohabiters. Especially, the female sample responds very strongly to the age eligibility criteria. At age 61, 90 per cent of the females in the sample have withdrawn to retirement. Male in general respond slower with 80 per cent being retired at age 62. The model slightly over predicts early retirement and after that under predicts old age pension. However, overall the fit is reasonably good.

In Figure 4a and 4b we compare actual unemployment and simulated unemployment. The overall pattern is an increase in the unemployment degree by age until age 59 and then a very fast decline after early retirement becomes available at age 60. However, it also clear that the unemployment picture is more variable than for retirement. Thus it is quite challenging to fit these long-term unemployment states. The increase in unemployment around age 57 to 59 is overstated in the model. And likewise for age 60+ the model overstates unemployment. The problem is that the model predicts unemployment duration to be too long.

Figures 5a and 5b for continued work in an existing job show that 80-90 per cent of individuals stay with their plant up from age 51 to around age 58 with a very small decline. The model predicts this reasonably well. After age 58 individuals start to go on UI benefit and retire and this reflected in the sharp decline from age 58+. The figures are close to the reverse of Figure 3a and 3b. Around 60 there is a sharp drop in the number of individuals staying in their existing job due to retirement. It is more pronounced for females than males.

Finally, jobs in new plants are described in Figures 6a and 6b. Again the overall impression is that it is quite variable, which is unsurprising, as transitions out of long-term unemployment is to a new job before the age 50. Again from age 58 the number of individuals in a new job decreases.

In Table 5a and 5b employment transitions probabilities can be seen for age groups 52-59 in Table 5a and for age 60+ in Table 5b. The employment transitions from new and old job to all other states before age 60 are reasonably well described. Though, in the data the transition to unemployment is higher in a new job than in an old job, which the model is not able to reproduce. Also the actual levels of separation from a new job to another new job are higher than the prediction from the model. On the other hand the number of transitions from an old job to a new job is slightly overstated. The same pattern between new and old job is repeated for the age group 60+. However, the transition to

unemployment is reversed in this case. Transitions out of unemployment reveal some greater differences. The model predicts much less outflow to a new job than the data. This explains a much higher unemployment rate in Figures 4a and 4b. Finally, we see that the transitions from and to retirement are well predicted with the model after age 60, but not very well before age 60. But the number of retirement sequences is extremely low before age 60.

The results showed estimated parameter values within plausible ranges. Unemployment is valued less than retirement; but search lowered the value of leaving unemployment. The predictions of the model revealed some very interesting things. First, the model fits participation rates well. The exception is unemployment, which is clearly overstated. Second, from Table 5a and Table 5b, we saw that the model predicts too long unemployment spells, which explain the very high unemployment. Third, new jobs are more volatile in the data than the model predicts. Fourth, retirement is well predicted in the model.

7. Experiments

We have previously indicated that once preferences are estimated we are very free to choose a simulation experiment. In this section we simulate the outcome of two experiments. The first is aimed at evaluating the effects of introducing the unemployment related pension scheme, which we will refer to as transitional allowance (“Overgangsydelsen”) that was mentioned in the introduction. The second experiment measures the difference between the value of leisure in unemployment and in retirement in money equivalent units.

Experiment 1: Unemployment related retirement

In 1992 a new scheme, transitional allowance, for early retirement appeared in Denmark. It was directed towards the long-term unemployed. Such new initiatives should be followed by careful analysis of the incentives build into the program. With the model in

this paper we can introduce an unemployment related retirement program. We do this by implementing the details of the program and predict the outcome from the model under two different assumptions on preferences for leisure, which we return to below.

Implementation of this scheme requires new transition equations for unemployment and retirement duration (equation (7) and (8)). After one period of unemployment individuals become eligible for transitional allowance. The benefit level is identical to UI benefit, but is reduced to 80% after the three periods on maximum benefit has expired. Likewise the Early retirement benefit will be 80% of the maximum for all years. This means a loss in early retirement benefit. Individuals are allowed to work but suffer a reduction in transitional allowance. Re-entry to UI system is possible by fulfilling the work requirement in the model of one period of work.

First, we assume leisure time for individuals on transitional allowance is as unemployment. The popularity of scheme is quite limited. In Figure 8 (Simulation, T1), the maximum fraction of individuals on transitional allowance is 2.8 per cent at age 59. As we expect that unemployment might increase due to the transitional allowance requires an UI spell, we compare unemployment with and without transitional allowance. In Figure 7, we can see that they are quite identical except for the few individuals who leave for transitional allowance. Distortions are almost absent in this case because unemployment insurance clearly dominates transitional allowance.

Next, we evaluate the transitional allowance under the alternative assumption that individuals on transitional allowance enjoy the leisure preferences of retired. Figure 7, simulation T2, now shows a very huge inflow to unemployment followed by a massive entry to the transitional allowance scheme (Figure 8, simulation T2). Moreover, there is an increase in the early retirement scheme at age 60, which was not present before (not shown).

Experiment 2: Leisure in income equivalent units

We will in this section try to compute the willingness to pay (WTP) from the quality change of leisure, where the quality change comes from removing the search cost in unemployment by using the retirement parameters for leisure. This exercise, although not necessarily in discrete labour supply models, is frequently done in static random utility models. We can define C_{ij} as the income compensation necessary to equate utility of choice i before the reform with the utility of choice j after the reform.

$$(20) U(y(i), i, X_i) + \varepsilon_i = U(y(j) - C_{ij}, j, \tilde{X}_i) + \varepsilon_j$$

where U is instantaneous utility and we have used a static model with additive error term. The state variable X includes observed state variables as well as parameters of the utility function. \tilde{X} contains state variables and parameters quality adjusted. Computing willingness to pay in this model is detailed in McFadden (1999). Two problems are present in these calculations. First the error term induces a probability distribution on the willingness to pay and second income effects might be present in the conditional choice probabilities under the quality improvement and compensation.¹⁵

To be able to compute the willingness to pay it is required that utility is monotone and increasing in income. Since our utility function is strictly increasing in income and the value function inherits this property (Bertsekas, 1999), we note that it directly extends to our dynamic model above. It would be possible to calculate these values for a change in quality for example for the next period or for all periods, and the willingness to pay for this. However, our model does not allow for transferring money across time so it is not obvious how to interpret the resulting willingness to pay. Instead we will calculate the willingness to pay for a quality change in the current period. While this is similar to the static analysis, we should emphasize that our model is estimated taking into account all the dynamic incentives of social insurance and labour market opportunities.

Let us define the measure of willingness to pay in our model for individuals choosing i before the change and j after the quality improvement.

¹⁵ If utility is a nonlinear function of income or marginal utility of income differs across alternatives.

(21)

$$U(y(i), i, X_t) + \varepsilon_{it} + \beta(\gamma + \log(\sum_k \exp(V_{t+1}^k))) = U(y(j) - C_{ij}, j, \tilde{X}_t) + \varepsilon_{jt} + \beta(\gamma + \log(\sum_k \exp(V_{t+1}^k)))$$

which is identical to the static discrete choice setting except that we have included the expected value of future utility. A standard assumption is that the random state variable is invariant with respect to the quality improvement, which makes it particular easy to calculate the willingness to pay for individuals, who do not change alternative after the quality improvement and compensation. This is because the error in equation (?) drops out under $j=i$.

Let us define C , the income adjustment that equates maximum utility before quality improvement and maximum utility after the quality adjustment, where we allow substitution between alternatives. It is intuitive to see that this will be bracketed by C_{ii} and C_{kk} . On one hand the income compensation required for keeping initial utility constant at choice i after the quality adjustment, C_{ii} , must be less than C because individuals are allowed to substitute away from i after quality improvement and thereby increase ex post utility. Utility under alternative k must be higher than i after the quality adjustment and less is needed to equalize utilities. On the other hand C must be less C_{kk} because individuals are allowed under initial utility to choose i instead of k . Ex ante utility under alternative i is higher than utility for alternative k . Therefore it takes more to equalize utility between k and k . This argument requires only that the value functions are monotone and increasing in income (see McFadden, 1999, and De Palma et al., 2003, for a proof).

We have

$$(22) \quad C_{ii} \leq C \leq C_{kk}$$

if case i is chosen before improvement and k after the improvement. Note that if the improvement – in this case value of leisure in the current period – do not influence a

particular choice k , C_{kk} must be zero. Taking expectation on both sides of this inequality, we get bounds on the expected compensating variation.

$$(23) \sum_i P_i^0 C_{ii} \leq E[C] \leq \sum_k P_k^1 C_{kk}$$

where P^0 is the conditional choice probability ex ante and P^1 is the conditional choice probability ex post. A problem with computing these values are that the upper bound depends on the true income compensation in the conditional choice probability. Herriges and Kling, 1999, provide some computable bounds on these. Here we choose to note that the choice probabilities are bounded by the uncompensated choice probabilities. This is because the income effect present in the choice probability decreases the probabilities due to utility is monotone increasing in income.

In Table 6 we have computed the willingness to pay for the unemployed at age 51 (results for other age groups show very similar willingness to pay), whom is affected by a change in the quality improvement of leisure. Note the characteristic of the individual only varies with respect to taste shifters in these calculations. This we can do by solving the non-linear programming problem in (21). We could compute the overall population willingness to pay by weighting each group appropriately. Instead, we report result across sex and cohabitation status. The results show quite high willingness to pay for this improvement. The value of DKK 51,000-52,000 (after tax) should be compared with an unemployment benefit level after tax of approximately DKK 53,000 in 1980 DKK. As expected single women are willing to pay the most for this and single men the least. However, the differences are quite small.

These values should not be taken at face value. The point is that individuals are willing to reduce their income to a level very close to minimum subsistence level to avoid the search costs of unemployment if they are unemployed. In our model the subsistence level is zero. In a more realistic version of the model, the subsistence level should be included.

By using the estimated choice probabilities we can compute bounds on the expected willingness to pay for avoiding search costs. This might be a more realistic measure in our model as the choice probabilities take into account utility derived in the other states. The bounds are given in Table 6 as well. They seem to be quite broad. Since we are using uncompensated ex-post conditional choice probabilities, the large substitution to unemployment after quality improvement increases the upper limit. Remember the theoretical correct bound should decrease income by the true willingness to pay. The expected willingness to pay reflects the average over the states and do therefore not depend on which state the individuals occupy.

It should be noted that the welfare calculation depends highly on the functional form used for computing these measures. Something that becomes evident from the discussion on static random utility models. The virtue of the present calculations is of course that we take into account the dynamic incentives in public support to retiree and unemployed.

This section has clearly shown that the distortions of early retirement benefit and transitional allowance can be very large.

8. Conclusion

In this paper we have seen that unemployment and retirement are two distinct states contrary to what previous structural retirement models have presumed. The main conclusion is that the early retirement scheme is highly distorting incentives mainly because of lack of search requirements. This is even the case for a period with general high unemployment and therefore loose administration or enforcement of search requirements.

We estimated a dynamic programming model with incomplete control and found parameter values within plausible ranges. The model predicted retirement and job transitions well, but was overstating unemployment. Policy simulation revealed large distortions from introduction of unemployment related retirement schemes. Moreover, we

saw that individual's willingness to pay for abolishing search cost in unemployment was very large.

Literature

Aguirregabiria, Victor (2006): “Another Look at the Identification of Dynamic Discrete Decision Processes, with an Application to Retirement Behavior”, Working paper, September 2006.

Berkovec, James and Steven Stern (1991): “Job Exit Behaviour of Older Men”, *Econometrica* vol. 59, No. 1, (Jan 1991), pp. 189-210.

Bertsekas, Dimitri (1999): “Dynamic Programming and Optimal Control”, Volume 1, Athlean Scientific, Ed. 2, 1999

Blöndall, Sveinbjörn and Stefano Scarpetta (1999): “The Retirement Decision in OECD Countries”, OECD Economic Department working papers 202, OECD Economics Department

Börsch-Supan, Axel; Reinhold Schnabel; Simone Kohnz and Giovanni Mastrobuoni (2005): “Micro-modeling of Retirement Decision in Germany”, in Gruber and Wise “Social Security Programs and Retirement Around the World: Microestimation”, University of Chicago Press

Browning, Martin and Søren Leth-Petersen (2003): “Imputing Consumption from Income Tax Registers”, *The Economic Journal*, 113(488): F282-301

Davis, Steven; John Haltiwanger and Scott Schuh (1996): “Job Creation and Destruction”, MIT Press, 1996

Deaton, Angus (1991) “Savings and Liquidity Constraints”, *Econometrica*, Vol. 59, No. 5, pp. 1221-1248

Dempster, Laird and Rubin (1977): "Maximum Likelihood from Incomplete Data via the EM-algorithm", *Journal of the Royal Statistical Society. Series B*

Eckstein, Zvi and Gerard J. van den Berg (2006): "Empirical Labor Search: A Survey", *Journal of Econometrics*, 136 (2007), pp. 531-564

French, Eric (2005): "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behaviour", *Review of Economic Studies*, Vol. 72, No. 2 (April 2005), pp. 395-427

French, Eric and John Bailey Jones (2004): "The Effects of Health Insurance and Self Insurance on Retirement Behavior", Federal Reserve Bank of Chicago, WP. 2001-18, Revised March, 2004

Heckman, James and Salvador Navarro-Lozano (2005): "Dynamic Discrete Choice and Dynamic Treatment Effects", IZA Discussion Paper No. 1790

Herriges, Joseph and Catherine Kling (1999): "Nonlinear Income Effects in Random Utility Models", *The Review of Economics and Statistics*, Vol 81, No. 1. pp. 62-72

Heyma, Arjan (2004): "A Structural Dynamic Analysis of Retirement Behaviour in the Netherlands", *Journal of Applied Econometrics*, No. 19, pp. 739-759

Junge, Martin (2007): "An Econometric Model of Early and Ordinary Retirement", Working paper January 2007.

Magnac, Thierry and David Thesmar (2002) "Identifying Dynamic Discrete Decision Processes", *Econometrica*, Vol. 70. NO. 2, pp. 801-816

Mahieu, Ronan and Didier Blanchet (2005): “Estimating Models of Retirement Behaviour on French Data”, in Gruber and Wise “Social Security Programs and Retirement Around the World: Microestimation”, University of Chicago Press

McFadden, Daniel (1999): “Computing Willingness to Pay in Random Utility Models”, In J. Melvin, J. Moore, and R Riezman: “Trade, Theory and Econometrics, Essays in Honor of John Chipman”, Routledge, London, 1999

Oshio, Takashi and Akiko Sato Oishi (2005) “Social Security and Retirement in Japan: An Evaluation Using Micro-Data”, in Gruber and Wise “Social Security Programs and Retirement Around the World: Microestimation”, University of Chicago Press

De Palma, André and Karim Kilani (2003): “Compensating variation in discrete choice”, Working paper 2003

Rust, John (1994) “Structural Estimation of Dynamic Decision Processes”, Handbook of Econometrics IV, Chapter 51.

Rust and Phelan (1997): “Social Security in a World with Incomplete Markets”, *Econometrica*, Vol 65. No. 4, pp. 781-831

Stock and Wise (1990): “The Option Value Approach to Retirement”, *Econometrica*

Van der Klaauw, Wilbert and Kenneth I. Wolpin (2006): “Social Security and the Retirement and Savings Behavior of Low Income Households ”, mimeo, March 2006

Zaidi, Asghar (2006): “Pension Policy in EU25 and its Possible Impact on Elderly Poverty”, Policy Brief September 2006, European Centre

Økonomiske Råd, Det (2005): “Dansk Økonomi, forår 2005”.

Table 1: Wage regression

	Dependent variable log real hourly wage	
	Coefficient	T-stat
Experience	0,001494	(0.66)
Experience squared	-0,000168	(-3.84)
Tenure	0,0137	(10.37)
F-test (11729, 34312)	7,85	0,000
No. of observations	46045	
No. of individuals	11730	

Source: IDA and own calculations

Table 2: Descriptive statistics of estimation sample

	Mean	Std. Dev	min	max
Females	0,19	0,40	0	1
Single	0,12	0,32	0	1
log real wage	4,18	0,28	3,21	6,36
Age	60,41	5,17	50	71
Tenure	1,76	2,75	0	16
Experience	19,17	5,29	0	32
DHS	0,02	0,59	-2	2
Number of observations			7092	
Number of individuals			432	

Source: IDA and own calculations

Table 3: Most typical employment histories

Sequence	Per cent
FN	50,7
FWFN	14,8
FUN	6,1
F	3,9
FWFWFN	3,3
FWN	3,3
FWFWN	1,8
FUWFN	1,7
FWF	1,3
FWFUN	1,3
Other	11,8

F: Full time job, W: New full time job

U: Unemployment, N: Retirement

Source: IDA and own calculations

Table 4: Parameter estimates

	Unemployment	Retirement	Work	Layoff probabilities
ν			0.51 θ_1	1.83
			(1.12)	(4.31)
ρ	1.16	1.16	1.16 θ_2	0.07
	(6.23)	(6.23)	(6.23)	(0.40)
α^{single}	0.62	0.02		
	(1.4)	(0.01)		
α^{female}	0.70	0.15		
	(1.96)	(1.25)		
$\alpha^{\text{female*single}}$	-0.76	0.042		
	(-3.1)	(0.46)		
α^{constant}	-1	0.87		
	(-2.11)	(1.9)		
ϕ^{age}	1.75	1.75		
	(2.22)	(2.22)		
$l(h=w, h=r)$			-6	
			(-1.34)	
$l(h=w, h=u)$			-2.42	
			(-2.19)	
$l(h=w, h=w)$			-2.47	
			(-0.61)	
Number of observations		7092		
Loglikelihood		2239.9		
Note: T-stat in parenthesis				
Source: IDA and own calculations				

Table 5a: Employment transition probabilities, age 51-59

From	To	New job	Old job	Unemploy- ment	Retirement
New job	Actual	0,146	0,786	0,064	0,004
	Predicted	0,109	0,843	0,043	0,005
Old job	Actual	0,058	0,913	0,027	0,002
	Predicted	0,115	0,842	0,041	0,002
Unemployment	Actual	0,592	-	0,398	0,010
	Predicted	0,225	-	0,672	0,103
Retirement	Actual	0,714	-	0,143	0,143
	Predicted	0,477	-	0,502	0,021

Table 5b: Employment transition probabilities, age 60+

From	To	New job	Old job	Unemploy- ment	Retirement
New job	Actual	0,156	0,398	0,019	0,427
	Predicted	0,037	0,607	0,063	0,293
Old job	Actual	0,036	0,623	0,015	0,326
	Predicted	0,059	0,559	0,059	0,323
Unemployment	Actual	0,078	-	0,109	0,813
	Predicted	0,009	-	0,124	0,867
Retirement	Actual	0,000	-	0,000	1,000
	Predicted	0,000	-	0,001	0,999

Source: IDA and own calculations

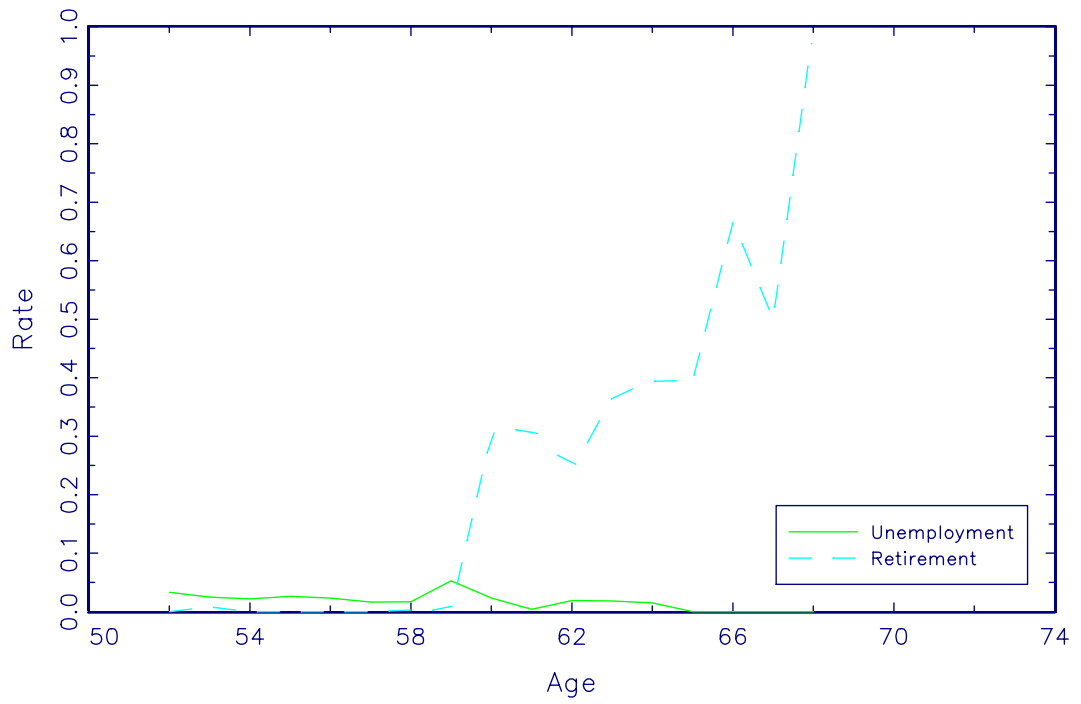
Table 6: Willingness to pay for removing search cost

	WTP for unemployed	Expected WTP	
		Lower bound	Upper bound
Male, single	51381	11251	30550
Male, couple	51952	7729	29147
Female, single	52149	11203	36864
Female, couple	51863	12061	34971

Source: IDA and Own calculations

GAUSS Sun Feb 11 19:35:49 2007

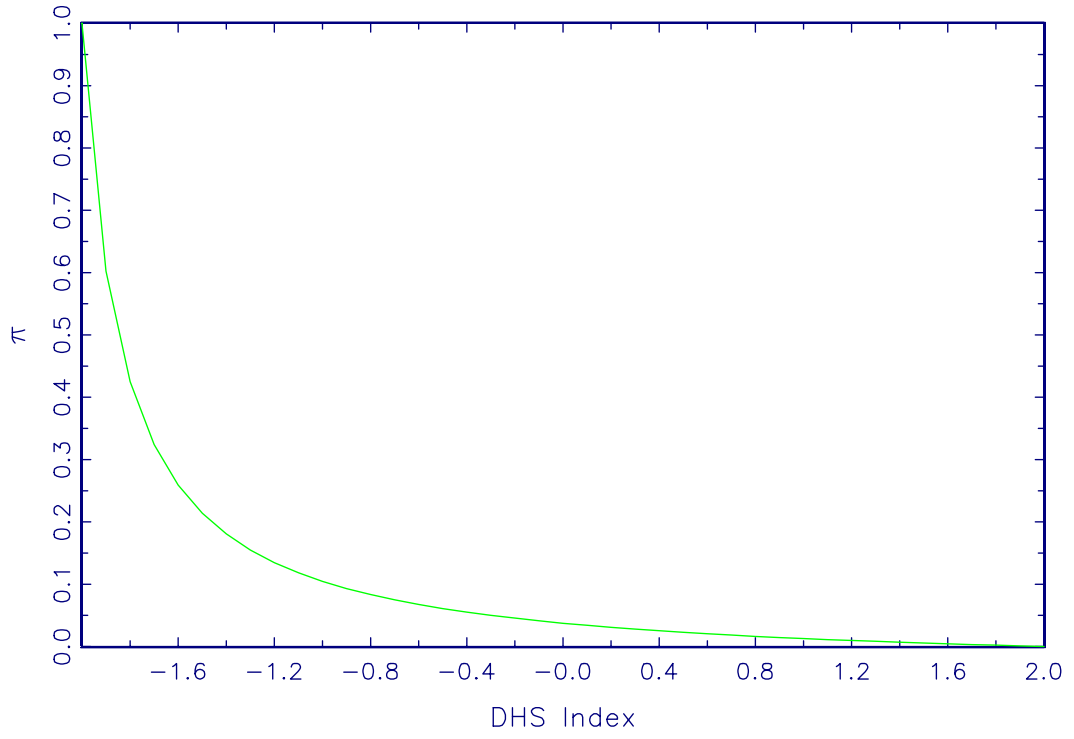
Figure 1: Transitions from work to unemployment and retirement



Source: IDA and own calculations

GAUSS Tue Mar 06 09:08:06 2007

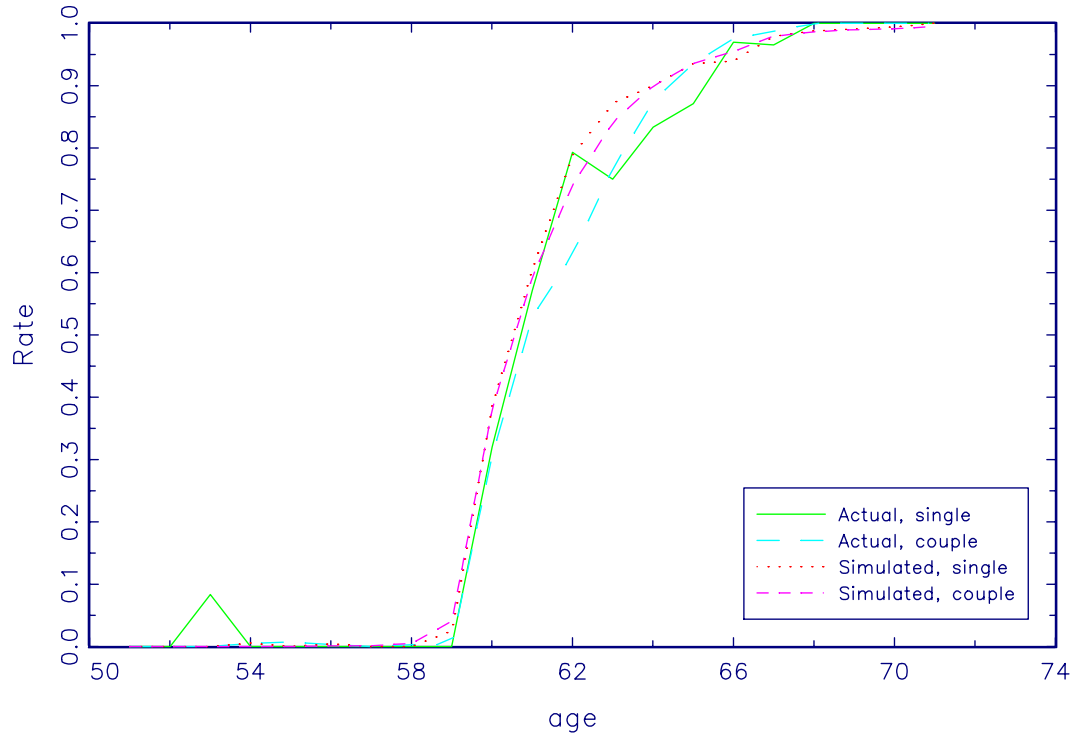
Figure 2: Layoff probability



Source: IDA and own calculations

GAUSS Tue Mar 06 09:08:37 2007

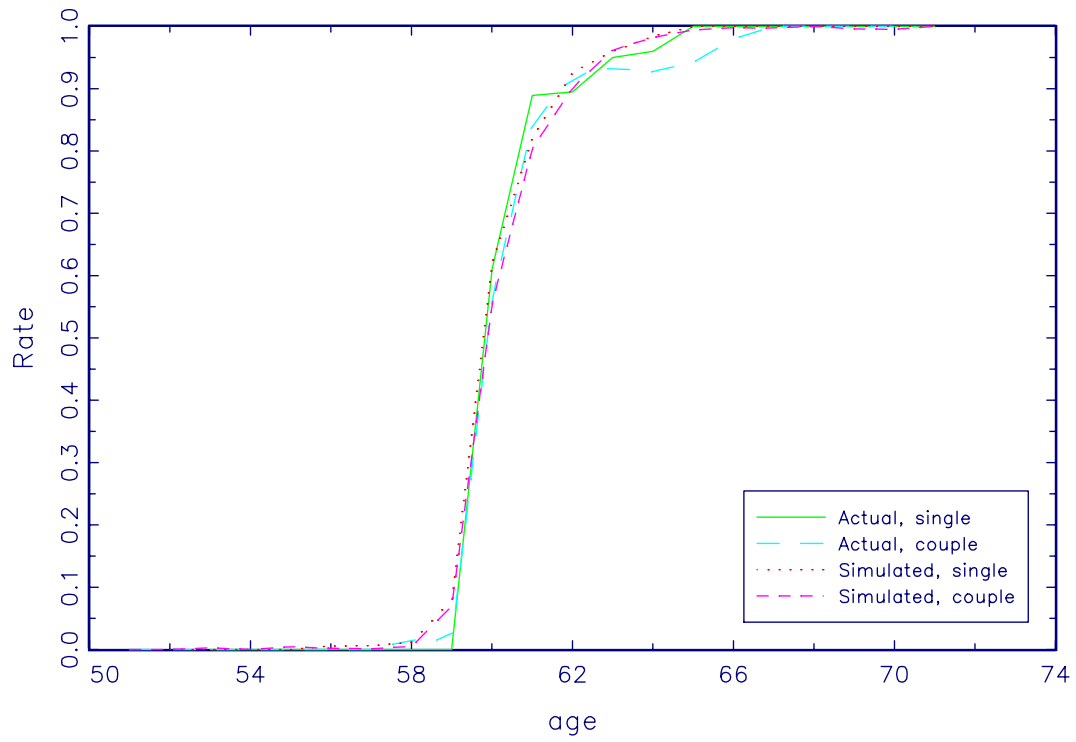
Figure 3a: Number of retired, Male



Source: IDA and own calculations

GAUSS Tue Mar 06 09:08:51 2007

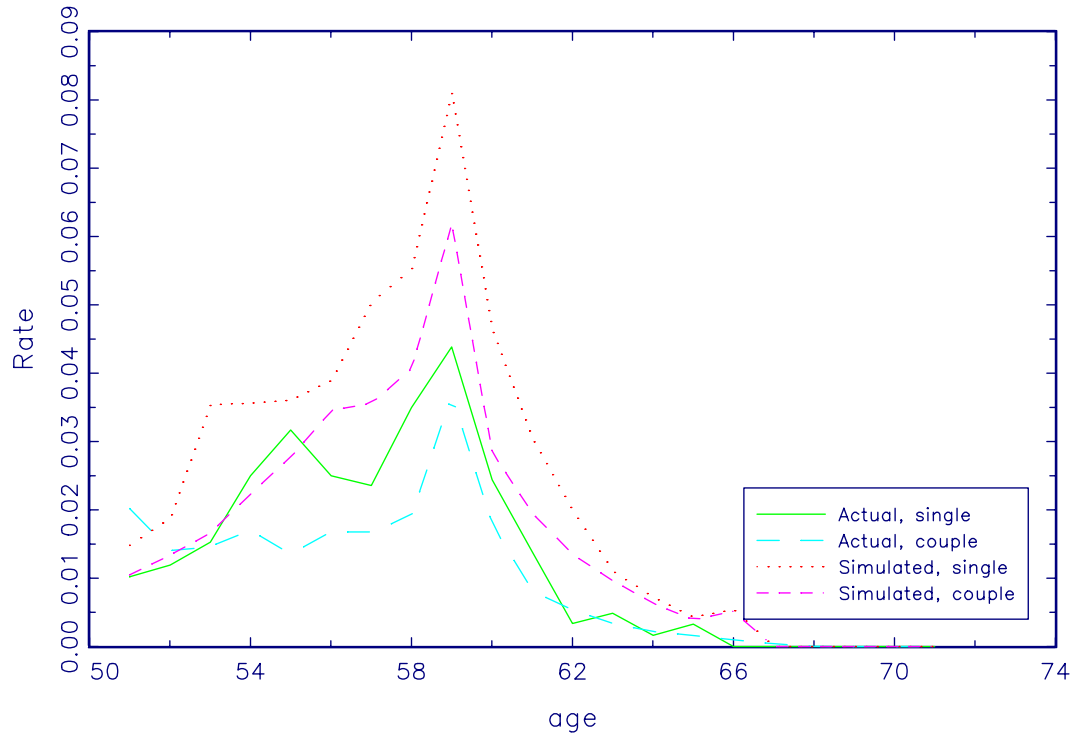
Figure 3b: Number of retired, Female



Source: IDA and own calculations

GAUSS Tue Mar 06 09:13:40 2007

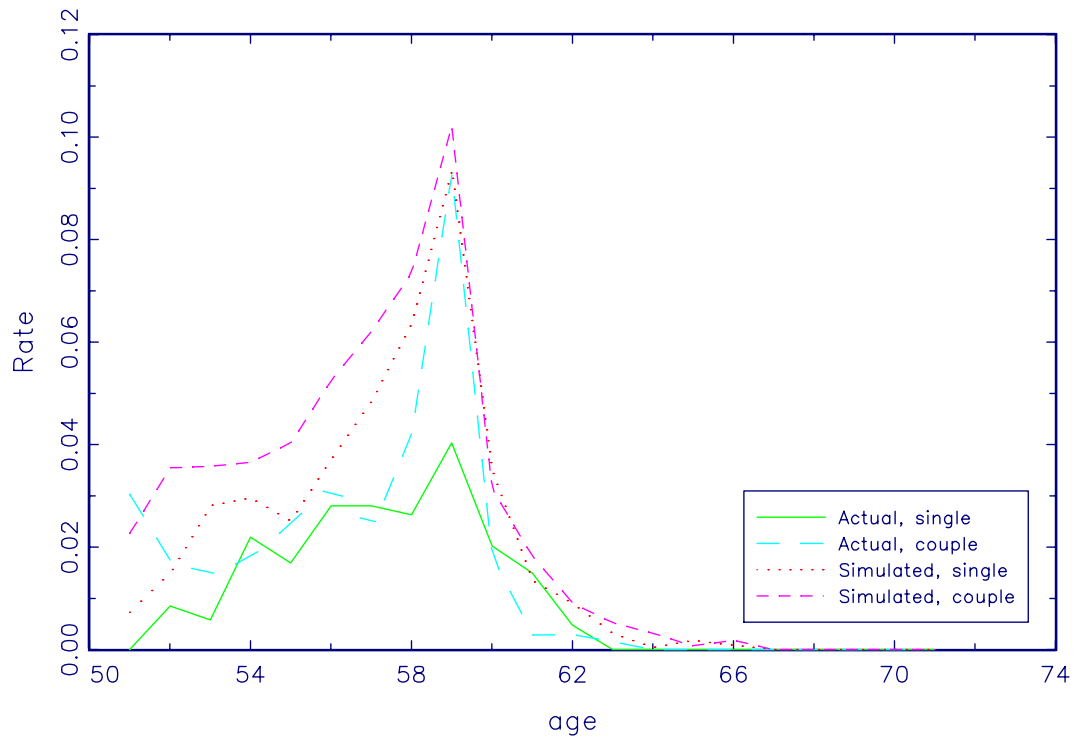
Figure 4a: Number of unemployed, Male



Source: IDA and own calculations

GAUSS Tue Mar 06 09:09:14 2007

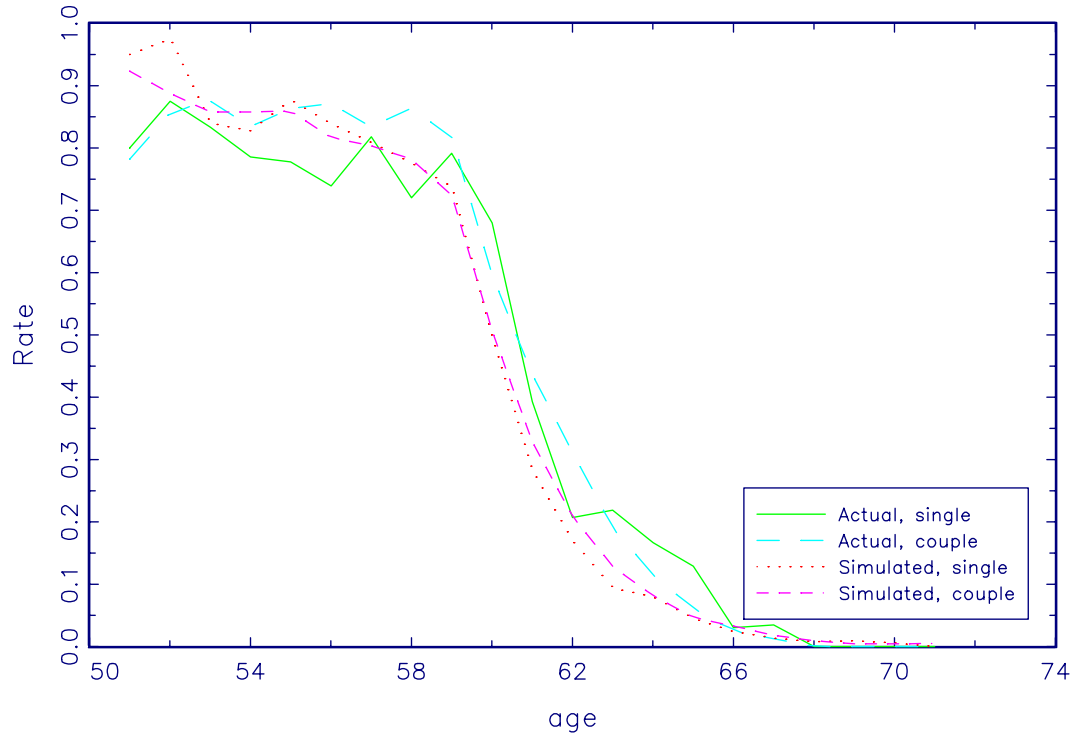
Figure 4b: Number of unemployed, Female



Source: IDA and own calculations

GAUSS Tue Mar 06 09:09:21 2007

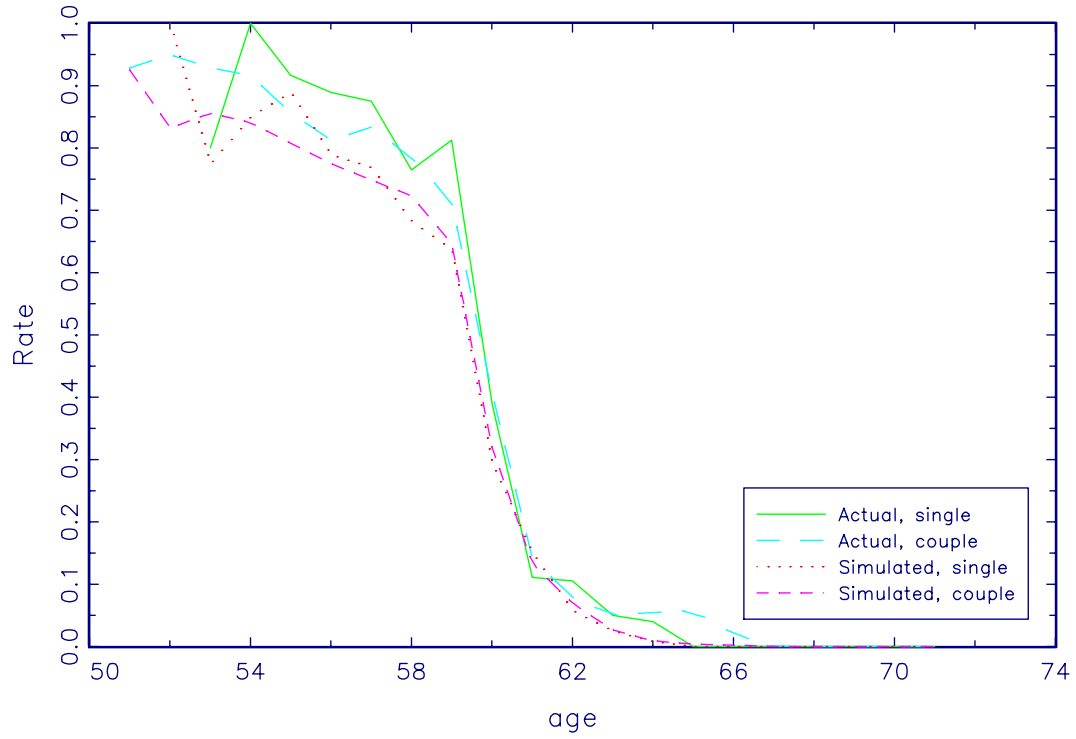
Figure 5a: Workers in old job, Male



Source: IDA and own calculations

GAUSS Tue Mar 06 09:09:33 2007

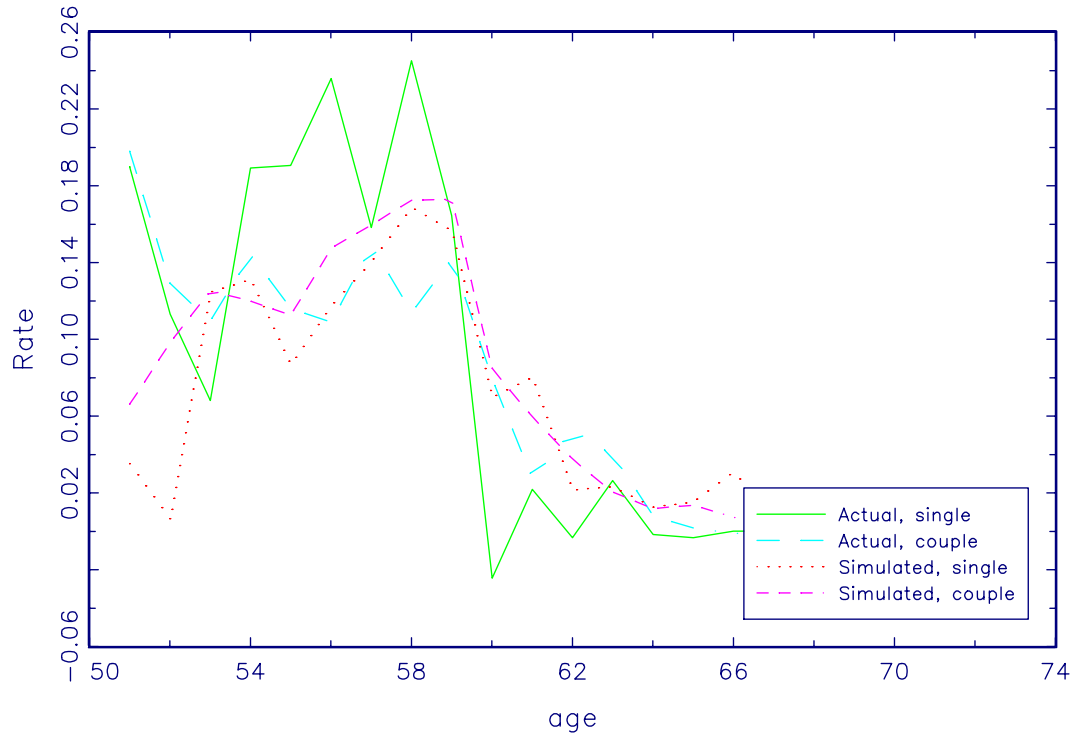
Figure 5b: Workers in old job, Female



Source: IDA and own calculations

GAUSS Tue Mar 06 09:09:41 2007

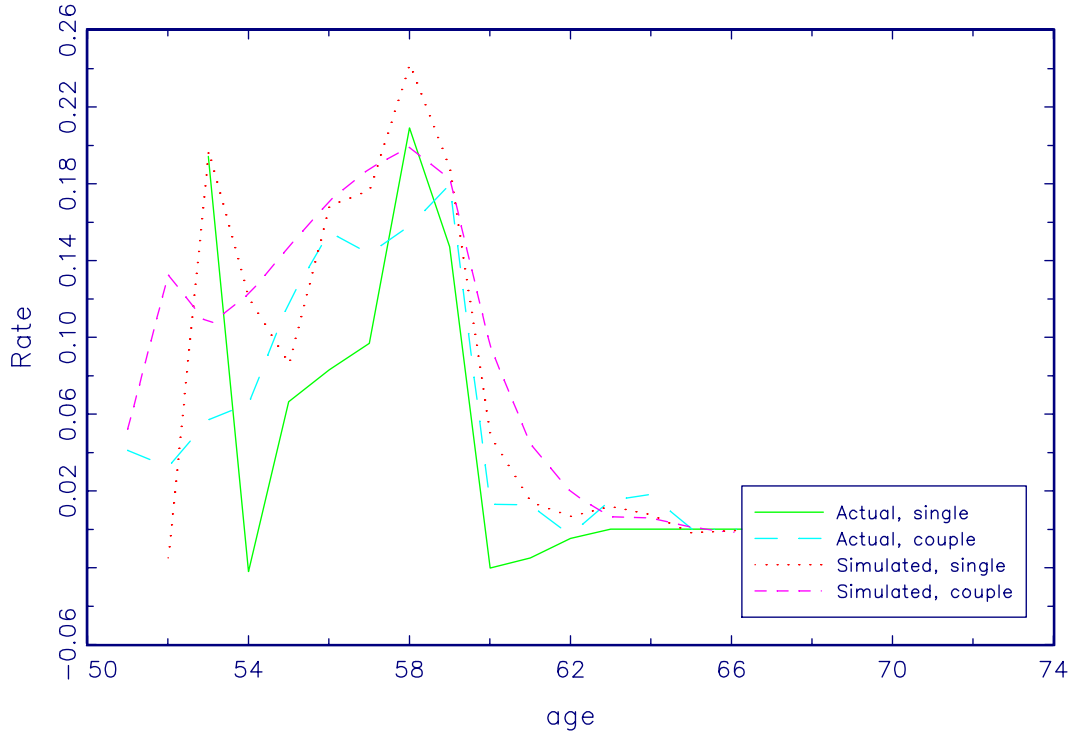
Figure 6a: Workers in new job, Male



Source: IDA and own calculations

GAUSS Tue Mar 06 09:09:53 2007

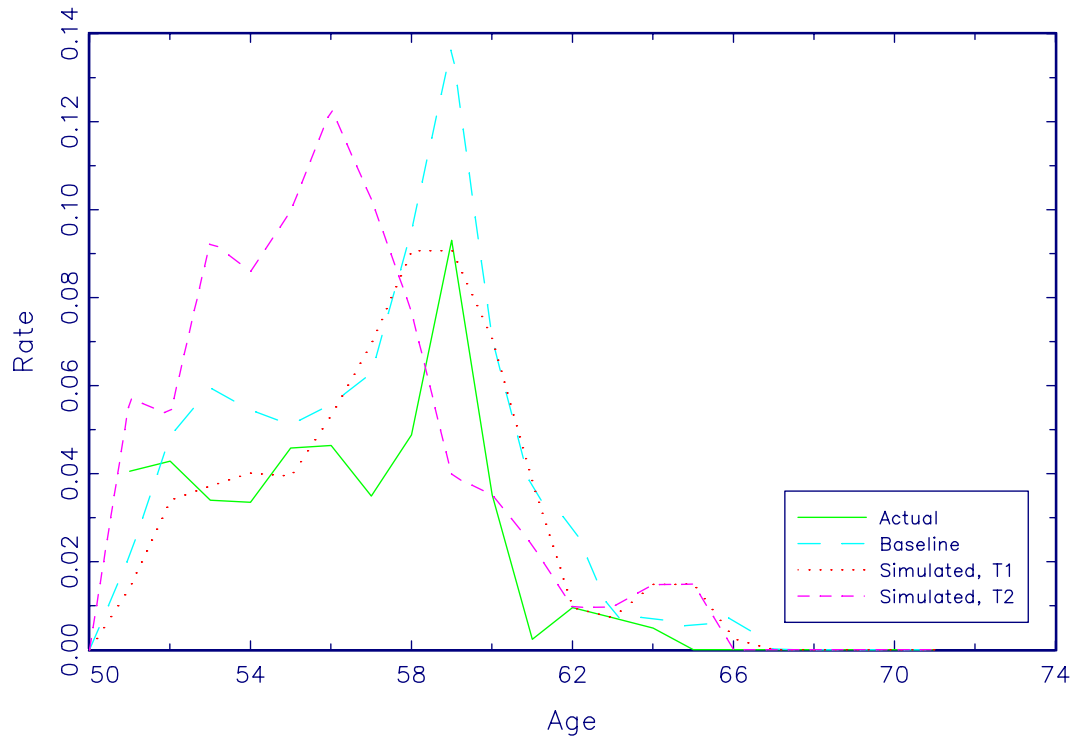
Figure 6b: Workers in new job, Female



Source: IDA and own calculations

GAUSS Tue Mar 06 09:10:42 2007

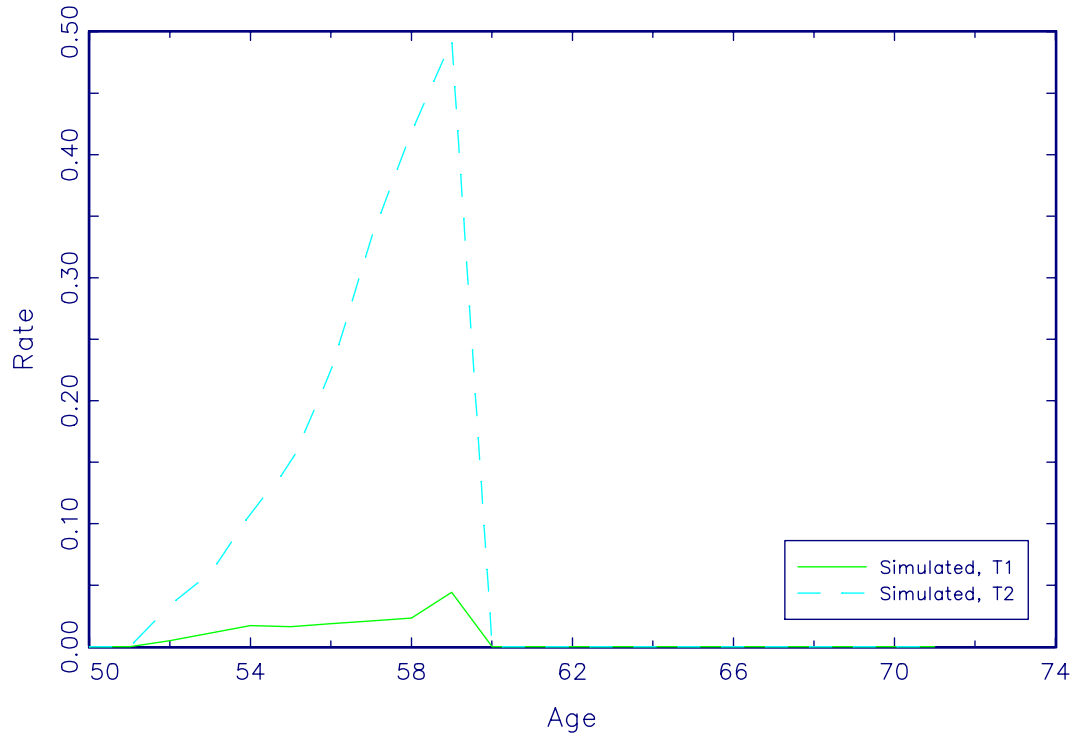
Figure 7: Unemployment



Source: IDA and own calculations

GAUSS Tue Mar 06 09:11:00 2007

Figure 8: Transitional allowance



Source: IDA and own calculations

Labour supply in an economy with a generous UI system.

Martin Browning Raquel Carrasco
Department of Economics, Carlos III, Madrid
University of Oxford

Martin Junge
CEBR, Copenhagen

February 2006

1 Introduction.

In Denmark in the 1980's, a low wage worker who had worked for six months was entitled to thirty months of Unemployment Insurance (UI) benefits at 90% of their wage. By any standards this seems likely to distort incentives. There are two major distortions as compared to the situation in which UI benefits and entitlement periods are much lower. First, workers who would normally work all the time have strong incentives to cut back work substantially. Indeed, if we allow for costs of going to work and some substitution between market goods and home production then many (low wage) workers who worked more than six months in any three year period were effectively paying to go to work (see Pedersen and Smith (1998)). Second, some people who would not normally be in the labour force now have a very strong incentive to work for six months to gain entitlement to thirty months of high benefit. In this paper we consider the effect of this system on labour market behaviour by analysing the labour force participation decisions of a sample of the Danish male population followed over an extended time period.

The emphasis in this paper is on the need to allow for considerable heterogeneity in the distaste for work that is correlated with productivity when considering the design of a UI system. Hamermesh (1980) tests the implications of a model that accounts for UI in the worker's labour supply decision. Using US data on married women, he focuses on the effects of higher benefits, an increased duration of benefits and easier eligibility criteria on average weeks worked and on the duration of unemployment. McCall (1996) examines the effects of UI on part-time work and finds a significant effect of disregard on the probability of part-time employment during the first three months of joblessness. Cullen and Gruber (2000) check whether unemployment insurance substitutes for the potential insurance role of female labor supply and find a strong crowding-out

effect. Beker, Gruber, and Milligan (2001) evaluate the effects of the Canadian income security system for retirement, and show that the work disincentives inherent in the system have large impacts on retirement. Saez (2002) analyzes optimal income transfers for low incomes and concludes that the nature of labor supply responses to taxes and transfers (intensity of work on the job or participation into the labor force) is critical to design optimal income transfer programs. Chone and Laroque (2005) study optimal taxation in an economy where the only choice of individuals is to work or to stay out of the labor force. Allowing for heterogeneity in productivity and tastes for work, they show how the optimal incentive schemes depend on the underlying structure of preferences. The optimal joint design of unemployment insurance and employment protection has also been studied by Blanchard and Tirole (2006); they show that layoff taxes should be different from unemployment benefits depending on certain distortions from a benchmark model, as for example the presence of unobserved heterogeneity across workers. Krueger and Meyer (2002) provide a survey on the empirical evidence on the labor supply effects of social insurance programs.

Our main empirical finding is that amongst prime age single men who have a low potential wage and a high disincentive to work, market work is the usual state. However, a substantial fraction do not participate at all (and consequently do not receive UI benefits). Finally, a small fraction seem to ‘play the system’ and alternate long periods of unemployment with short periods of work that are designed to re-invigorate UI entitlement. In a pure labour supply model with labour sold on a spot market and no human capital formation, the only way to rationalise this is by positing very substantial heterogeneity in preferences over work relative to consumption. Indeed, our analysis suggests that the distribution of ‘tastes for market work’ is bimodal. The largest group comprises those who will work in the market even when the financial incentives are low (they have nearly horizontal indifference curves in the usual labour supply/income picture). The second group are the polar group who will not participate in labour market work even for short periods with high remuneration (they have almost vertical indifference curves). Both of these first two groups display low labour supply elasticities. Finally, there is a third, smaller group who are in between and who display high labour supply elasticities. If this finding is correct and heterogeneity is as diverse and as biased towards work as we find, then this has strong implications for the design of UI systems. Very broadly, if a large fraction of people are willing to work even if the financial incentives are low and those who are out of the labour force are very reluctant to enter it, then some of the widespread fears concerning the disincentive properties of UI are unfounded. Our finding also calls into doubt the validity of findings for the design of UI programs that assume homogeneous attitudes to work (see, for example, Baily (1977), Gruber (1997), Hansen and Imrohoroglu (1992), Hopenhayn and Nicolini (1997) and Wang and Williamson (1996)).

In our empirical implementation, we consider discrete labour supply choices

for each month over a 36 month period.¹ We take a sample of low education single males for our empirical work. The selection of low educated workers (and consequently low wage) is to isolate those who are most likely to face very high disincentives to work due to the UI system. We consider only single males since our theory model has preferences over labour supply and consumption and the latter depends on other family members if we consider married males. In our analysis we allow for heterogeneity in the taste for work that may be correlated with the wage. We present results using a semiparametric estimator and a parametric estimator. For the former, we put enough parametric structure on individual preferences so that we can non-parametrically identify the distribution of tastes conditional on wages. This identification requires that workers can choose their desired labour supply with no constraints from the supply side and no work requirements from the UI authorities. Even though such a strong assumption is justifiable for environment for our data period, we recognise that actual labour supply may diverge from desired participation. One possibility is involuntary unemployment which gives that actual employment is below desired employment. We could also, however, have involuntary employment since the UI rules specify that workers should take a suitable job if offered one. Another source of involuntary employment is that a job taken to ensure qualification may last for more than six months. To allow for this we have to specify a parametric model with a parametric assumption on the conditional distribution of tastes for work. For this model we require that actual labour supply is close to desired labour supply.

Section 2 provides details of the Danish UI system as it operated in our sample period. We show that because of the UI system and the tax system, low wage workers faced very high implicit tax rates on work; for most workers in our sample this implicit rate was above 85%. Nonetheless, most workers worked most of the time. This is the most important finding in our study. However, some workers worked for much less than 36 months over our 3 year sample period and it is this distribution that we are interested in modelling. In section 3 we present a parametric intertemporal allocation model with flexible work and borrowing and saving between periods. We allow for non-additivities between consumption and labour supply but the main focus of interest is on an additive 'distaste for work' parameter which is allowed to be heterogeneous. We show that we can partially identify the distribution of this parameter conditional on wage if workers can achieve exactly their desired months of work.² In section 4 we present the fit of the parametric model to statistics that capture the marginal distribution of months of work and to statistics that capture the dependence of months of work on the wage. We do not need to present goodness of fit results

¹Our implicit time aggregation over a given 36 month period is imposed on us since the period by period dynamic program for choice with constraints depending on the choices in the preceding 36 periods (which is what is needed to model the Danish system) turns out to be very complex (see, for example, the program in Wang and Williamson (1996) for a four period model).

²Partial identification here means that for each worker we identify bounds on their 'distaste for work' parameter.

for the semiparametric model since it fits exactly by construction (that is, it predicts exactly the observed months of work for each observed wage). In section 5 we consider four hypothetical reforms to the UI system, keeping the tax system constant. These hypothetical changes either decrease the benefit period (holding the entitlement period constant) and/or reduce the benefit level. We present predictions for a range of outcomes including the unemployment rate, the incidence of long run unemployment and the cost of the UI system. For each of these reforms we have both parametric and semiparametric predictions. The semiparametric estimates are in the form of bounds; one feature of our predictions are that these bounds are very tight and usually contain the predictions from the parametric model. An important modelling conclusion we draw from this analysis is that allowing for non-additivities between consumption and labour supply makes very little difference to the results, so that in this context it may be reasonable to consider the simple case of an additive utility function. Our main substantive finding is that changes in the benefit rate have a much larger impact than changes in the entitlement period and a change to a ‘US’ style system (with six months of entitlement and a 45% replacement rate) would reduce unemployment rate for our sample of low education single males from the observed 19% to around 4%. A final section concludes.

2 Institutions and the distribution of work and wages.

2.1 The Danish UI and tax system.

The Danish UI system is a voluntary public system. It is administered by trade unions, but membership of a trade union is not a prerequisite for membership of the UI fund. Benefit levels, maximum duration and search requirements are regulated by the government and are identical across all UI funds. Despite being voluntary, 80% of the working population are members. The government makes a substantial contribution to the UI funds so that they are actuarially very attractive to most workers. In Denmark in the mid-1980’s eligibility for Unemployment Insurance (UI) benefit required 6 months of work within the last 36 months and membership of a UI fund. The replacement rate was 90 percent up to a maximum benefit level of DKK 104,766 in 1984. As we shall see shortly this maximum benefit cap is low and very few workers were entitled to the top replacement. In the 1980’s the Danish economy was in a recession and active labour market policies were in place so that individuals could stay entitled for UI benefit if they exhausted their benefit. A job offer was provided at the end of the UI cycle (The Economic Council, 2005). Anecdotal evidence also points to a very low search requirement for benefit receipt in this period.

The Danish tax system in our data period of 1984–1986 was individual (not family) based. There was a flat *local* tax of approximately 29%, which applied to all labour (including UI benefit) and capital income after an deduction of 14,000 DKK. State taxes were highly progressive. After a deduction, which is

Wage Percentile	Wage (DKK)	Wage (normalised)	UI Replacement rate
10%	7,365	2.78	85%
25%	8,018	3.02	78%
<i>Median</i>	9,007	3.40	69%
75%	10,378	3.92	60%
90%	12,464	4.70	50%

Table 1: Wages and UI replacement rates

of same magnitude as the local deduction, all income was taxed by the bottom rate of 14.4 percent. Income above 77,550 DKK were taxed by another 14.4 percent, so that the middle tax rate was 28.8%. The top tax rate was 10.8% and applied to incomes above 128,450 DKK. These rates, levels and brackets are taken as our policy parameters in the model, where we allow for the small year to year adjustment in the system from 1984 – 1986, a period without large reforms in this area.

2.2 The distribution of work and wages.

The data are compiled from various administrative registers in Denmark. In Denmark all individuals have a personal ID number used in all type of transactions: tax returns, social welfare, unemployment, job centre, banks etc. which allows us to construct a panel census. We have a 10% random sample in 1980 and follow individuals up to 1997. From this population we draw a very homogenous sample of workers who are most likely to be affected by the high replacement rates for UI (that is, low education workers) and for whom we can equate personal and household income (that is, single people). Specifically, we select single males with less than 10 years schooling who have no children (at least living at home) and who are aged between 25 and 55 in 1984. We also select on income from UI, social assistance and earnings being higher than 20,000 (1980) DKK in each year. We also removed self-employed and employees in the agriculture, fishing and mining as both groups have large changes in income from year-to-year and possibly very seasonal dependent labour supply. We also require that there is no missing observations on covariates from 1984 – 1986. We only select workers who had at least six months of work in the three years 1984 – 1986; this is partly to focus only on those who are eligible for UI and partly to make sure that we can compute a monthly wage for each worker. Finally, we trimmed the sample by removing individuals in top and bottom 1% in the wage distribution. Monthly wages are constructed from annual earnings divided by months in the labour market over the estimation period. The final estimation sample size is 2,046 individuals.

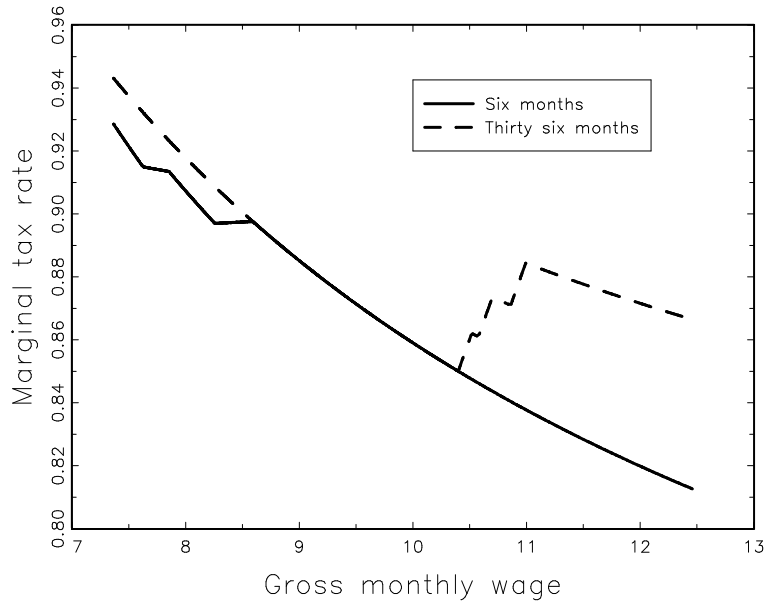


Figure 1: Marginal tax-benefit rates.

The distribution of wages and replacement rates for the sample is given in Table 1. The normalised wage is the wage divided by the monthly social assistance benefit that a single male would receive if they had no other source of income. One important feature the Table reveals is that the maximum UI benefit cap is very restrictive; in fact, about 4% of the sample are eligible for the maximum replacement rate of 90%. The return to working is partly determined by the UI system but also by the tax system. Figure 1 gives the marginal tax-benefit rates for the distribution of wages for a worker who works only six months of the 36 and a worker who works all of the time. The values of the marginal rates are striking: a low wage worker faces a marginal rate of about 90%; that is, for each extra 100 DKK earned by extra work, he sees an increase of only 10 DKK in net income. For such a worker this largely reflects that the UI replacement rate is about 89%. The upwards shift in the marginal rate for medium wage workers who work the full 36 months is where they move into a higher tax bracket.

Given such strong disincentives to work, it is of obvious interest to see how much people actually work. The distribution for months of work is given in Table 2. The most striking feature of this distribution is that 40% of workers work over the whole period and about two thirds work at least 30 months. On the other hand, there is a significant proportion who work less than half of the three year period. The Table also shows the mean and standard deviation of the log wage for each group. A clearer picture of the joint distribution of work

Months	Proportion (%)	Mean log wage	Std log wage
6 – 11	9.43	9.07	0.21
12 – 17	7.48	9.10	0.20
18 – 23	7.04	9.11	0.20
24 – 29	8.55	9.13	0.21
30 – 35	28.10	9.10	0.17
36	39.39	9.13	0.20

Table 2: Distribution of months

and gross wages is shown in figure 2; this is a nonparametric quantile regression with a uniform kernel. The wage values are trimmed to avoid end effects. Since over 40% of workers at any wage level work 36 months, we present only the bottom five deciles. As can be seen, for all groups the median labour supply is 34 months or above. For the lower deciles labour supply is first increasing sharply in wage and decreasing for higher wage workers. For the higher deciles there are small rise at the extremes of the wage distribution with a flat segment between. The purpose of our econometric modelling below is to account for this joint distribution using the institutional setting and the joint distribution of wages and the distaste for work.

3 Theory and identification.

3.1 Preferences over labour supply and consumption.

We assume that preferences over consumption and labour supply are intertemporally additive (where we take the period to be a month) and that labour supply in any period, d , is discrete with a value of 0 or 1. Let the within period utility function over consumption and work be denoted $u(c, d)$ where c is consumption. In considering a functional form for this utility function we need to satisfy a number of criteria. First, the dependence on consumption for fixed labour supply should not be too far from those assumed in the consumption and savings literature; this suggests an iso-elastic or translated iso-elastic form. Second, we do not want to unduly restrict the preference interactions between consumption and labour supply. Although the literature is not very clear on what the form of this dependence should be (see Browning, Hansen and Heckman (1999) for a review) we certainly want to allow for the possibility of costs of going to work and (Frisch) complementarity or substitutability between consumption and work. Third, we want to avoid parameterising the taste for work in such a way that it unduly affects the substitutability discussed in the last sentence; we shall return to this later. Finally, for empirical tractability, we want a form with a small number of unknown parameters. After considerable

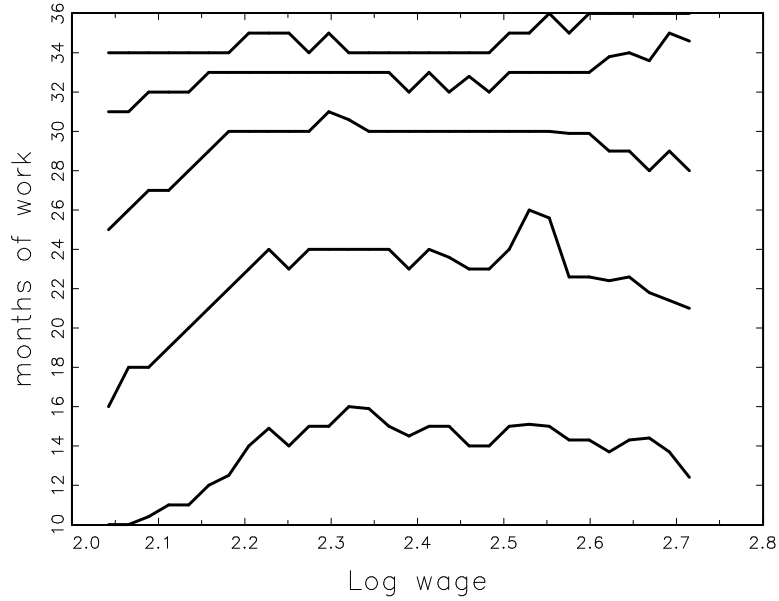


Figure 2: Bottom five deciles of months of work.

experimentation, we took the following within period utility function:

$$u(c, d) = \sigma^{-d\gamma} \left(\frac{(c - \tau d)^{1-\gamma}}{1-\gamma} \right) - \theta d \quad (1)$$

where $\gamma > 1$, $\theta \geq 0$, $\sigma > 0$ and $c > \tau > 0$. If the agent does not work in the period ($d = 0$) then the utility function takes the iso-elastic form:

$$u(c, 0) = \frac{c^{1-\gamma}}{1-\gamma} \quad (2)$$

In periods in which the agent does work the utility function for consumption is a translated power function:

$$u(c, d) = \sigma^{-\gamma} \left(\frac{(c - \tau)^{1-\gamma}}{1-\gamma} \right) - \theta \quad (3)$$

In terms of our analysis, the most important parameter is θ which feeds directly into the ‘distaste for work’ in the sense that higher values of θ denote a lower level of utility if working (holding consumption constant). The parameter τ captures any fixed costs of going to work. The parameter σ is required to allow flexibility in the link between preferences over consumption and labor supply. If we set $\sigma = 1$ and then intertemporal preferences over labour supply are linear

and additive between consumption and labour supply, which would violate the second criterion above. Setting $\sigma = 1$ and allowing $\tau > 0$ does give a non-additive form, but then all of the the interactions come only through a fixed cost, which seems restrictive. The parameter σ governs how complementary consumption is with labour supply, with values of $\sigma < 1$ giving higher complementarity. Thus different values of (σ, τ) allow that the preference interactions between consumption and labour supply can vary from low wage workers to high wage workers in a flexible way. The parameters σ and τ also partly govern the distaste for work but we do not want to use these parameters to have a primary role in this (that is, setting the additive distaste parameter, θ , to zero) since this would introduce a spurious connection between the relative consumption levels when employed or unemployed and tastes for work.

In our data we do not observe consumption in market work periods and periods of unemployment so we cannot hope to credibly estimate the coefficient of relative risk aversion parameter nor the parameters that control the complementarity of labour supply and consumption, (σ, τ) . There is no consensus in the literature on these parameters (see Browning, Hansen and Heckman (1999)) so we take ‘reasonable’ values that we use to assess the sensitivity of our estimates to the choice of these parameters. For the CRRA parameter we take a value of 3; this is a widely used value in the consumption literature. We set τ , the fixed costs of going to work, to be 20% of net income for the lowest wage worker in our sample. In our case this gives $\tau = 0.69$. Finally we choose the σ parameter to be 1.1 so that labour supply and consumption are (Frisch) substitutes, net of the fixed costs. expenditures in unemployed periods are about 10% lower than in employed periods for the highest wage person. Setting $\sigma > 1$ means that the mue in unemployed periods exceeds the mue in employed periods if there are no fixed costs of being employed ($\tau = 0$):

$$\begin{aligned} u(c, 0) &= -\frac{1}{2c^2} \\ u(c, 1) &= -\frac{1}{2\sigma^3(c-\tau)^2} - \theta \end{aligned} \tag{4}$$

which has three unknown parameters: (σ, τ, θ) . This utility function is strictly increasing and strictly concave in consumption, for fixed labour supply. Work and consumption are defined to be (Frisch) complements if $u_c(c, 1) > u_c(c, 0)$ (the marginal utility of consumption is higher when in work); a sufficient condition for this is $\tau > 0$ and $\sigma \leq 1$.

We now turn to intertemporal allocation. Consider a series of L periods (where L is relatively small so that, for ease of notation, we take the discount rate and the real rate to be equal to zero). The agent has income Y to allocate over the L periods and works for M periods and takes $L-M$ periods of non-work. An optimal program has the same level of consumption in all work periods, c_1 , and the same within all non-work periods, c_0 . It is more convenient to solve everything in per period terms so let $m = M/L$ (so that $0 \leq m \leq 1$) and

$y = Y/L$. Then the (mean) indirect utility function is given by:

$$V(y, m) = \max_{c_0, c_1} \left\{ \begin{array}{l} mu(c_1, 1) + (1 - m)u(c_0, 0) \\ \text{subject to } mc_1 + (1 - m)c_0 = y \end{array} \right\} \quad (5)$$

This indirect utility function and the functional form (4) provide the basis for our empirical analysis. The optimal values for c_0 and c_1 satisfy the usual Euler condition that the agent keeps the marginal utility of consumption constant³:

$$\begin{aligned} u_c(\hat{c}_1, 1) &= u_c(\hat{c}_0, 0) \\ \Rightarrow \hat{c}_0 &= \sigma(\hat{c}_1 - \tau) \end{aligned}$$

Substituting in the budget constraint from (5) we have:

$$\begin{aligned} \hat{c}_0 &= \frac{\sigma y - m\sigma\tau}{m + (1 - m)\sigma} \\ \hat{c}_1 &= \frac{y + (1 - m)\sigma\tau}{m + (1 - m)\sigma} \end{aligned} \quad (6)$$

Note these consumption levels are independent of the additive parameter θ , which fulfills the third of our criteria listed above. If both utility functions are iso-elastic ($\tau = 0$) then workers spend more in work periods than in non-work periods if and only if $\sigma < 1$. The associated indirect utility function over income and labour supply is given by:

$$\begin{aligned} V(y, m) &= m \left\{ \frac{-1}{2\sigma^3(\hat{c}_1 - \tau)^2} \right\} + (1 - m) \left\{ \frac{-1}{2(\hat{c}_0)^2} \right\} - m\theta \\ &= \left\{ \frac{m + \sigma(1 - m)}{\sigma^3} \right\} \left\{ \frac{-1}{2(\hat{c}_1 - \tau)^2} \right\} - m\theta \end{aligned} \quad (7)$$

where \hat{c}_0 and \hat{c}_1 are given by equation (6). Note that if we take the additive case with $\sigma = 1$ and $\tau = 0$ then we have:

$$V(y, m) = \left\{ \frac{-1}{2y^2} \right\} - m\theta \quad (8)$$

The indirect utility function given in (7) is analogous to the utility function taken in ‘static’ labour supply analysis which is defined over income and labour supply. Note, however, that our form allows for many periods and also allocation between different blocks of L periods if y includes saving or dissaving between blocks.

When considering the choice of the proportion of time to work, m , we need to allow that net income depends on m . Denote this by $y = y(m)$. This function

³Assuming that the interest rate and the discount rate are equal and that the agent can save and borrow at the fixed interest rate.

depends on the gross wage, the tax system and the UI system. The optimal level of labour supply is given by:

$$\hat{m} = \arg \max_m \{V(y(m), m) \text{ subject to } 0 \leq m \leq 1\} \quad (9)$$

If the net income function $y(m)$ is continuous, a maximum always exists. If $y(m)$ is smooth and the optimum is interior ($0 < m < 1$) then this gives the conventional first order conditions:

$$-\frac{V_m}{V_y} = y'(m) \quad (10)$$

where the left hand side is the marginal rate of substitution and $y'(m)$ is the ‘marginal real wage’. In general, however, the mapping from work to income will be non-differentiable so that we need to take into account ‘mass points’ in our empirical work; we return to this in the next subsection.

Although our parameterisation has been chosen according to a number of criteria, the values of θ do not have any immediate interpretation. We thus define the *distaste for work*, δ , to be the proportional increase in income that is needed if the agent is moved one percent away from their preferred labour supply:

$$\delta = \frac{d \ln y}{d \ln m} \Big|_V = \left(\frac{dy}{dm} \Big|_V \right) \frac{m}{y} = -\frac{V_m}{V_y} \frac{m}{y} \quad (11)$$

If we have an interior solution then equation (10) gives:

$$\delta = y'(m) \frac{m}{y} \quad (12)$$

For corner solutions, we must use (7). If preferences are additive ($\sigma = 1, \tau = 0$) the distaste for work is given by:

$$\delta = \theta m y^2 \quad (13)$$

In this case, the distaste for work is increasing in work and income. More importantly, the distaste for work is increasing in the parameter θ . For the general case, δ is affine in θ with a positive slope (so long as τ is small relative to y).⁴

3.2 Identification.

We now turn to the relationship between choices made using (5) and heterogeneity in the taste for work in the presence of a non-linear transfer system. We consider a transfer system with a cycle length of 36 months and a UI entitlement period of 6 months. That is, the agent is eligible for UI if they have worked 6 of the previous 36 months. Let w be the per period gross monthly earnings

⁴The explicit equation is too long to present here; the calculations leading to this conclusion were made with Maple.

(‘wage’) in work, ρ be the UI replacement rate (so that $\rho \in [0, 1]$). We have the following mapping from the number of periods of work to net income for everyone with $M \geq 6$:

$$Y = \Psi(w(M + (36 - M)\rho)) \quad (14)$$

where $\Psi(\cdot)$ is the tax system that converts gross earnings over L periods into net income.⁵ Consider, for example, a worker who has enough months of work to entitle them to UI. For this worker we assume that the wage is constant over the L periods and the value, w , is equal to the mean observed gross wage (that is, gross labour income divided by M). Our goal is to find the distribution of heterogeneity conditional on wages, $f(\theta|w)$. Once we have this then we can compute outcomes for any new policy (for UI or for the tax system or for a combination) that maps work and gross wages to net income. In our empirical work we take the Danish tax and benefit system described in section 2. This gives a net income function that is strictly increasing and continuous but non-differentiable in gross income for a given wage.

We adopt two approaches to estimating the preference parameters, one fully parametric and the other semiparametric. The advantage of the semiparametric approach is that it fits the data exactly. The disadvantages are that it requires stronger assumptions concerning the slippage between desired work and observed work (details are given below) and it only gives intervals for outcomes in out hypothetical policy simulations. The advantage of the fully parametric approach is that it gives a functional form for the dependence of taste parameters on wages which allows us to conduct out of sample predictions. The disadvantage is that we have to make assumptions about the functional form of $f(\theta|w)$.

We use somewhat different identification assumptions for the two approaches. Denote the observed proportion of months of work for worker h by m_h^* . The model in the previous subsection gives desired labour supply, \hat{m}_h , given a wage w_h , the tax and benefit system ($\Psi(\cdot)$ and ρ) and the utility parameters (σ, τ, θ) . For the semiparametric estimation scheme our principal assumption is a strong one that that agents are unconstrained in their choice:

Axiom 1 $m_h^* = \hat{m}_h$.

We do not have direct evidence for or against the assumption that workers could work a good deal less than the maximum if they wished. We note, however, that over the several years in which the UI scheme in Denmark took the form described here, many workers did achieve close to a 6/30 work/UI plan over several years (so that it was feasible for some workers). Moreover, anecdotal evidence suggests that work tests were not enforced with any rigour for our selected sample of low education, low wage males. More importantly for research purposes; it is of interest to consider the benchmark case in which work requirements do not have any effect, so that moral hazard is maximised.

⁵When we compute this in our empirical work we take the most advantageous patterns for labour supply for net income purposes. Since workers face a progressive tax system this usually means spreading the periods of unemployment across the three tax years.

The assumption that the observed choice is the optimal choice ($m_h^* = \hat{m}_h$) rules out frictions such as involuntary unemployment ($M_h^* < \hat{m}_h$). It also rules out any ‘availability for work’ test for unemployed workers. If workers have to take jobs and these are available then we may have $m_h^* > \hat{m}_h$. In our fully parametric estimation approach we make limited allowance for such frictions. Specifically, our identifying assumption is that if a worker wanted to work a good deal less than the maximum ($\hat{m} \ll 36$) then he could do so ($m^* \ll 36$), even if the actual work is not quite equal to the desired value. Conversely, if the worker wished to work the maximum ($\hat{m} = 36$) then he could find enough work to be close to this. This is captured by

Axiom 2 *The agent is approximately unconstrained in their choice: $m_h^* \simeq \hat{m}_h$.*

We shall make clear the approximation in the empirical section, but it allows for involuntary unemployment ($\hat{m}_h = 36$ and $m_h^* < 36$) and hence gives a rationale for the insurance function of the UI system.

We now consider nonparametric identification of the distribution of the principal parameter which governs labour supply, θ . Since we are considering a discrete choice model ($M = 6, 7 \dots 35, 36$) we shall only be able to identify sets for the latent parameters. We use the indirect utility function given in (7), $V(y, m; \theta)$ where we show the explicit dependence on the preference parameter θ . For a given pair of preference parameters (σ, τ) and a given wage w_h we define a rationalising θ value, $\hat{\theta}_h$, by:

$$\arg \max_m V(y, m; \hat{\theta}_h) = m_h^* \quad (15)$$

For various values of (σ, τ) , w_h and the tax and benefit system, there may not be any values of θ which satisfy this condition. In general it is difficult to write down sufficient conditions for $\hat{\theta}_h$ to exist. In our empirical work we found that for the additive case ($\sigma = 1$ and $\tau = 0$) the sets of $\hat{\theta}_h$ were always non-empty for every (w_h, m_h^*) . Even more conveniently, these sets always turned out to be intervals of the form $(\theta_h^{\min}, \theta_h^{\max})$. If the worker is in work for the whole three years ($m^* = 1$) then $\theta_h^{\min} = 0$. Conversely, if the worker is only employed for 6 months (the minimum in our selected sample) then we have $\theta_h^{\max} = \infty$.

For the fully parametric estimation we shall take a parametric form for $f(\theta|w)$. Specifically we assume that the distribution is a mixture of two logNormals in which both the means and variances of the component logNormal and the mixing parameter depend on the log wage.⁶ Specifically we take the

⁶We tried many other distributions (for example, the generalised gamma). The distribution used here gave the best fit. A single logNormal is decisively rejected by the data.

following 10 parameter conditional distribution:

$$\begin{aligned}
\theta^1 &= \exp(\pi_{11} + \pi_{12} \ln(w) + \exp(\pi_{13} + \pi_{14} \ln(w))U) \\
\theta^2 &= \exp(\pi_{21} + \pi_{22} \ln(w) + \exp(\pi_{23} + \pi_{24} \ln(w))U) \\
\phi &= \frac{\exp(\pi_1 + \pi_2 \ln(w))}{1 + \exp(\pi_1 + \pi_2 \ln(w))} \\
\theta &= \theta^1 \text{ with probability } \phi \\
&= \theta^2 \text{ with probability } (1 - \phi)
\end{aligned} \tag{16}$$

where U is a standard Normal random variable.

4 Results for the parametric model.

We estimate the parametric model above using Simulated Minimum Distance (SMD)⁷. To implement this estimator we need to specify auxiliary parameters (ap's) that are used to match the observed data and simulated data. We have two sets of ap's. The first set of 5 ap's describes the univariate distribution of the months of work. Specifically we take the proportions with months in the intervals $[6, 12)$, $[12, 18)$, $[18, 24)$, $[30, 36]$ and the aggregate unemployment rate for our sample over the three years. The focus on intervals allows for frictions as in assumption 2. In particular, pooling all those who have between 30 and 36 months of work allows for short spells of involuntary unemployment amongst those who would prefer to work all the time. The second set of 6 ap's captures the dependence of the first two moments of the months of work on the wage. Specifically we regress months on a constant, the log wage (centred on zero) and the squared log wage and record the three parameters, denoted $(\beta_1, \beta_2, \beta_3)$. We then take the OLS residuals, square them and regress them on the same three variables to capture the conditional heteroskedasticity; we denote these three coefficients by $(\delta_1, \delta_2, \delta_3)$. We use a nonparametric bootstrap to calculate an estimate of the covariance matrix for these ap's. In Table 3 we present the means and the 95% confidence bands for the ap's. As can be seen, the regression results indicate that the mean months are increasing in the wage ($\beta_1 < 0$) and the dispersion is decreasing ($\delta_1 > 0$).

As discussed in the previous section, the conditional distribution of the dis-taste for work parameter, θ , is modelled as a mixture of two logNormals. For SMD we have to replicate the H workers in our sample. This involves a trade-off between precision and speed. Since the simulation involves a maximisation step for each worker it is relatively slow and we take only one replication with an antithetic second draw. To do this we first draw two H -vectors of pseudo-random

⁷This estimator is often known as 'indirect inference' (Gouriéroux, Monfort and Renault (1993), Gouriéroux, C. and A. Monfort, (1996)). We take the more descriptive term SMD used in Pakes and Pollard (1989) and Hall and Rust (2003).

numbers:

$$\begin{aligned} u &\sim N(0, 1) \\ s &\sim U[0, 1] \end{aligned} \tag{17}$$

We then take a second replication with antithetic values $-u$ and $(1 - s)$. Thus we have $2H$ workers in our simulated data set. Taking antithetic draws is a conventional procedure to reduce the variance of the estimator and to stabilise the optimisation routine when we have a small number of replications. For given parameters and wage, we simulate the distribution given in (16) by:

$$\begin{aligned} d &= 1 \text{ if } \frac{\exp(\pi_1 + \pi_2 w)}{1 + \exp(\pi_1 + \pi_2 w)} > s \\ &= 0 \text{ otherwise} \end{aligned} \tag{18}$$

$$\begin{aligned} \theta &= d * \exp(\pi_{11} + \pi_{12} \ln(w) + \exp(\pi_{13} + \pi_{14} \ln(w))u) + \\ &\quad (1 - d) * \exp(\pi_{21} + \pi_{22} \ln(w) + \exp(\pi_{23} + \pi_{24} \ln(w))U) \end{aligned} \tag{19}$$

where w now denotes the log wage. Given values for (σ, τ) we use this θ value in the indirect utility function ((7) and (6)) and find the optimal months of work for a given worker (w, θ) . We then calculate the ap's for the $(2H \times 2)$ matrix of wages and months. Since we have 11 ap's and 10 parameters we have one degree of overidentification.

The last two columns of Table 3 present the fit of the parametric model with consumption and labour supply additivity imposed $((\sigma, \tau)=(1, 0))$ and with the non-additive model with $(\sigma, \tau)=(1.1, 0.69)$. The most important feature of the two sets of estimates is that the additive model fits very well and the non-additive model less well, but still adequately.

The parameter values for the two sets of estimates are of little interest and we do not present them. Of more interest is the distribution of θ . The marginal distribution for the additive model is shown in figure 3. As can be seen, the two mixing distributions are very different. The first has mixing distribution has a high mode and a low variance whereas the second has a low mode and a high variance. The consequent mixed distribution is close to being bimodal which 'explains' the decisive rejection of the single logNormal model. The probability of being in the first distribution is sharply increasing in the wage so that low wage workers have θ 's that are drawn from the more dispersed mixing distribution.

5 Policy simulations

The principal virtue of a structural approach is that we can consider the effects of policies that have never been implemented. In this section we consider three alternative policy scenario. In these we keep the qualifying period at six months and vary the benefit period and the maximum replacement rate. The policies are:

Thus in policy 1 we leave the benefit level unchanged but now a worker is entitled to only 12 months of UI benefit for each six months of work. Policies

Auxiliary parameter	Mean	Confidence interval		Parametric model fit	
				Additive	Non-additive
[6, 12) months	10.1	8.8	11.3	9.1	11.2
				[1.0]	[-1.3]
[12, 18) months	7.3	6.3	8.4	6.6	8.4
				[1.0]	[-1.5]
[18, 24) months	7.3	6.2	8.4	7.7	6.6
				[-0.6]	[0.9]
[30, 36] months	66.3	64.3	68.3	67.0	65.5
				[-0.5]	[0.5]
Unemployment rate	19.1	17.9	20.2	18.2	20.1
				[1.0]	[-1.2]
β_1	-73.0	-113	-32.5	-88.0	-128
				[0.5]	[1.9]
β_2	86.5	51.1	122	99.9	135
				[-0.5]	[-1.9]
β_3	-18.0	-25.8	-10.4	-20.9	-28.6
				[0.5]	[1.9]
δ_1	12.6	8.2	17.	12.7	17.7
			0	[-0.0]	[-1.7]
δ_2	-9.8	-13.6	-6.0	-9.9	-14.0
				[0.0]	[1.5]
δ_3	2.0	1.2	2.8	2.0	2.9
				[-0.0]	[-1.5]
$\chi^2(1)$ OI test	-	-	-	4.4	9.9

The value in [.] is the t-value for the difference of the fitted value from the observed value for the auxiliary parameter.

Table 3: Auxiliary parameters and parametric fit

Policy	Benefit period	Replacement rate
Current	30 months	90%
1	12 months	90%
2	12 months	60%
3	6 months	45%

Table 4: Alternative policies

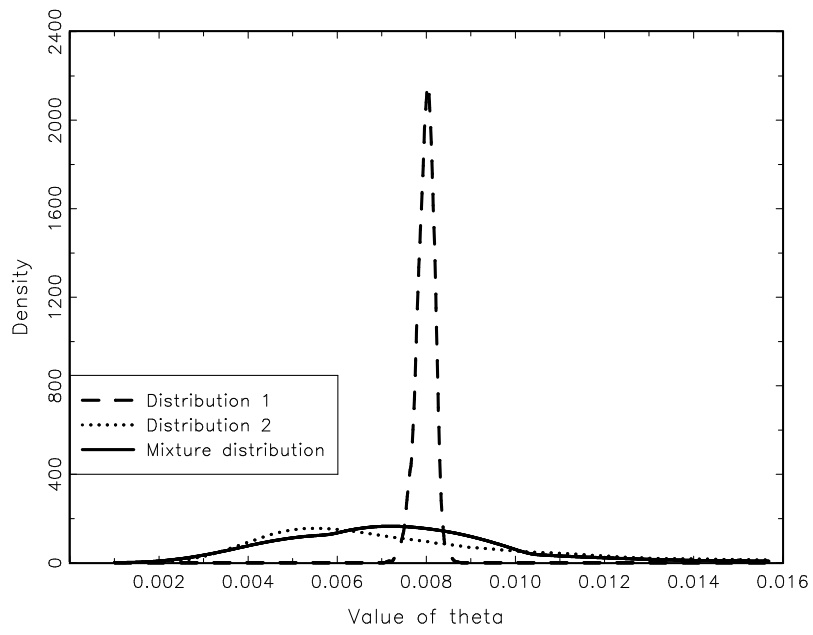


Figure 3: The distribution of θ .

1 to 3 are increasingly less 'generous' (policy 4 is akin to the US UI system). We shall present three sets of predictions. The first are the bounds from the partially identified additive model. The second and third are the point predictions from the additive and non-additive parametric models, respectively. We present results for six outcomes: our choice of outcome variables is motivated partly by the perceived policy concerns of finance ministers and those concerned with the functioning of the labour market and partly by the welfare of the workers themselves. Specifically we consider: the unemployment rate for this group of workers; the proportion of long term unemployed amongst the unemployed⁸; UI benefits paid; gross earnings; net income (including UI benefits) accruing to the sample and the total taxes paid. The units for the last four of these are all in millions of 1980 DKK but only their relative magnitudes are of any interest. It would be desirable to consider a broader social welfare function for this group, but that would require taking a stand on how changes in government net receipts (taxes minus UI benefits paid) are re-distributed to the group, which we are reluctant to do since this requires considering changes in tax policy.

Table 5 presents the predictions. For each outcome there are three rows. The first gives the bounds from the partially identified additive model and the second and third rows give the point predictions from the parametric models. The first column gives the 'predictions' for the current policy; for the partially identified model this is, of course, the actual value calculated from the data since this model fits perfectly whether we take the upper or lower bounds on θ . One general point to note about the predictions in Table 5 is that for the bounds from the partially identified models are very tight; this is in contrast to many partially identified models. A second point to note is that the estimates for the non-additive parametric model are generally quite close to those for the additive parametric model. This suggests that the predictions if we make 'large' changes to the UI system are insensitive to 'sensible' assignments of values for the 'nuisance' preference parameters (that is, σ and η in 1). Finally, the predictions from the additive parametric model are usually outside the bounds for the partially identified additive model but this is largely because the predictions for the current policy are different.

Turning to policy 1 we see that shortening dramatically the benefit period (from 30 months to 12 months) has little impact on any of the outcomes. Indeed, the proportion of long term unemployed is unaffected since the proposed policy change impacts only those who take at least 12 months of unemployment under the current policy. The most significant change is in the unemployment rate and, consequently, the UI benefits paid out. Policy 2, which lowers the replacement rate as well as the benefit period, has much more impact. The unemployment rate is cut by two thirds and UI benefits paid are massively reduced. Moreover, gross earnings increase by about 10% which represents a considerable increase in output. On the other hand, net income for this group of workers does not rise very much so that many of them would be worse off under this new policy, some

⁸Strictly, the proportion who have 12 months or more unemployment relative to the proportion who have at least one month of unemployment.

Outcome	Policy			
	Current	1	2	3
Unemployment rate (%)	19.1	[18.1, 18.3]	[5.5, 6.2]	[2.3, 3.2]
	18.2	17.3	5.9	2.6
	20.1	19.0	7.4	2.4
Proportion of long term unemployed (%)	41.9	[41.9, 41.9]	[38.3, 39.0]	[35.2, 38.0]
	45.5	45.5	27.1	24.1
	50.1	50.1	33.3	22.5
UI benefits paid	100	[95.9, 95.9]	[29.0, 32.8]	[12.2, 16.7]
	100	95.0	32.3	14.1
	100	94.6	36.9	12.1
Gross earnings	100	[101, 101]	[114, 115]	[118, 119]
	100	101	113	118
	100	101	113	121
Net income	100	[100, 100]	[101, 102]	[102, 103]
	100	100	102	102
	100	100	102	103
Total taxes paid	100	[100, 100]	[104, 104]	[107, 107]
	100	100	104	106
	100	100	104	107
<p>The first row for each outcome are the bounds from the additive partially identified model. The second row is for the additive parametric model. The third row is for the non-additive parametric model.</p>				

Table 5: Policy outcomes

considerably so.⁹ Moving to something close to a US system, policy 3, leads to even lower unemployment and UI benefits with an even greater increase in gross earnings. At this point it may be thought that we have moved too far ‘outside sample’ and that the values predicted are implausibly low. Finally, policy 4 goes the opposite direction to the other three and makes the current system even more generous in terms of the benefit rate. As can be seen, the predicted results are catastrophic. Almost half of workers simply work the minimum 6 months and output (gross earnings) collapses. Of course, no one is worse off under the new policy but this is at considerable cost to whoever is financing the UI system (as can be seen taxes raised from this group do not even cover payments made through UI).

6 Conclusions.

We have investigated the supply of labour if agents face the very ‘generous’ UI system that existed in Denmark in the mid-1980’s. Our most important finding is the simplest: despite facing very high implicit tax rates, most of our sample of low educated single males work most of the time. Even so, the average unemployment rate for this group over the period 1984 – 1986 is 19%. We present a parametric and semiparametric analysis of the data that allows for heterogeneity in the tastes for work that is correlated with the wage. Given these estimates we simulate the responses to a series of decreasingly generous hypothetical policies. We draw the same conclusions from all of our estimates: the outcomes of interest are much more sensitive to cuts in the replacement rate than to decreases in the entitlement period. Introducing a US style system would have cut unemployment in our group of low educated single males from 19% to around 3%.

References

- [1] Baily, M.N. (1977), “On the Theory of Layoffs and Unemployment”, *Econometrica*, 45, 1043-1064.
- [2] Baker, M., J. Gruber, and K. Milligan (2001), “The Retirement Incentive Effects of Canada’s Income Security Programs”, NBER Working Paper n° 8658, forthcoming *Review of Economics and Statistics*.
- [3] Blanchard, O. and J. Tirole (2006), “The Joint Design of Unemployment Insurance and Employment Protection. A First Pass, IDEI Working Paper, 258.

⁹At this point it would be useful to have a supplementary policy that returned the extra government revenue to the workers in this group (see the discussion of the outcomes we consider).

- [4] Browning, M., L.P. Hansen and J. Heckman (1999), "Micro Data and General Equilibrium Models", *Handbook of Macroeconomics*, Volume 1A, edited by J. Taylor and M. Woodford, North Holland, Amsterdam.
- [5] Chone, P. and G. Laroque (2005), "Optimal Incentives for Labor Force Participation", *Journal of Public Economics*, 89, 395-425.
- [6] Cullen, J.B. and J. Gruber (2000), "Does Unemployment Insurance Crowd Out Spousal Labor Supply?", *Journal of Labor Economics*, 18, 546-572.
- [7] Gouriéroux, C. and A. Monfort, (1996) *Simulation-Based Econometric Methods*, CORE Lectures, Oxford University Press.
- [8] Gouriéroux, C., Monfort, A. and E. Renault (1993), "Indirect Inference", *Journal of Applied Econometrics*, 8, 85-118.
- [9] Gruber, J. (1997), "The Consumption Smoothing Benefits of Unemployment Insurance", *American Economic Review*, 87, 192-205.
- [10] Krueger A., and B. Meyer (2002), "Labor Supply Effects of Social Insurance", *Handbook of Public Economics*, Vol. 4, North-Holland, Amsterdam.
- [11] Hall, G. and J. Rust (2003), "Econometric Methods for Endogenously Sampled Time Series: The Case of Commodity Price Speculation in the Steel Market". Unpublished Manuscript, University of Yale.
- [12] Hamermesh, D.S. (1980), "Unemployment Insurance and Labor Supply", *International Economic Review*, 21, 517-527.
- [13] Hansen, G. and A. Imrohorglu (1992), "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard", *Journal of Political Economy*, 100, 118-142.
- [14] Hopenhayn, H. and J.P. Nicolini (1997), "Optimal Unemployment Insurance", *Journal of Political Economy*, 105, 412-438.
- [15] McCall, B. (1996), "Unemployment Insurance and Unemployment Spells", *Econometrica*, 64, 647-682.
- [16] Pakes, A. and D. Pollard (1989), "Simulation and the asymptotics of optimization estimators", *Econometrica*, 57(5), 1027-1057.
- [17] Pedersen, P.J. and N. Smith (2002), "Unemployment Traps: Do Financial Disincentives Matter?", *European Sociological Review*, 18, 271-288.
- [18] Saez, E. (2002), "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses", *Quarterly Journal of Economics*, 117, 1039-73.
- [19] Wang, C. and S. Williamson (1996), "Unemployment Insurance with Moral Hazard in a Dynamic Economy", *Carnegie-Rochester Conference Series on Public Policy*, 1-41.