The Value of Joint Retirement of Married Couples

Thomas H. Jørgensen[†]

August 1, 2012

Abstract

I estimate the value of joint retirement of elderly Danish households using a collective structural life cycle model of consumption and retirement. The model encompass non-separability between consumption and leisure and income uncertainty, correlated across spouses. I find positive valuation of joint retirement of both males and females. Point estimates indicate that males tend to value joint retirement more than their female counterpart.

To illustrate the importance of the value of joint retirement, I compare policy responses from changes in financial incentives from the collective model with nested unitary models. Labor market responses predicted by the two groups of models diverge. The unitarian models seem to overestimate the policy responses.

Keywords: Joint Retirement, Household Consumption, Labor Supply, Collective Model, Dynamic Stochastic Programming, Structural Estimation.

JEL-codes: D13, D91, J13, J22, J26.

[†]Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen, and Centre for Applied Microeconometrics (CAM). Contact: thomas.h.jorgensen@econ.ku.dk.

I am deeply grateful to Bertel Schjerning, Mette Ejrnæs, John Rust, Søren Leth-Petersen, Martin Browning, and Christopher Carroll for valuable comments and discussions. Also, I thank the Danish Economic Counsil for providing data and insights regarding implementation of the Danish insitutional settings. The usual disclaimer applies.

1 Introduction

I estimate the value of joint retirement of elderly Danish households using a collective structural life cycle model of consumption and retirement. The model encompass non-separability between consumption and leisure and income uncertainty, correlated across spouses. I find positive valuation of joint retirement of both males and females while the results indicate that males tend to value joint retirement more than their female counterpart. I also provide evidence that unitarian models should be abandoned when evaluating policy proposals.

One regularity often found in the retirement pattern of elderly couples is the tendency to retire at the same time.[1] However, most structural models estimated in the retirement literature are based on single males not taking the joint decision of multi-agent households into account.[2] Leaving out the joint decision potentially lead to miss-specification and little out of sample relevance since the wast majority are married at the age of retirement. For example, evaluating the effect of increasing the age of eligibility for early retirement with, say, two years in an unitarian model will likely produce biased effects of such a policy change. This study present evidence that this in fact the case, pointing to the importance of couples' joint retirement behavior.

Some studies does allow agents to be married but do not directly model the spouse of an individual. See, e.g., Rust and Phelan (1997) and Iskhakov (2010) who include marital status, but does not model the couples' joint decision process. The focus of these papers are the effect of health insurance on retirement. This topic is highly relevant and has received, and continue to receive, attention in the retirement literature.[3] However, since the model in this paper is not aimed at describing the US population or the transition into disability pension effects from health-related issues are not included.

Another strand of literature focus exclusively on couples. See, e.g., Hurd (1990); Blau (1998, 2008); Gustman and Steinmeier (2000, 2004, 2005, 2009); and Blau and Gilleskie (2006, 2008). In these studies, important information from single's behavior is excluded. Neglecting the effects from singles is the opposite extreme and cannot be expected to produce trustworthy policy evaluations.

This study include both singles and married couples' consumption and retirement

[1] See, e.g., Hurd (1990); Blau (1998, 2008); An, Christensen and Gupta (1999); Gustman and Steinmeier (2000, 2004, 2005); Mastrogiacomo, Alessie and Lindeboom (2004); Blau and Gilleskie (2006); and van der Klaauw and Wolpin (2008).

[2] See, e.g., Gustman and Steinmeier (1986); Stock and Wise (1990); Berkovec and Stern (1991); Lumsdaine, Stock and Wise (1992, 1994); Blau and Gilleskie (2008); Belloni and Alessie (2010); Haan and Prowse (2010); and Bound, Stinebrickner and Waidmann (2010)

[3] Consult, e.g., Blau and Gilleskie (2006, 2008); van der Klaauw and Wolpin (2008); and Iskhakov (2010) as well as the recent working papers of Casanova (2010); Gallipoli and Turner (2011); and Ferreira and Santos (2012) for studies of the effect from health insurance on retirement in the US. Christensen and Kallestrup-Lamb (2012) find evidence that health does effect the early retirement in Denmark as well.

choices, as is also done in van der Klaauw and Wolpin (2008); Mastrogiacomo, Alessie and Lindeboom (2004); and Michaud and Vermeulen (2011). The latter two, however, does not incorporate the important dynamics of the household retirement choices. van der Klaauw and Wolpin (2008) restrict their analysis to only include low income households and exclude all who have ever had a defined contribution (DC) plan. I include a separate income state for each spouse, providing a much more comprehensive analysis of the retirement behavior across the income distribution. Further, I include high quality information on private pension wealth. These data are rarely available and to my best of knowledge no dynamic programming model of couples has included private pension wealth.[4]

Finally, this is the first dynamic programming model of couples estimated on high quality Danish register data on third party reported income and wealth information. Almost all studies on joint retirement are based on the Health and Retirement Study (HRS) and empirical evidence of the importance of joint retirement from other sources are therefore valuable.[5] Further, on a technical note, I do not have knowledge of any other studies estimating a model with both discrete and continuous choices solved with the EGM method of Carroll (2006).

The present study is also related to the recent working papers of Casanova (2010) and Gallipoli and Turner (2011). Casanova (2010) focus on health insurance effects on couples' retirement and consumption behavior. However, the behavior of singles and households with private pension wealth are excluded from her analysis. Gallipoli and Turner (2011) formulate three models; one for singles, one for couples with complimentarities in leisure, and one model where couples solve a non-cooperative game with respect to retirement. They find that the non-cooperative model fit female retirement behavior, while the model with complimentarities fit the male retirement behavior the best. They do, however, not estimate either of the models but calibrate parameters using the US Panel Study of Income Dynamics (PSID).

The paper proceeds as follows: Section 2 present the data used for estimation along with sample selection criteria and empirical "facts" motivating the model specification. Section 3 discuss the implemented institutional settings. In Section 4 the collective household model is presented and Section 5 discuss the endogenous grid (EGM) method applied to solve the model. In Section 6 the estimation strategy, results and model fit is discussed and Section 7 present policy experiments from the collective model and unitary versions. Finally, Section 8 concludes and suggests further research.

[4]Bound, Stinebrickner and Waidmann (2010) include private pension wealth in a model of single's choice of leaving the workforce and applying for Disability Pension.

[5]See as exceptions, the reduced form studies of An, Christensen and Gupta (1999); Jia (2005) and Mastrogiacomo, Alessie and Lindeboom (2004) using Danish, Norwegian and Dutch data, respectively.

2 Data and Empirical Regularities

The data used throughout is based on high quality Danish administrative register data on the total Danish population in the years 1996-2008. The pension wealth data (PERE) is based on information relating to a wealth test regarding early retirement (in Danish »Pensionsrettigheder«) collected for the Danish tax authority. Pension wealth information are collected for all individuals at the age of $59\frac{1}{2}$ independent of eligibility for early retirement. The pension wealth test on early retirement was introduced in 1999. Therefore, individuals aged 61 or above in 2000 are not included, leaving the oldest individuals in the data to be 68 years old.

The sample is further restricted to households with no cohabiting children where all members are wage workers at the first data entry and no younger than 57 years old (married females are allowed to be as young as 57-6=51 years old). No more than six years of age difference between spouses and households who are net-borrowers (excluding private pension wealth) one year are excluded. Further, if one member of a household is eligible for the Danish equivalence of a defined benefit plan (DB) in the US (in Danish »Tjenestemandspension«) or leaves the workforce through disability pension, the household is excluded from the analysis.

The above criteria yield a population consisting of $150,323$ households, summarized in Table A1 in Appendix A. Throughout the analysis, income and wealth are measured in 2008 prices.[6]

Clasification of retirement is based on labor market status the end of November a given year. The measure of the retirement age is accurate, since age is measured at the same time. However, potential timing problems regarding income can arise since an individual retiring, say, in the beginning of November has potentially earned nearly a full year of labor market income while observed as retired by this definition.

Eligibility for early retirement depend upon many years (10-30 years, depending on the cohort) of payments to the program. Hence, the actual eligibility is not observed in the data but is approximated by the last year of payment to the program. If an individual stops the payment to the program at, say, the age of 61, the age of eligibility is then assumed to be 61.

2.1 Empirical Regularities the Model Should Capture

Danish couples tend to retire jointly. Figure 1 display density plots for nine different spousal age differences (age of male - age of female) with retirement age difference on the horizontal axis. The mass under "Δ retirement $= 0$" (red bin) illuminate the joint

[6] Income and wealth is adjusted into 2008 prices using the change in old age basis pension (B_t). This measure of inflation is chosen in order to make the implemented retirement scheme for 2008 compatible with years before 2008. The change (ΔB_t) has roughly been 2-3 pct. each year in the years 1998-2008.

retirement while the mass under the green bin illustrates couples retiring with the same number of years difference as the difference in age.

For example, panel (a) plot the difference in retirement age for households in which the male is four years younger than the female spouse. Nearly 1 out of 3 of such households retire within the same year while in less than 1 out of 10 households the male retire four years later than the spouse. In panel (i), households in which the male is four years older than the female spouse is considered. The same pattern emerges albeit less profound with only 1 out of 4 couples retiring jointly.

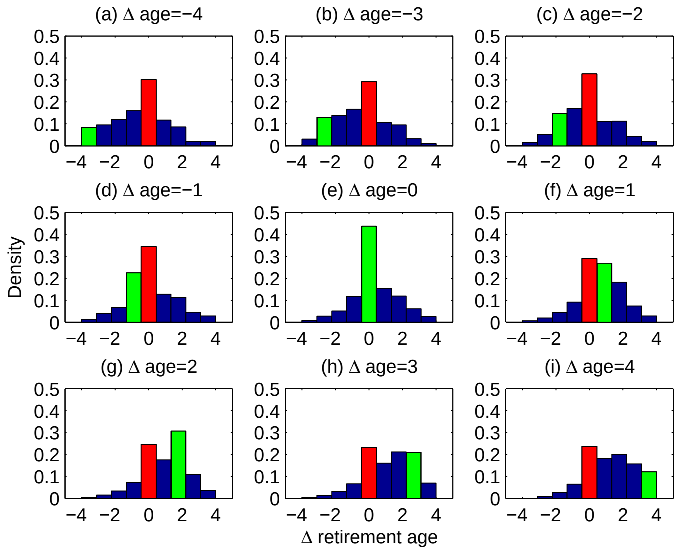

Figure 1 – Retirement Pattern of Danish Couples.

Figure 2 illustrate married male's and female's retirement age distribution for combinations of low/high income and pension wealth and eligibility for early retirement.[7] To illustrate how spousal characteristics affect the retirement age, the distributions conditional on the four combinations of income, wealth and eligibility in married households are presented. The first (blue) bin is both male and female low value, the second (red) bin is male low and female high, the third (green) bin is male high and female low, and the fourth (yellow) bin is both high. p-values from Pearson's χ^2 test of independence from spousal charactreristics are presented, all tests being significant.

Individuals with a relatively high level of income or private pension wealth prior to retirement tend to postpone retirement relative to low income/wealth individuals, c.f. Figure 2. Further, if an individual is not eligible for early retirement by the age of 60, retirement is postponed significatly with only a small fraction retiring at the age of 60. These behavioral features are identical across marital status and gender, but strongest for

[7]Low income is defined as annual pretax income of less than 250,000 DKK the year before retirement and low pension wealth is defined as less than 1,000,000 DKK the year before retirement.

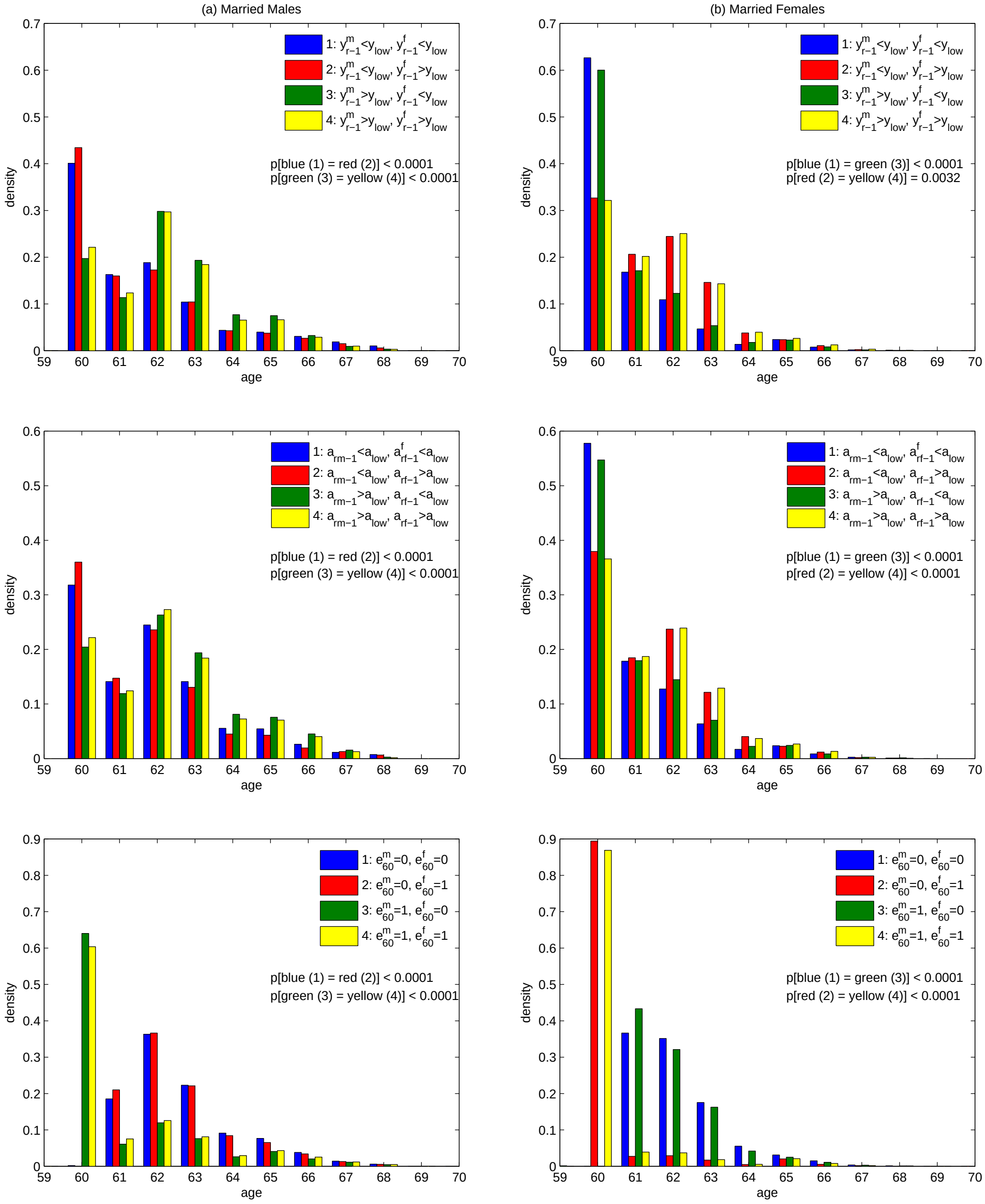

Figure 2 – Couple's Retirement Age Across Income, Wealth and Eligibility for Early Retirement.

females.[8]

The effect of spousal characteristics vary across gender. Married males tend to postpone retirement if the spouse has relatively low income or private pension wealth. Married females, on the other hand, tend to advance retirement if their male spouse have a relatively low level of income or pension wealth. If both are eligible to early retirement by the age of 60, both married males and females tend to postpone retirement to the age of 61-62.

The danish retirement scheme facilitates retirement at the age of 60 for individuals being eligible to early retirement. However, if an individual chose to work two years while eligible to early retirement, a higher level of transfers are received, as well as milder pension wealth tests. This could explain the spikes at the age of 62, since people often become eligible at the age of 60. Further, old age pension is available to everybody at the age of 65 (subject to income and wealth tests) resulting in a (small) spike at the age of 65.

3 Institutional Settings

Pension wealth affect the level of retirement benefits an individual can receive when retiring. Retirement savings can be administrated by the employer or privately by the employee and three main types of retirement saving opportunities are available:

1. *Lifelong Annuity (LA)*, in Danish »Livsvarige Pensionsordninger«, is an insurance guaranteeing a monthly payment when retired. The amount guaranteed (commitment value) is received until death and is therefore increased (decreased) if the owner postpone (advance) retirement.

2. *Annuitized Individual Retirement Arrangement (AIRA)*, in Danish »Ratepension«, is a pension deposit commited by the owner to be distributed through annuities of 10 through 25 years. If the owner initiate the distribution of the funds after the early retirement age, a 40 pct. tax payment of the withdrawn amount will be collected by the government. If the funds are withdrawn earlier than the early retirement age, a tax of 60% is collected. Hence, the distribution of funds does not necessary start at the age of retirement although this is most common practice. The annuitization must be initialized by the age of 77.

3. *Individual Retirement Arrangement with no restrictions (IRA)*, in Danish »Kapitalpension«, is a AIRA with no committed to annuitize the pension wealth. There is no upper age limit to when the owner must withdraw the funds.

[8]See Figure A1 in Appendix A for retirement pattern of singles.

In order to keep the model tractable while maintaining incentives in the early retirement scheme, I assume all pension wealth is held in IRAs. Hence, I do not need to worry about commitment values and annuities of pensions. Further, the early retirement scheme does not discriminate between privately and employer administrated IRAs (see Table 1). I will hereafter refer to the pension wealth deposit in IRA as *private pension wealth,* whether the pension wealth is privately or employer administrated.[9]

The implemented transfer function, $\mathbf{T}(\cdot)$, does not include unemployment benefits, since all individuals are assumed working or retired,

$$\mathbf{T}^j(\mathbf{z}_t;\tau_{\mathbf{T}}) = \begin{cases} \mathbf{ER}(\mathbf{z}_t;\tau_{\mathbf{T}}) & \text{if} \quad 60 \leq age_t^j < 65 \text{ and } d_t^j = 0, \\ \mathbf{OA}(\mathbf{z}_t;\tau_{\mathbf{T}}) & \text{if} \quad age_t^j \geq 65 \text{ and } d_t^j = 0, \end{cases}$$

where $\mathbf{ER}(\mathbf{z}_t;\tau_{\mathbf{T}})$ and $\mathbf{OA}(\mathbf{z}_t;\tau_{\mathbf{T}})$ summarize the early retirement pension and old age pension, respectively. All institutional settings are implemented for the year 2008 applying to the cohorts used for estimation (born 1940-1948). Since it would be far out of the realm of a stochastic dynamic programming model to incorporate all aspects determining the level of transfers, approximations are applied. See, e.g., Jørgensen (2009) and Forsikring & Pension (2008) for a description of the Danish pension system.

3.1 Early Retirement Pension

The ER is determined by eligibility and pension wealth at the *time of retirement*. Furthermore, the ER depends on the type of pension fund the saving is placed in, whether the pension fund has been managed by the employer, and wage rate and hours worked if working while retired. Once the ER has been calculated based on this information at the time of retirement, the ER received in years until old age pension age is fixed.

Here, the ER is recalculated each year using present information. Since the model is solved by backwards induction the retirement age and previous information on income and pension wealth is not known.

Table 1 illustrate the testing of private pension wealth deposits and payouts in the ER scheme for the three different types of pension wealth (LA, AIRA, IRA) across privately and employer administrated types. The fulfillment of the two-years rule is indicated by $e_t = 2$. As can be seen, assuming all pension wealth held in IRAs simplifies matters substatially. Combining the assumption that all pension wealth is held in IRAs with the assumption of zero hours worked when retired, the early retirement scheme implemented

[9]Private pension funds not based on deposits but rather on, e.g., predicted annuities from life expectancy are converted by The Danish Economic Counsil into deposits by discounting the annuities with a survival and inflation adjusted interest rate.

Table 1 – Early Retirement Wealth Test for Types of Pension Wealth, Retirement Age and Administrative Type.

		$60 \leq age_t < 62$		$age_t \geq 62$, and $e_t = 2$	
		Employer[‡]	Private	Employer	Private
LA[†]	payout	Tested	Not	Tested	Not
	deposit	*Tested*	*Tested*	*Not*	*Not*
IRA	payout	Not	Not	Not	Not
	deposit	*Tested*	*Tested*	*Not*	*Not*
AIRA	payout	Tested	Not	Tested	Not
	deposit	*Tested*	*Tested*	*Not*	*Not*

[†] "LA" refers to »Livrente« in Danish and the "deposit" is the commitment value of the LA, "IRA" (Individual Retirement Account) refers to »Kapital pension« in Danish, and "AIRA" (Annuitized Individual Retirement Account) refers to »Ratepension« in Danish.
[‡] "Employer" refers to employer administrated and "Private" refers to pension wealth administrated by the individual in an private retirement account.

here can be formulated as

$$\text{ER}_t = \begin{cases} 0 & \text{if } e_t = 0, \\ \overline{\text{ER}}_0 - .6 \cdot (.05 \cdot (\text{IRA deposit}_t) - \underline{\text{ER}}) & \text{if } e_t = 1 \text{ and } 60 \leq age_t < 62, \\ \overline{\text{ER}}_1 & \text{if } e_t = 2 \text{ and } age_t \geq 62, \end{cases}$$

where $\overline{\text{ER}}_0 = 166,400 \approx \$30,250$ is the maximum early retirement pension in 2008 if the two year rule is not fulfilled, $\overline{\text{ER}}_1 = 182,780 \approx \$33,250$ is the maximum early retirement pension if the two year rule is fulfilled, and $\underline{\text{ER}} = 12,600 \approx \$2,300$ is a deduction.

3.2 Old Age Pension

The most important factor determining the level of OA is the individual's annual income. Marital status, potential labor market status and income of the spouse also affect the level of pensions received. Further, the wealth (excluding private pension wealth, housing and debt) also affect whether households are eligible for supplementary transfers. These supplementary transfers are aimed at households with very low wealth, such that households with more than approximately $10,000 worth of wealth is not eligible for these supplementary benefits. Not only wealth but also information on square feet of residence, whether the residence is owned or rented, and the number of children residing are used to determine the actual level of supplementary transfers.

The implementation of old age pension benefits only include the two main parts of the old age pension scheme in Denmark, ignoring the supplementary transfers directed to very low wealth households. I will refer to these as the base (OA_B) and additional (OA_A) part, in Danish »Grundbeløbet« and »Pensionstillæget«.

Due to these simplifying assumptions, the OA only depend upon individual income, potential spousal income and whether the spouse is retired, $\mathbf{OA}(y^m, y^f, d)$. Interestingly, for some combinations of own and spousal income, the OA system facilitates joint retirement, while at other combinations punishes joint retirement. Appendix B contain the implemented old age penseion rules.

3.3 Tax System

The Danish tax system is a piece-wise linear tax schedule. In 2008 there where three main tax brackets: 1) lower tax bracket, 2) middle tax bracket, and 3) upper tax bracket. Throughout the analysis, income equal labor market income ruling out capital gains and loses.

If a spouse does not utilize the full deduction (41,000DKK≈7,500USD) the remainder is transferred to the spouse. This creates an incentives for married couples *not* to retire simultaneously. Appendix B contain the implemented tax rules.

4 A Collective Model of Consumption and Retirement

Here, I formulate a collective model of married couples and singles in order to capture the complex and simultaneous influences from own and spousal income, pension wealth, and eligiblity for early retirement on the decision to retire, presented in Section 2.1.

Households are maximizing expected lifetime utility:

$$\begin{aligned} \max_{\{c_t, l_t\}_1^T} \quad & \mathbb{E}_0 \left[\textstyle\sum_{\tau=0}^{T} \beta^\tau \mathbf{U}(c_\tau, l_\tau, \mathbf{z}_\tau) \right] \\ \text{s.t.} \quad & c_\tau + s_\tau = m_\tau(\mathbf{z}_\tau), \forall \tau \in \{1, \ldots, T\}, \\ & s_\tau \geq 0, \end{aligned}$$

where β is the between-period discount factor, consumption and leisure (c, l) are the choice variables, $\mathbf{z}$ contains the different state variables, and $m(\mathbf{z})$ is the available household cash-on-hand.

4.1 State Space and Choice Set

State variables are partitioned into observed, $\mathbf{z}_t$, and (to the researcher) unobserved, ε_t, state variables, following Rust (1987). The observed states at time t are given by

$$\mathbf{z}_t = (a_t, d_t^m, d_t^f, age_t^m, age_t^f, y_t^m, y_t^f, e_t^m, e_t^f),$$

where $a_t \in \mathcal{R}_+$ is the available (household) assets in the beginning of period t, $d_t^j \in \{0, 1\}$ is the labor market status of spouse j, $age_t^j \in [57, 100]$ is the age of spouse j, $y_t^j \in \mathcal{R}_+$ is the pretax income of spouse j, and $e_t^j \in \{0,1,2\}$ indicates whether spouse j is eligible for early retirement benefits ($e_t^j = 1$) and if the individual fulfills the two year rule ($e_t = 2$).

The choice of retirement can be represented as the binary choice

$$d_{t+1}^j = \begin{cases} 1 & \text{if person } j \text{work at time } t+1, \\ 0 & \text{if person } j \text{retire at time } t+1, \end{cases}$$

where $d_{t+1} = (d_{t+1}^m, d_{t+1}^f) \in \{0,1\} \times \{0,1\}$ is the vector of household labor market choice at time t. The timing of this model is different than the existing literature, since each spouse's labor market status the *following* period, $d_{t+1} = (d_{t+1}^m, d_{t+1}^f)$ are the choices this period. As elaborated further in Section 5, this is done for computational reasons only and should not affect the results.

Alternatively, as done in French and Jones (2011), hours worked could be a (continuous) choice variable. However, the available data on hours worked are clustered at 37 hours a week (the norm in Denmark) and zero hours (not working). French and Jones (2011) argue that introducing a fixed cost to work will help explain this type of behavior. It is, however, questionable how much information there is gained from using hours worked instead of the more easily handled binary choice. Therefore, the labor market decision is modeled as a discrete variable, albeit it's continuous features in reality.

Aggregate *household* consumption, c_t, is endogenous in the model since the labor market participation decision is interrelated with the consumption decision through retirement savings, possibly binding budget constraints, and uncertainty about the future (Deaton, 1991 and Cagetti, 2003).

The marriage decision is assumed exogenous. Single individuals remain single until they die and couples can only become single due the death of the spouse.

4.2 Preferences

The household choices are assumed to be the outcome of NASH-bargaining (Bourguignon and Chiappori, 1994),

$$\mathbf{U}(c_t, d_{t+1}, \mathbf{z}_t; \theta_\mathbf{U}) = \lambda \mathbf{U}^m(c_t, d_{t+1}, \mathbf{z}_t; \theta_\mathbf{U}) + (1-\lambda)\mathbf{U}^f(c_t, d_{t+1}, \mathbf{z}_t; \theta_\mathbf{U}) + \varepsilon_t(d_{t+1}), \quad (4.1)$$

where $\varepsilon_t(d_{t+1}) \sim GEV(1)$ summarize the household choice-specific unobserved states and $\lambda \in [0,1]$ represents the Pareto weight/household power by each spouse, as argued in Browning and Chiappori (1998).[10]

[10] This approach is widely used in the literature on joint retirement of couples. See, e.g., An, Christensen and Gupta (1999); Mastrogiacomo, Alessie and Lindeboom (2004); Jia (2004, 2005); van der Klaauw and Wolpin (2008) and Casanova (2010). As an alternative, one could estimate the model as a cooperative dynamic game, incorporating the intra-household bargaining directly, as done in Gallipoli and Turner (2011).

The Pareto weight in (4.1) is a function of characteristics,

$$\lambda(\mathbf{z}_t;\theta_\lambda)=\frac{\exp(\lambda_0+\lambda_1 y_t^m/(y_t^m+y_t^f)+\lambda_2(age_t^m-age_t^f)+\lambda_3 a_t)}{1+\exp(\lambda_0+\lambda_1 y_t^m/(y_t^m+y_t^f)+\lambda_2(age_t^m-age_t^f)+\lambda_3 a_t)},$$

where $\lambda=.5$ if $\lambda_0=\lambda_1=\lambda_2=\lambda_3=0$. If the household power is a function of outcome variables, e.g., the difference in income between spouses (affected by labor market status), the outcome is generally not efficient anymore. This inefficiency arises since a spouse could undertake more labor, than what would be efficient, in order to gain household power. Therefore, the model would not be a result of NASH-bargaining, and equation (4.1) would merely be a "household welfare-function" given as the weighted sum of individual utilities.

Individual preferences are of the CES type, allowing for non-separability between leisure and consumption,[11]

$$\mathbf{U}^j(c_t,d_{t+1},\mathbf{z}_t;\theta_\mathbf{U})=\frac{1}{1-\rho}\left(c_t(d_{t+1},\mathbf{z}_t)^\eta l^j(d_t)^{1-\eta}\right)^{1-\rho}e^{\alpha'\mathbf{x}_t^j}-\alpha_2\mathbf{1}(d_t^j=0,d_{t+1}^j=1),\quad(4.2)$$

where ρ is the relative risk aversion, η is the share of consumption to the utility and $\alpha'\mathbf{x}_t^j=\alpha_3\mathbf{1}(single_t^j=0)+\alpha_4 age_t^j$. $\mathbf{1}(\cdot)$ is the indicator function, equal to one when the statement in the parentheses is true and the parameter $\alpha_2>0$ thus measures the dis-utility of entering the labor market *next period* if retired *this period*. Here, α_2 is not estimated, but rather fixed at a value insuring no re-entry into the labor market. Hence, retirement is absorbing in the present model.

Leisure depend on own and potential spousal labor market status,

$$l^j(d_t)=\bar{l}(1+\alpha_1^j\mathbf{1}(d_t^j=0,d_t^k=0))-h\mathbf{1}(d_t^j=1),\ k\neq j,\qquad(4.3)$$

where $\bar{l}=17\cdot 7\cdot 52=6,188$ is the endowment of (awake) hours a year, $h=37\cdot(52-7)=1,665$ is the (assumed) hours worked a year when working, and α_1^j is the *value of joint retirement* measured in leisure units. If $\alpha_1^j>0$ spouse j tend to value time together with the spouse.

The household budget constraint takes the form:

$$\begin{aligned}c_t+s_t &= \underbrace{a_t+\mathbf{Y}(\mathbf{z}_t^m,y_t^f;\tau_\mathbf{Y})+\mathbf{Y}(\mathbf{z}_t^f,y_t^m;\tau_\mathbf{Y})+\mathbf{T}(\mathbf{z}_t;\tau_\mathbf{T})}_{m_t} &(4.4)\\ a_t &= (1+r)s_{t-1} &(4.5)\\ s_t &= m_t-c_t\geq 0 &(4.6)\end{aligned}$$

[11] Structural models of consumption *and* leisure estimated in the literature often assume an utility function with separability between consumption and leisure. See, e.g., Gustman and Steinmeier (2004, 2005, 2009); Blau and Gilleskie (2006); and Blau (2008). However, Browning and Meghir (1991) show evidence that (at least in the UK) separability in consumption and leisure is rejected.

where $s_t \geq 0$ is the savings constraint, m_t is the "cash-on-hand", r is the per period net interest rate fixed at $r = 0.03$, and $\mathbf{Y}(\cdot)$ and $\mathbf{T}(\cdot)$ are the tax-function and government transfers, respectively, where $(\tau_{\mathbf{Y}}, \tau_{\mathbf{T}})$ summarize the elements of the institutional settings presented in Section 3 and Appendix B.

4.3 Private Pension Wealth

In order not to include separate (continuous) state variables for each spouse's private pension wealth, the fraction of wealth held by each spouse in private pension funds, $\wp_t^j$, are estimated a function of the included state variables,

$$\wp_t^j = \wp(\mathbf{z}_t^j).$$

The estimated equations are presented in Table A6 in Appendix D. The fit of the model is reasonable, albeit a slight tendency to underestimating the private pension shares for singles.

4.4 Death and Bequests

The survival probability is assumed to depend only on age and sex,[12]

$$\pi_t^j \equiv \Pr(\text{survival}_t^j | age_t^j, j),\ j \in \{m, f\}. \tag{4.7}$$

If spouse j dies at time t the widowed spouse receives all the household assets and is assumed single untill death.[13] If both individuals die at time t, the bequest function for the household is assumed to be of a similar form as the utility function in (4.2):

$$\mathbf{B}(a_t) = \gamma \frac{1}{1-\rho} (a_t + \kappa)^{\eta(1-\rho)}, \tag{4.8}$$

where a_t is the household assets left at time t, γ measures the value of bequest, and κ is a parameter determining the curvature of the bequest function.

4.5 Beliefs

Since the present model incorporates uncertainty about the future, beliefs regarding future income and eligibility for early retirement have to be specified.[14]

[12]The estimated death probabilities are presented in Appendix C.
[13]This approach is similar to the one of van der Klaauw and Wolpin (2008) while in Blau and Gilleskie (2006) and Casanova (2010) the widowed spouse is not included in the model.
[14]Age evolve deterministicly (unfortunately in real life but practical here), and labor market status next period is a choice variable.

4.5.1 Income Process

The pretax income processes of husband and wife are potentially correlated. Hence I model a simultaneous system with both spouse's income and labor market choices:

$$\begin{aligned} \ln y_t^m &= \mathbf{x}_t^m \theta_y^m + \eta_t^{y_m}, \\ \ln y_t^f &= \mathbf{x}_t^f \theta_y^f + \eta_t^{y_f}, \\ d_t^m &= \mathbf{1}(\mathbf{z}_t^m \delta^m + \eta_t^{d_m} > 0), \\ d_t^f &= \mathbf{1}(\mathbf{z}_t^f \delta^f + \eta_t^{d_f} > 0), \end{aligned}$$

where $\mathbf{x}_t^j$ contains different state variables such as labor market status, age, eligibility for early retirement and lagged log income. $\mathbf{z}_t^j$ contains, in addition to the ones in $\mathbf{x}_t^j$, the household wealth as an identifying restriction. η is assumed independent across individuals and time and identical multivariate normal distributed, $\eta \sim \mathcal{N}_4(\mathbf{0}, \mathbf{\Omega})$, with covariance matrix given by

$$\mathbf{\Omega} = \begin{pmatrix} \Omega_y & \Omega_{dy} \\ \Omega_{yd} & \Omega_d \end{pmatrix} = \begin{pmatrix} \sigma_{y_m}^2 & & & \\ \sigma_{y_m y_f} & \sigma_{y_f}^2 & & \\ \sigma_{y_m d_m} & 0 & 1 & \\ 0 & \sigma_{y_f d_f} & \sigma_{d_m d_f} & 1 \end{pmatrix}.$$

The specification allows for correlation between spousal income and labor market choices. This is a crucial element of the model. If, say, the male experience a negative shock to his income the probability of the pension benefits being attractive enough to "force" him into retirement is larger. When the male is retired the value of retirement is greater for the female ($\alpha_1 > 0$) and since the correlation between spousal income is estimated positive the female's probability of retiring is also greater.[15] On the other hand, a negative shock to male income could potentially offset the female to work longer in order to sustain the preferred level of consumption. Hence, there are several elements of the model driving the retirement pattern in potential opposite directions.

4.5.2 Eligibility for Early Retirement

The domain of future eligibility status is restricted by the institutional settings to be given by

[15] Couples have a tendency to have the same level of education (Nielsen and Svarer, 2006). This could result in a positive correlation between spousal income.

$$
\begin{aligned}
e_{t+1} \in & \quad \{0,1\} \quad \text{if } e_t = 0 \text{ and } 60 \leq age_{t+1} < 65, \\
e_{t+1} = & \quad 1 \quad \text{if } e_t = 1 \text{ and } 60 \leq age_{t+1} < 62, \\
e_{t+1} \in & \quad \{1,2\} \quad \text{if } e_t = 1 \text{ and } 62 \leq age_{t+1} < 62, \\
e_{t+1} = & \quad 2 \quad \text{if } e_t = 2 \text{ and } 62 \leq age_{t+1} < 65.
\end{aligned} \tag{4.9}
$$

I assume that individuals are aware of the institutional settings, but do not fully keep track of their eligibility. Therefore, individuals form beliefs about future eligibility,

$$
\begin{aligned}
P^j_{e=1} &\equiv \Pr(e^j_{t+1} = 1 | e^j_t = 0, \mathbf{z}^j_t), \\
P^j_{e=2} &\equiv \Pr(e^j_{t+1} = 2 | e^j_t = 1, \mathbf{z}^j_t),
\end{aligned}
$$

for $j = m, f$.

5 Solving the Model

The model for a single individual (j) can be formulated as the solution to the Bellman equation:

$$
\begin{aligned}
\mathbf{V}^j_t(\mathbf{z}_t, \varepsilon_t) &= \max_{\substack{0 \leq c_t \leq m(\mathbf{z}_t) \\ d^j_{t+1} \in \{0,1\}}} \left\{ \mathbf{U}^j(c_t, d^j_{t+1}, \mathbf{z}^j_t) + \varepsilon(d^j_{t+1}) + \beta \mathbb{E}_t \left[\mathbf{V}^j_{t+1}(\mathbf{z}_{t+1}, \varepsilon_{t+1}) | \mathbf{z}_t, c_t, d_{t+1} \right] \right\} \\
&= \max_{\substack{0 \leq c_t \leq m(\mathbf{z}_t) \\ d^j_{t+1} \in \{0,1\}}} \left\{ \mathbf{v}^j_t(\mathbf{z}^j_t, d^j_{t+1}) + \varepsilon(d^j_{t+1}) \right\},
\end{aligned}
$$

where the assumption of Extreme Value Type I error terms yields (Rust, 1994)

$$
\begin{aligned}
\mathbf{v}^j_t(\mathbf{z}^j_t, d^j_{t+1}) \equiv \; & \mathbf{U}^j(c_t, d^j_{t+1}, \mathbf{z}^j_t) + \beta \Bigg[(1 - \pi^j_{t+1}) \mathbf{B}(a_{t+1}) \\
& + \pi^j_{t+1} \int \underbrace{\log \left(\sum_{d^j_{t+2} \in \mathcal{D}(\mathbf{z}_{t+1})} \exp \left(\mathbf{v}^j_{t+1}(\mathbf{z}^j_{t+1}, d^j_{t+2}) \right) \right)}_{\equiv EV^j_{t+1}(\mathbf{z}^j_{t+1})} F(\mathrm{d}\mathbf{z}^j_{t+1} | \mathbf{z}^j_t, c_t, d^j_{t+1}) \Bigg]
\end{aligned} \tag{5.1}
$$

For couples, the Bellman equation is:

$$
\mathbf{V}_t(\mathbf{z}_t, \varepsilon_t) = \max_{\substack{0 \leq c_t \leq m(\mathbf{z}_t) \\ d_{t+1} \in \{1,2,3,4\}}} \left\{ \mathbf{v}_t(\mathbf{z}_t, d_{t+1}) + \varepsilon(d_{t+1}) \right\}, \tag{5.2}
$$

where

$$
\begin{aligned}
\mathbf{v}_t(\mathbf{z}_t, d_{t+1}) &= \lambda \mathbf{U}^m(c_t, d_{t+1}, \mathbf{z}_t) + (1-\lambda)\mathbf{U}^f(c_t, d_{t+1}, \mathbf{z}_t) + \beta \Bigg[(1-\pi^f_{t+1})(1-\pi^m_{t+1})\mathbf{B}(a_{t+1}) \\
&+ \pi^m_{t+1}\pi^f_{t+1} \int EV_{t+1}(\mathbf{z}_{t+1}) F(\mathrm{d}\mathbf{z}_{t+1}|\mathbf{z}_t, c_t, d_{t+1}) \\
&+ \pi^m_{t+1}(1-\pi^f_{t+1}) \int EV^m_{t+1}(\mathbf{z}^m_{t+1}) F(\mathrm{d}\mathbf{z}^m_{t+1}|\mathbf{z}^m_t, c_t, d^m_{t+1}) \\
&+ \pi^f_{t+1}(1-\pi^m_{t+1}) \int EV^f_{t+1}(\mathbf{z}^f_{t+1}) F(\mathrm{d}\mathbf{z}^f_{t+1}|\mathbf{z}^f_t, c_t, d^f_{t+1}) \Bigg].
\end{aligned}
$$

Since death is treated as exogenous, the (expected) value functions etc. for singles, $\{EV^j_s, c^j_s, d^j_s \,\forall\, j \in \{m, f\}, 1 \le s \le T\}$, can be found by solving the model for singles in a first step.

The consumption- and labor supply functions are uncovered numerically using the Endogenous Grid Method (EGM) proposed by Carroll (2006). The method is a modification of time-iterations, where the grid is defined over savings this period instead of savings last period. This trick replaces, for each value of the state space, a root-finding operation of a non-linear system with interpolation, reducing the computation time dramatically.

Since the present model includes a discrete choice-variable, a combination of Euler equation and value-function evaluation is used. The approach here is different than the one proposed by Barillas and Fernández-Villaverde (2007) or by Fella (2011) since I include an unobserved choice-specific state, ε, smoothing out the kinks from the discrete choices. Here, the consumption problem is solved using EGM conditioning on the discrete labor market choice. This leads to four (for couples, two for singles) choice-specific consumption functions. These functions are interpolated on the same grid and inserted (via interpolation) into the value function from the next period in order to calculate the conditional probability of each labor market choice. The solution method applied here is based on code from Schjerning (2006) and formally proven to be applicable by Clausen and Strub (2012). Consult Appendix E for a detailed describtion of the solution method.

6 Estimation Results

The estimation procedure applied to uncover the parameters of the model is asymptotically equivalent to Full Information Maximum Likelihood (FIML). Since the number of parameters in the model is large, the two-step procedure proposed by Rust (1994) is applied.

The two-step approach splits up the parameters in two groups: *i)* Parameters regarding processes which do not require solving the DP problem, $\Theta_1 = (\theta_y, \theta_e)$, and *ii)* Parameters regarding processes which require numerical solutions to the DP problem,

$\Theta_2 = (\theta_{\mathbf{U}}, \theta_{\mathbf{B}}, \theta_\lambda)$. Hence, the parameters in the transition probabilities of the observed state variables summarized in $F_{\mathbf{z}}(\mathbf{z}_{it}|\mathbf{z}_{it-1};\Theta_1)$, are estimated using partial MLE in the first step. Secondly, the parameters in the transition probabilities of the choice variables summarized in $F_{c,d}(c_{it}, d_{it+1}|\mathbf{z}_{it};\Theta)$ are estimated, also using partial MLE.[16] For readability, I present the estimated preference parameters below and defer the results on beliefs to Section 6.2.

6.1 Preferences

The derivation of the likelihood function regarding the preference parameters, $\Theta_2 = (\theta_{\mathbf{U}}, \theta_{\mathbf{B}}, \theta_\lambda)$, are described in detail in Appendix F. In order to make the estimation of parameters feasible, I restrict the income process of retirees to zero labor market income. Another crucial assumption is the conditional independence (CI) assumption:

Assumption (CI)**.** *The transition density for the controlled Markov process* $\{c_t, \mathbf{z}_t, \varepsilon_t\}$ *factors as*

$$F_{c,\mathbf{z},\varepsilon}(c_{t+1}, \mathbf{z}_{t+1}, \varepsilon_{t+1}|\mathbf{z}_t, \varepsilon_t, c_t, d_t, d_{t+1}) = F_c(c_{t+1}|d_{t+1}, \mathbf{z}_t)F_\varepsilon(\varepsilon_{t+1}|\mathbf{z}_{t+1})F_{\mathbf{z}}(\mathbf{z}_{t+1}|\mathbf{z}_t). \quad (6.1)$$

The CI assumption restricts the processes in several severe ways. Most important is the assumption that the unobserved states, ε, does not affect any processes directly. This rules out auto correlation in ε and restricts the dynamics of the model to be captured solely by the observed state variables.

Since the additive unobserved states, $\varepsilon_{tj} \equiv \varepsilon(d_t)$, are assumed *iid* Extreme Value Type I, the probability of household i choosing labor status j at time $t+1$ is given by the Dynamic Multinomial Logit (MNL) formula,

$$F(d_{t+1} = j|\mathbf{z}_t;\Theta) = \frac{e^{\mathbf{v}_{tj}}}{\sum_{k\in\mathcal{D}(\mathbf{z}_t)} e^{\mathbf{v}_{tk}}}, \quad (6.2)$$

where $\mathbf{v}_{tj} \equiv \mathbf{v}_t(\mathbf{z}_t, d_{t+1} = j)$ from (5.2).

Assuming *independence across households*, the log likelihood function regarding the preference parameters can be written as

$$\mathfrak{L}(\Theta_2|\hat{\Theta}_1) = \sum_{i=1}^{N}\left[\sum_{t=1}^{T_i}\left[\sum_{j\in\mathcal{D}(\mathbf{z}_{it})} \mathbf{1}\left(d_{it+1} = j\right)\mathbf{v}_{itj} - \log\left(\sum_{k=1}^{K_{it+1}} e^{\mathbf{v}_{itk}}\right)\right]\right]. \quad (6.3)$$

Table 2 reports the ML-estimates of the preference parameters. Since estimation of all parameters within an acceptable time frame turned out to be impossible, most parameters

[16] Ideally, in order to correct the standard errors for the two-step approach, one iteration of the full information likelihood function should be performed.

are fixed and only the risk aversion, ρ, and the value of joint retirement, α_1^m, α_1^f, are estimated using the likelihood function in equation 6.3.[17]

The estimated value of joint retirement is about 23 pct. and 18 pct. "additional" leisure from joint retirement for males and females, respectively. The value of joint retirement in van der Klaauw and Wolpin (2008) is measured in utility units, not directly comparable to the leisure value of joint retirement, estimated here. They do, however, also find a positive significant value of joint leisure. The comparable analysis in Casanova (2010) yields significantly lower value of joint retirement of about 360 worth of "additional" leisure hours (8 pct.) if the spouse is retired. Although her model restricts married males and females to value joint leisure the same and does not include singles in the analysis, the difference is unexpectedly large.

The estimated risk aversion (based on singles only) of 2.3 seems reasonable albeit larger than 1.6 and 1.7 reported in van der Klaauw and Wolpin (2008) for males and females, respectively.

Table 2 – Estimated Preferences, Θ_1.

Parameter		Estimate	(SE)	t-value
Discount factor†	β	.975	–	–
		Utility function, $\theta_{\mathbf{U}}$		
Risk aversion‡	ρ	2.303	(.051)	45.039
Consumption share†	η	.330	–	–
Male value of joint retirement	α_1^m	.228	(.052)	4.419
Female value of joint retirement	α_1^f	.175	(.037)	4.687
Disutility of re-entry into labor market†	α_2	2.0E+3	–	–
Taste shifter: Married couples†	α_3	-.693	–	–
Taste shifter: Age†	α_4	.000	–	–
		Power function, θ_λ		
Constant†	λ_0	.000	–	–
Male income share in household†	λ_1	.900	–	–
Age difference	λ_2	-.032	(.713)	-.045
Household assets	λ_3	.022	(.027)	.815
		Bequest function, $\theta_{\mathbf{B}}$		
Value of bequest†	γ	1.0E-5	–	–
Curvature in bequest function†	κ	1.000	–	–
$\mathcal{L}(\Theta)$			51.301	
$\max_i\{\lvert\partial\mathcal{L}(\Theta)/\partial\Theta_i\rvert\}$			$1.4\text{E}-6$	
# Households			150,323	

† Parameter value fixed.
‡ Parameter value estimated based on singles only.
Standard errors based on the inverse of the Hessian.

[17] The relative risk aversion is estimated using singles only, since solving the model for singles is considerably faster than sovling the model for couples. The share of consumption in utility, η, as well as the distributional factor regarding male income share in the power function, λ_1, is calibrated by comparing actual and simulated retirement age distributions. The taste shifter regarding couples, α_3, is equal log(.5), such that utility is scaled by .5 for couples.

Identification

The value of joint retirement (α_1^m, α_1^f) is identified through *i)* variation in age-differences within households, *ii)* variation in eligibility for early retirement across households, and *iii)* couples where one individual dies. The retirement scheme has several "kinks" where retirement incentives change dramatically helping to identify the value of joint retirement. These kinks at the age of 60, 62 and 65 in 2008 also increase the need for age/eligibility variation, since the effect from changes in incentives cannot be disentangled from the value of joint retirement.

For example, say we observe a household retiring jointly when the male is 62 years old and the female is 60 years old. If both are eligible for early retirement at the age of 60 (such that the male is fulfilling the two years rule when retiring at age 62) we cannot say whether the choice to retire jointly is due to a high value of joint leisure or because the early retirement (ER) scheme facilitates their behavior. Imagine instead the female not being eligible for early retirement and still retires simultaneously with her husband. In such a case, her behavior could be driven by a positive valuation of joint retirement. Alternatively, imagine that the male is only one year older than the female and still retiring jointly at the age of 60 and 61 years old, respectively. Then, since the male did not chose to retire when eligible one year earlier but postponed retirement until the female spouse retired (because she became eligible for early retirement), the behavior can be attributed to males valuing joint retirement.

6.1.1 Model Fit

To asses the ability of the estimated model to predict actual outcomes, the number of single men in the data (25,984), single women (36,803) and couples (87,536) are simulated using the parameters in Table 2. The initial distribution of state variables are identical to the actual data.

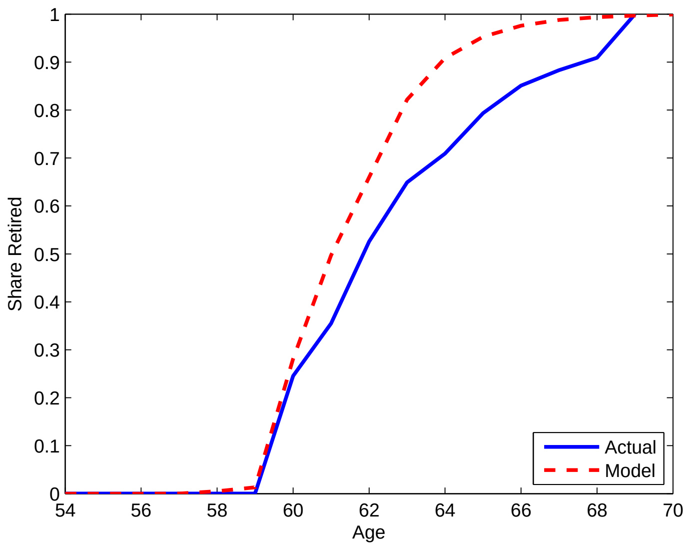

Figure 3 – Actual and Model Predicted Retirement.

Simulated and actual retirement age are illustrated in Figure 3. The model predictions are reasonable close to actual outcomes but seems to over predict retirement significantly at age 61. Table 3 investigates the fit for married and single males and females. The model over predict retirement of married males at age 61 and under predict at age 62. For married females and singles, significant under prediction at age 60 and over prediction at ages 63-64 are visible.

Table 3 – Actual and Predicted Retirement Age Distribution.

	Couples				Singles			
	Males		Females		Males		Females	
Age	Actual	Predicted	Actual	Predicted	Actual	Predicted	Actual	Predicted
57	.0	.0	.0	.0	.0	.0	.0	.0
58	.0	.0	.0	.0	.0	2.0	.0	1.6
59	.0	.0	.0	.0	.0	3.2	.0	3.0
60	33.3	34.8	52.8	27.7	33.3	9.2	52.8	11.5
61	14.0	24.8	18.6	21.9	14.0	13.9	18.6	18.5
62	25.0	12.9	15.6	15.3	25.0	23.2	15.6	22.6
63	14.3	12.9	7.8	17.0	14.3	21.4	7.8	19.6
64	5.1	7.3	2.1	9.0	5.1	12.1	2.1	1.7
65	4.7	3.6	2.1	4.5	4.7	6.8	2.1	5.8
66	2.3	1.8	.8	2.3	2.3	3.8	.8	3.1
67	.9	.9	.2	1.1	.9	2.3	.2	1.7
68	.4	.5	.1	.6	.4	1.2	.1	.9
69	.0	.2	.0	.3	.0	.5	.0	.4
70	.0	.1	.0	.2	.0	.5	.0	.4

notes: The numbers are fraction of retirees retiring at a given age.

The joint retirement pattern of couples is investigated in Table 4. The model is significantly overestimating the amount of joint retirement. This result hold true, even if the value of joint retirement is restricted to zero. Hence, I interpret this result as an artifact of the relatively poor predictive power of the model and not the large estimed value of joint retirement in Table 2.

Table 4 – Prediction Error of Joint Retirement, pct.

	ΔRetirement								
ΔAge	-4	-3	-2	-1	0	1	2	3	4
-4	-1.12	-3.23	-13.60	-13.48	43.66	-.87	-2.21	-1.62	-1.57
-3	-2.90	-2.73	-6.20	-2.12	31.97	-8.50	-2.31	-1.40	-1.13
-2	-1.64	-2.50	-.90	-9.74	43.71	-14.93	-.95	-4.68	-2.73
-1	-1.98	-3.48	-3.68	-7.85	4.64	-12.10	-1.48	-3.57	-2.11
0	-8.06	-13.67	-5.82	-2.58	66.44	-.23	-1.46	-11.57	-7.45
1	-9.32	-16.23	-9.32	-3.98	56.20	6.08	1.55	-6.66	-4.67
2	-9.09	-17.68	-11.06	-4.39	57.97	-4.09	6.93	-4.01	-3.22
3	-5.34	-9.60	-19.58	-12.11	62.67	-2.35	-3.79	-.42	-1.69
4	-3.03	-5.88	-11.69	-21.44	59.33	-3.77	-4.57	-3.64	-.58

notes: Percentage point deviation between actual and predicted fraction of a given retirement and age difference. ΔAge $\equiv age_t^m - age_t^f$. Similar definition for ΔRetirement.

Several reasons for the relatively poor fit of the model for the discussed groups are possible. First, the model could be a poor description of the actual decision process such that no parameter values can approximate the underlying data. The collective model presented here include several complex elements of the institutional settings as well as intra-household bargaining, suggesting that the overall model setup should be rich enough to describe the data, compared to the unitarian models in the existing literature. Secondly, I find evidence that couples and singles have very different preferences, indicating that parameters should be allowed to vary across marital status and possibly also gender. This would, however, more than double the number of preference parameters complicating the estimation further. Finally, the approximation of labor market income into ten discrete values might be to coarse. During calibration, I did find the solution to be sensitive to the number of discrete points used to approximate both income and wealth. Unfortunately, the complexity of the model does not permit increasing the number of points when estimating the preference parameters. Despite the issues, I consider the model well specified in what follows.

6.2 Beliefs

Here, the estimated beliefs, $\Theta_1 = (\theta_y, \theta_e)$, are presented. The main objective when estimating the beliefs is the ability to predict actual in sample outcomes. Hence, the performance of the estimated relations are evaluated on this margin.

6.2.1 Income Process

The parameters of the system is estimated by Maximum Likelihood, generalizing the approach in Heckman (1978) to be a four (two continuous, two binary) dimensional system:[18]

$$\begin{aligned}\mathcal{L}(\theta_y,\Omega) &= \frac{1}{\sum_i^N T_i}\sum_{i=1}^{N}\sum_{t=1}^{T_i}\log\Big[\phi_2(v_{it}^m,v_{it}^f,\Omega_y)\\ &\quad\times\Phi_2(r_{it}^m,r_{it}^f,\Omega_{d|y})^{\mathbf{1}(d_{it}^m=1,d_{it}^f=1)}\Phi_2(r_{it}^m,-r_{it}^f,\dot{\Omega}_{d|y})^{\mathbf{1}(d_{it}^m=1,d_{it}^f=0)}\\ &\quad\times\Phi_2(-r_{it}^m,r_{it}^f,\dot{\Omega}_{d|y})^{\mathbf{1}(d_{it}^m=0,d_{it}^f=1)}\Phi_2(-r_{it}^m,-r_{it}^f,\Omega_{d|y})^{\mathbf{1}(d_{it}^m=0,d_{it}^f=0)}\Big], \quad (6.4)\end{aligned}$$

where $\phi_2(x_1,x_2,\Omega_x)$ and $\Phi_2(x_1,x_2,\Omega_x)$ are the bivariate normal pdf and cdf, respectively, with zero mean and covariance Ω_x evaluated at (x_1,x_2), and

$$\begin{aligned}v_{it} &\equiv (\ v_{it}^m,\ \ v_{it}^f\ ) = (\ \ln y_{it}^m-\mathbf{x}_{it}^m\Theta_{\mathbf{I}}^m,\ \ \ln y_{it}^f-\mathbf{x}_{it}^f\Theta_{\mathbf{I}}^f\ ),\\ r_{it} &\equiv (\ r_{it}^m,\ \ r_{it}^f\ ) = (\ \mathbf{z}_t^m\delta^m,\ \ \mathbf{z}_t^f\delta^f\ )+v_{it}\Omega_{yd}\Omega_y^{-1},\\ \Omega_{d|y} &= \Omega_d-\Omega_{dy}\Omega_y^{-1}\Omega_{yd},\\ \dot{\Omega}_{d|y} &= \Omega_{d|y}\bigodot\begin{pmatrix}1 & -1\\ -1 & 1\end{pmatrix},\end{aligned}$$

where $\bigodot$ denotes element-wise multiplication. The estimated parameters are reported in table 5, using the δ-method to calculate the standard errors of the covariance parameters. I do not report the partial effects, since that is not of particular interest here. For singles, two-equation systems are estimated separately. The distribution of estimation errors, $\hat{\eta}^j=\log y^j-\log\hat{y}^j$, are plotted in Figure 4. The errors are roughly centered around zero, but there is substantial mass under the tails.

[18] Since $\mathbf{x}_{it}$ contains the lagged dependent variable, the likelihood function is conditional on initial values of the income processes. The estimation is based on people aged 57 or more and the conditional likelihood is, therefore, expected to be fairly similar to the unconditional.

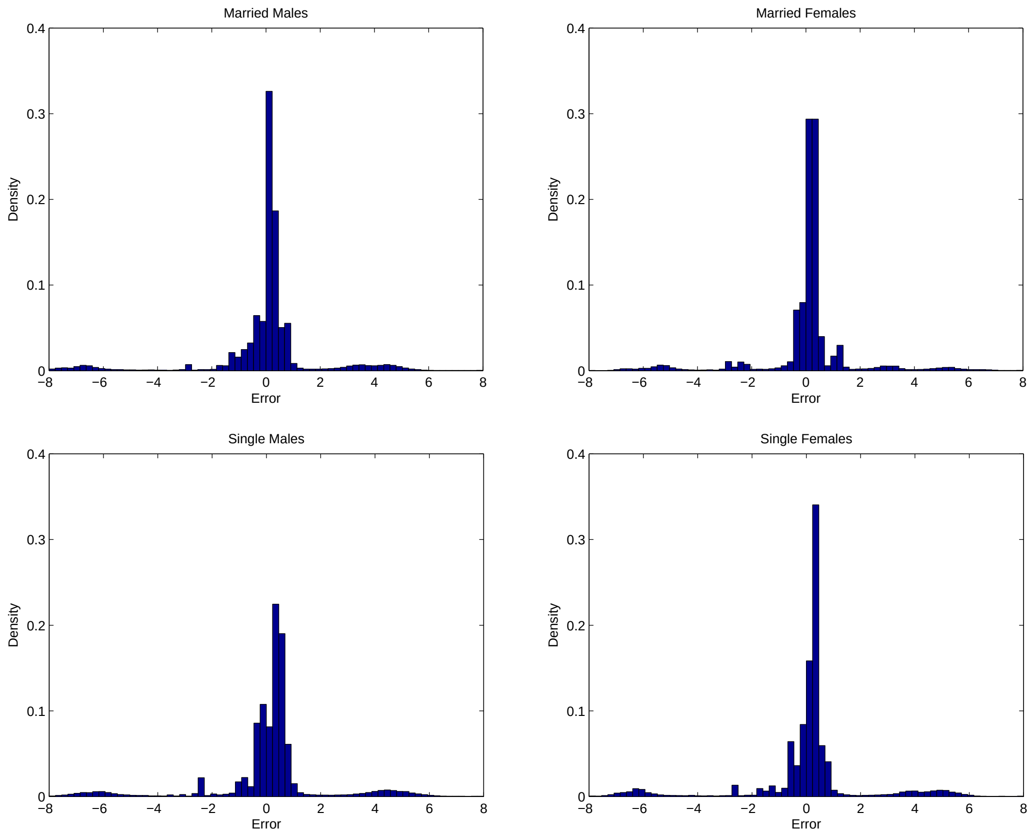

Figure 4 – Distribution of Prediction Error From Income Equations.

The estimated correlation between spousal labor market income, $\hat{\sigma}_{y_m y_f}$, is significant positive. This result underlines the importance of including the labor market income of each spouse and allowing for interdependence between the processes. Married female's labor market income tend to correlate with the age of the male spouse. This result is not true for males, indicating that females is more influenced by their male spouse than vice versa. Wealth is singificant in the selection equation, indicating that the instrument is valid and the pseudo R^2 of about 30 pct. is acceptable.

The income process estimated here is continuous in nature. When solving the model I descretize income. Therefore, I follow the approach of Rust (1990) and construct an income transition matrix using the estimated (continuous) income process. Say income is discretized in N_{inc} points $\vec{y} = (y_1, \ldots, y_{N_{inc}})$, where $y_1 < y_2, \ldots, y_{N_{inc}-1} < y_{N_{inc}}$. The probability of a single individual's income to fall in the interval $[y_{k-1}; y_k]$ is found by

$$P_k \equiv \begin{cases} \Phi(\hat{\eta}_t^j \leq y_1 | \mathbf{z}_t^j, d_{t+1}^j) & \text{if } k = 1, \\ \Phi(\hat{\eta}_t^j \leq y_k | \mathbf{z}_t^j, d_{t+1}^j) - \Phi(\hat{\eta}_t^j \leq y_{k-1} | \mathbf{z}_t^j, d_{t+1}^j) & \text{if } 1 < k < N_{inc}, \\ 1 - \Phi(\hat{\eta}_t^j \leq y_{N_{inc}-1} | \mathbf{z}_t^j, d_{t+1}^j) & \text{if } k = N_{inc}. \end{cases}$$

The probability associated with each of the $N_{inc} \times N_{inc}$ possible income states for couples at time $t+1$ are calculated by similar two-dimensional rules.

Table 5 – System Estimates of the Income and Labor Supply Processes, θ_y.

	Couples		Singles	
	Males	Females	Males	Females
Dep.: $\ln y_t^j$	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
$constant$	.280 (.020)***	-.127 (.023)***	-.308 (.052)***	.005 (.035)
$\ln y_{t-1}^j$	.571 (.001)***	.531 (.001)***	.550 (.002)***	.540 (.002)***
$d_t^j = 1$	2.597 (.032)***	2.818 (.029)***	2.669 (.062)***	2.541 (.050)***
$\ln y_{t-1}^j, d_t^j = 1$	.189 (.002)***	.232 (.002)***	.228 (.004)***	.234 (.003)***
$e_t^j = 1$	.100 (.142)	-.061 (.166)	.967 (.209)***	.565 (.166)**
$e_t^j = 2$	.802 (.141)***	.341 (.166)*	.985 (.205)***	.994 (.163)***
$age_t^j = 60$	.138 (.012)***	.391 (.014)***	.187 (.031)***	.149 (.021)***
$age_t^j = 61$	-.888 (.065)***	-2.465 (.080)***	-.539 (.121)***	-.663 (.093)***
$age_t^j = 62$	-.606 (.108)***	-1.815 (.129)***	-.196 (.176)	-.256 (.141)
$age_t^j = 63$	-.867 (.141)***	-.353 (.165)*	-.867 (.204)***	-.986 (.163)***
$age_t^j = 64$	-.567 (.141)***	-.089 (.166)	-.648 (.205)*	-.791 (.163)***
$age_t^j = 65$	.351 (.021)***	.437 (.028)***	.717 (.052)***	.477 (.036)***
$age_t^j > 65$	.117 (.020)***	.328 (.029)***	.460 (.052)***	.050 (.036)
$age_t^j = 60, e^j > 0$	1.228 (.142)***	2.139 (.167)***	.122 (.210)	.920 (.167)***
$age_t^j = 61, e^j > 0$	.253 (.156)	1.527 (.184)***	-.774 (.240)*	-.278 (.189)
$age_t^j = 62, e^j > 0$	1.010 (.178)***	2.308 (.208)***	-.067 (.271)	.284 (.216)
$age_t^m > age_t^f$	-.015 (.013)	.034 (.012)*		
$age_t^m < age_t^f$	-.010 (.009)	.054 (.009)***		
Labor Supply Parameters				
$constant$	5.651 (.078)***	4.631 (.044)***	3.479 (.093)***	3.761 (.122)***
$wealth_t$	.734 (.008)***	.494 (.010)***	.929 (.019)***	.781 (.016)***
$e_t^j = 1$	-2.685 (.212)***	-1.787 (.162)***	-3.720 (.235)***	-4.342 (.342)***
$e_t^j = 2$	-.904 (.212)***	.155 (.161)	-1.792 (.235)***	-2.288 (.342)***
$age_t^j = 60$	-4.129 (.146)***	-3.459 (.097)***	-.328 (.185)	-.277 (.218)
$age_t^j = 61$	-3.803 (.182)***	-3.002 (.130)***	-1.046 (.264)***	-1.173 (.260)***
$age_t^j = 62$	-3.722 (.198)***	-2.836 (.146)***	-1.327 (.281)***	-1.464 (.276)***
$age_t^j = 63$	-4.720 (.225)***	-4.729 (.167)***	-1.492 (.252)***	-1.192 (.363)*
$age_t^j = 64$	-5.066 (.225)***	-5.042 (.167)***	-1.745 (.252)***	-1.485 (.363)***
$age_t^j = 65$	-4.190 (.078)***	-4.287 (.045)***	-4.570 (.095)***	-4.811 (.122)***
$age_t^j > 65$	-4.528 (.078)***	-4.593 (.046)***	-4.903 (.095)***	-5.148 (.122)***
$age_t^j = 60, e^j > 0$	.532 (.245)*	-.377 (.184)*	.243 (.285)	.381 (.387)
$age_t^j = 61, e^j > 0$	1.022 (.273)**	-.071 (.203)	1.306 (.342)**	1.937 (.412)***
$age_t^j = 62, e^j > 0$	.329 (.278)	-.735 (.215)**	.870 (.355)*	1.597 (.423)**
$age_t^m > age_t^f$	-.036 (.011)*	.156 (.010)***		
$age_t^m < age_t^f$	.089 (.008)***	-.141 (.009)***		
Covariance Parameters				
σ_{y_j}	2.364 (.002)***	2.305 (.002)***	2.746 (.005)***	2.549 (.004)***
σ_{y_j,d_j}	-.167 (.016)***	-.313 (.018)***	-.204 (.026)***	-.109 (.016)***
σ_{y_m,y_f}	.386 (.007)***	.386 (.007)***		
σ_{d_m,d_f}	.373 (.004)***	.373 (.004)***		
$1 - \mathcal{L}(\Theta)/\mathcal{L}(0)$	.304		.279	.296
$\max_i\{\lvert\partial\mathcal{L}(\Theta)/\partial\Theta_i\rvert\}$	$1.2e-7$		$1.2e-7$	$4.7e-8$
# Obs	579,501		145,079	223,641
# Households	87,760		24,773	35,901

Notes: Since lagged variables are included, the number of observations used here is less than reported in Table A1 on page 32. Wealth is measured in 10,000,000 DKK. Standard errors based on the inverse of the hessian. The δ-method is used to calculate the standard errors of the covariance parameters. *: $p < .05$, **: $p < .001$, ***: $p < .0001$.

6.2.2 Eligibility for Early Retirement

The estimated parameters of the two individual logit equations $P^j_{e=1} \equiv \Pr(e^j_{t+1} = 1 | e^j_t = 0, \mathbf{z}^j_t)$ and $P^j_{e=2} \equiv \Pr(e^j_{t+1} = 2 | e^j_t = 1, \mathbf{z}^j_t)$ are presented in table 7. An alternative probit specification was estimated yielding similar results with a slight decrease in performance.

The model is capable of predicting the correct eligibility status of more than 80 pct. of the relevant sample, c.f. Table 6.

Table 6 – Predicted Eligibility.

		$\hat{e}^m_t$			$\hat{e}^f_t$		
		0	1	2	0	1	2
e^j_t	0	87.2	12.8	.0	82.0	18.0	.0
	1	12.9	86.3	0.7	10.8	88.5	0.7
	2	.0	20.4	79.6	.0	16.4	83.6

Notes: Row percentages. Estimated eligibility status clasification is based on $\hat{e}^j_t = k$ if $P^j_{e=k} > .5$ combined with the information in (4.9).

The estimated parameters indicate that couples have a higher probability of being eligible for early retirement (and fulfilling the two years rule). Wealth has a negative and diminishing effect on the probability of being eligible at age 60 as well as fulfilling the two years rule at age 62. The reason for this is most likely, that people who know they are not going to be eligible for early retirement save more in order to be able to retire roughly at the same time as if they where eligible.

Table 7 – Logit Estimates of Beliefs Regarding Eligibility for Early Retirement, θ_e.

	$\Pr(e_t^j = 1 \mid e_{t-1}^j = 0, \mathbf{z}_t^j)$		$\Pr(e_t^j = 2 \mid e_{t-1}^j = 1, \mathbf{z}_t^j)$	
	Males, $P_{e=1}^m$	Females, $P_{e=1}^f$	Males, $P_{e=2}^m$	Females, $P_{e=2}^f$
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
$age_t = 60$				
$constant$	3.666 (.190)***	3.129 (.199)***		
$single_t = 0$	-.256 (.030)***	.399 (.030)***		
$wealth_t$	-1.545 (.079)***	-.362 (.073)***		
$wealth_t^2$	.911 (.046)***	.431 (.050)***		
$wealth_t, single_t = 0$	.284 (.065)***	-.174 (.063)*		
y_{t-1}	-.871 (.013)***	-1.205 (.017)***		
y_{t-1}^2	.076 (.002)***	.125 (.003)***		
$age_t = 61$				
$constant$	3.389 (.210)***	3.454 (.217)***		
$single_t = 0$	.455 (.094)***	.243 (.106)*		
$wealth_t$	.933 (.243)**	-.299 (.256)		
$wealth_t^2$	-.580 (.153)**	.301 (.208)		
$wealth_t, single_t = 0$	.233 (.185)	.299 (.209)		
y_{t-1}	.384 (.037)***	.554 (.047)***		
y_{t-1}^2	-.040 (.005)***	-.056 (.008)***		
$age_t = 62$				
$constant$	.224 (.290)	-.112 (.295)	.113 (.095)	.005 (.100)
$single_t = 0$	-.010 (.217)	-.242 (.286)	-.204 (.045)***	.035 (.056)
$wealth_t$	1.759 (.520)**	.924 (.619)	-2.092 (.108)***	-.366 (.116)*
$wealth_t^2$	-1.188 (.347)**	-.836 (.548)	1.228 (.060)***	.511 (.077)***
$wealth_t, single_t = 0$	.420 (.407)	.414 (.529)	.213 (.087)*	-.042 (.095)
y_{t-1}	.385 (.083)***	.590 (.103)***	-.405 (.019)***	-.984 (.029)***
y_{t-1}^2	-.030 (.011)*	-.042 (.016)*	.040 (.002)***	.125 (.005)***
$age_t = 63$				
$constant$	-.450 (.472)	-.105 (.398)	1.449 (.189)***	1.724 (.191)***
$single_t = 0$	-.375 (.377)	-.671 (.484)	1.378 (.159)***	1.168 (.230)***
$wealth_t$	-.388 (.995)	-.023 (.989)	-.434 (.373)	-.198 (.404)
$wealth_t^2$	-.744 (.666)	-.492 (.845)	.253 (.244)	.202 (.322)
$wealth_t, single_t = 0$	1.574 (.744)	.742 (.854)	.130 (.249)	.104 (.357)
y_{t-1}	.654 (.179)**	.665 (.179)**	.259 (.063)***	.442 (.084)***
y_{t-1}^2	-.022 (.023)*	-.062 (.029)*	-.042 (.007)***	-.054 (.011)***
$constant$	-1.672 (.187)***	-1.469 (.196)***	.582 (.080)***	.302 (.084)**
$1 - \mathcal{L}(\Theta)/\mathcal{L}(0)$	.303	.269	.310	.404
# Obs	143,027	124,137	63,871	45,554

Notes: Estimates in column one and two are based on individuals aged over 59 and under 65 with $e_{t-1}^j = 0$ and $d_{t-1}^j = 1$. Estimates in column three and four are based on individuals aged over 61 and under 65 with $e_{t-1}^j = 1$ and $d_{t-1}^j = 1$. Household wealth is measured in 10,000,000 DKK and income in 100,000 DKK. *: $p < .05$, **: $p < .001$, ***: $p < .0001$.

7 Policy Evaluation Comparison

To illustrate the importance of joint retirement of couples when performing policy evaluations, I compare policy simulations from the collective model, described throughout the paper, with three nested unitarian models. The first unitarian model (UNI1) is simply using the model for single males only and the second unitarian model (UNI2) include also single women. The third unitarian model (UNI3) include also couples but is based on a restricted version of the collective model with:

$$\lambda_0 = \lambda_1 = \lambda_2 = \lambda_3 = \alpha_1^m = \alpha_1^f = 0.$$

UNI3 is unitarian in the terminology of Browning, Chiappori and Lechene (2006), although spousal characteristics can influence the retirement decision of individuals through the household budget constraint.[19]

I compare policy simulations from a reduction in the early retirement benefit by 25 pct. while increasing the benefit received if the two years rule is fulfilled by 25 pct.. This policy is one possible way to foster postponement of retirement through financial incentives. When simulating data, actual initial values are used, as done when evaluating the model in Section 6.

Figure 5 plot the predicted retirement responses from the collective model, COL, and the three unitarian models, UNI1, UNI2 and UNI3 from reducing the financial incentive to retire at ages 60 and 61. The estimated responses based on UNI2 and UNI3 are in the opposite direction from what would be expected, suggesting an increase in retirement at age 60 of 0.1 percentage point. The model for single males is by far the most common model used in the existing literature, suggesting a decrease in retirement at age 60 of roughly 0.15 percentage points and at ages 62-63 of about 0.1 percentage point in total. This decrease in retirement at ages before 65 offset increased retirement at age 65 by more than 0.2 percentage points.

[19] Ideally, the three unitarian models should have been estimated independently. In stead, I use the same (relevant) parameters in all four models.

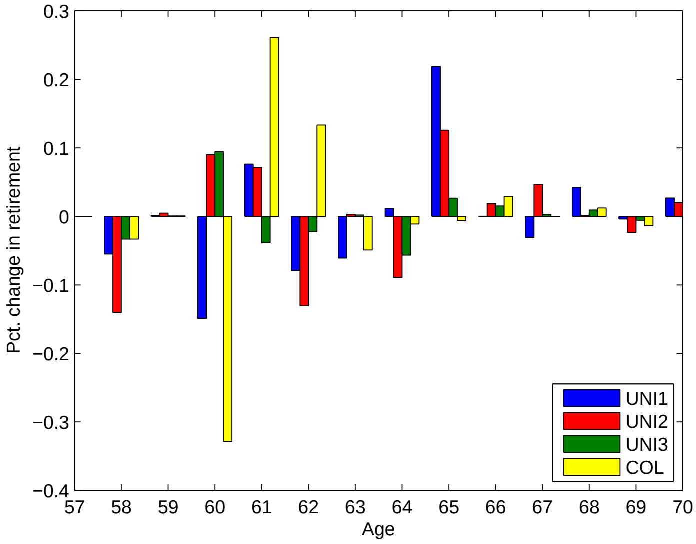

Figure 5 – Predicted Policy Responses.

The collective model, COL, predict by far the largest behavioral effect at age 60 by a decrease of more than 0.3 percentage points. The decrease at age 60 is fully absorbed in the next two years, such that there is no effect on the fraction retiring at age 63 and later.

The unitarian model based on single males only, UNI1, roughly predicts that 0.3 percentage points of people who would otherwise have retired at ages below 65 will postpone retirement until age 65 as a result of the policy change. Although the collective model predict a 0.3 percentage points drop at the age of 60, the overall effect of the policy is smaller, since almost all of the drop is postponed only one year. Hence, the *unitarian model over predict* the policy response. This result is in accordance with the finding in van der Klaauw and Wolpin (2008), where singles tend to have the larger response to policy changes.

8 Conclusion and Further Research

A thorough analysis of couple's joint retirement and saving choices have been conducted. Throughout the analysis great care has been taken to formulate a structural model capturing the incentives of elderly Danish households. The resulting estimated preferences fit the high quality register data reasonably well.

The estimation of the preferences from complex models as the one presented in this paper requires a lot of computation time. The results in this paper does, however, suggest that the computational burden is worth the effort.

The estimated value of joint retirement strongly suggest that non-monetary elements do in fact play an important role when households chose whether to retire or work. The results suggest that using unitarian models to perform policy predictions are most likely flawed and will result in overestimation of such policy effects.

The difficulties encountered during estimation of preferences illustrate significant differences between married couples and singles. Further research in estimating couples' preferences for joint retirement is, therefore, necessary.

References

AN, M., B. CHRISTENSEN AND N. GUPTA (1999): "A Bivariate Duration Model of the Joint Retirement Decisions of Married Couples," CLS Working Papers 99–10, Centre for Labour Market and Social Research, Denmark.

BARILLAS, F. AND J. FERNÁNDEZ-VILLAVERDE (2007): "A generalization of the endogenous grid method," *Journal of Economic Dynamics and Control*, 31(8), 2698–2712.

BELLONI, M. AND R. ALESSIE (2010): "Retirement choices in Italy: what an option value model tells us," Netspar Discussion Paper 10/2010-058, Network for Studies on Pensions, Aging and Retirement.

BERKOVEC, J. AND S. STERN (1991): "Job Exit Behavior of Older Men," *Econometrica*, 59(1), 189–210.

BLAU, D. M. (1998): "Labor Force Dynamics of Older Married Couples," *Journal of Labor Economics*, 16(3), 595–629.

——— (2008): "Retirement and Consumption in a Life Cycle Model," *Journal of Labor Economics*, 26(1), 35–71.

BLAU, D. M. AND D. B. GILLESKIE (2006): "Health Insurance and Retirement of Married Couples," *Journal of Applied Econometrics*, 21(7), 935–953.

——— (2008): "The Role Of Retiree Health Insurance In The Employment Behavior Of Older Men," *International Economic Review*, 49(2), 475–514.

BOUND, J., T. STINEBRICKNER AND T. WAIDMANN (2010): "Health, economic resources and the work decisions of older men," *Journal of Econometrics*, 156(1), 106–129.

BOURGUIGNON, F. AND P.-A. CHIAPPORI (1994): "The collective approach to household behaviour," in *The Measurement of Household Welfare*, ed. by R. Blundell, I. Preston and I. Walker, chap. 3, pp. 70–84. Cambridge University Press.

BROWNING, M. AND P.-A. CHIAPPORI (1998): "Efficient Intra-Household Allocations: A General Characterization and Empirical Tests," *Econometrica*, 66(6), 1241–1278.

BROWNING, M., P.-A. CHIAPPORI AND V. LECHENE (2006): "Collective and Unitary Models: A Clarification," *Review of Economics of the Household*, 4, 5–14.

BROWNING, M. AND C. MEGHIR (1991): "The Effects of Male and Female Labor Supply on Commodity Demands," *Econometrica*, 59(4), 925–951.

CAGETTI, M. (2003): "Wealth Accumulation Over the Life Cycle and Precautionary Savings," *Journal of Business and Economic Statistics*, 21(3), 339–353.

CARROLL, C. D. (2006): "The Method of Endogenous Gridpoints for Solving Dynamic Stochastic Optimization Problems," *Economics Letters*, 91(3), 312–320.

——— (2011): "Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems," Lecture notes, Department of Economics, Johns Hopkins University, Baltimore.

CARROLL, C. D. AND M. S. KIMBALL (1996): "On the Concavity of the Consumption Function," *Econometrica*, 64(4), 981–992.

Casanova, M. (2010): "Happy Together: A Structural Model of Couples' Joint Retirement Choices," Working paper, Department of Economics, UCLA.

Christensen, B. J. and M. Kallestrup-Lamb (2012): "The Impact of Health Changes on Labor Supply: Evidence from Merged Data on Individual Objective Medical Diagnosis Codes and Early Retirement Behavior," *Health Economics*.

Clausen, A. and C. Strub (2012): "Envelope Theorems for Non-Smoth and Non-Convace Optimization," Working Paper 62, University of Zurich.

Deaton, A. (1991): "Saving and Liquidity Constraints," *Econometrica*, 59(5), 1221–1248.

Fella, G. (2011): "A Generalized Endogenous Grid Method for Non-concave Problems," Working Papers 677, Queen Mary, University of London, School of Economics and Finance.

Ferreira, P. C. and M. R. d. Santos (2012): "The effect of social security, health, demography and technology on retirement," mimeo, Graduate School of Economics, Rio de Janeiro, Brazil.

Forsikring & Pension (2008): *Sociale Ydelser 2008: Hvem, hvad & hvornår*. Forlaget Forsikring, (In Danish).

French, E. and J. B. Jones (2011): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, 79(3), 69–732.

Gallipoli, G. and L. Turner (2011): "Social Security, Endogenous Retirement and Intrahousehold Cooperation," Working paper.

Gustman, A. L. and T. L. Steinmeier (1986): "A Structural Retirement Model," *Econometrica*, 54(3), 555–584.

——— (2000): "Retirement in Dual-Career Families: A Structural Model," *Journal of Labor Economics*, 18(3), 503–545.

——— (2004): "Social Security, Pensions and Retirement Behaviour within the Family," *Journal of Applied Econometrics*, 19(6), 723–737.

——— (2005): "The social security early entitlement age in a structural model of retirement and wealth," *Journal of Public Economics*, 89(2-3), 441–463.

——— (2009): "Integrating Retirement Models," Working Paper 15607, NBER.

Haan, P. and V. Prowse (2010): "A structural approach to estimating the effect of taxation on the labour market dynamics of older workers," *The Econometrics Journal*, 13(3), 99–125.

Heckman, J. J. (1978): "Dummy Endogenous Variables in a Simultaneous Equation System," *Econometrica*, 46(4), pp. 931–959.

Hurd, M. D. (1990): "The Joint Retirement Decision of Husbands and Wives," in *Issues in the Economics of Aging*, ed. by D. A. Wise, chap. 8, pp. 231–258. University of Chicago Press.

Iskhakov, F. (2010): "Structural dynamic model of retirement with latent health indicator," *The Econometrics Journal*, 13(3), 126–161.

Jia, Z. (2004): "Retirement Behavior of Working Couples in Norway: A Dynamic Programming Approach," Discussion Paper 405, Statistics Norway, Research Department.

——— (2005): "Labor Supply of Retiring Couples and Heterogeneity in Household Decision-Making Structure," *Review of Economics of the Household*, 3, 215–233.

Jørgensen, M. (2009): "An effect evaluation of the reform of the early retirement scheme - The reforms effect on the labour supply," SFI report 09:22, The Danish National Center for Research, (In Danish).

Lumsdaine, R. L., J. H. Stock and D. A. Wise (1992): "Three Models of Retirement: Computational Complexity versus Predictive Validity," in *Topics in the Economics of Aging*, ed. by D. A. Wise, chap. 1, pp. 21–60. University of Chicago Press.

——— (1994): "Pension Plan Provisions and Retirement: Men and Women, Medicare, and Models," in *Studies in the Economics of Aging*, ed. by D. A. Wise, chap. 6, pp. 183–222. University of Chicago Press.

Mastrogiacomo, M., R. Alessie and M. Lindeboom (2004): "Retirement Behaviour of Dutch Elderly Households," *Journal of Applied Econometrics*, 19(6), 777–793.

Michaud, P.-C. and F. Vermeulen (2011): "A collective labor supply model with complementarities in leisure: Identification and estimation by means of panel data," *Labour Economics*, 18(2), 159—167.

Nielsen, H. and M. Svarer (2006): "Educational Homogamy: Preferences or Opportunities?," Working Paper 2006-10, Aarhus Universitet.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55(5), 999–1033.

——— (1990): "Behavior of Male Workers at the End of the Life Cycle: An Empirical Analysis of States and Controls," in *Issues in the Economics of Aging.*, ed. by D. A. Wise, chap. 11, pp. 317–382. University of Chicago Press.

——— (1994): "Structural estimation of markov decision processes," in *Handbook of Econometrics*, ed. by R. F. Engle and D. McFadden, vol. 4, chap. 51, pp. 3081–3143. Elsevier, 1 edn.

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

Schjerning, B. (2006): "A Dynamic Model for Occupational Choice, Savings and Formation of Entrepreneurial Human Capital under Liquidity Constraints and Income Uncertainty," Working paper, University of Copenhagen, Center for Applied Microeconometrics.

Schumaker, L. L. (1983): "On Shape Preserving Quadratic Spline Interpolation," *SIAM Journal on Numerical Analysis*, 20(4), 854–864.

Stock, J. H. and D. A. Wise (1990): "Pensions, the Option Value of Work, and Retirement," *Econometrica*, 58(5), 1151–1180.

van der Klaauw, W. and K. I. Wolpin (2008): "Social security and the retirement and savings behavior of low-income households," *Journal of Econometrics*, 145(1-2), 21–42.

A Descriptive Statistics

Table A1 – Descriptive Statistics.

	Married Males		Married Females		Single Males		Single Females	
	Mean	Std.	Mean	Std.	Mean	Std.	Mean	Std.
Age	60.360	2.685	58.283	3.324	60.557	2.788	60.734	2.813
Income	279	194	210	140	221	188	208	164
Net Wealth (household)	5,033	3,309	5,033	3,309	3,036	2,722	2,862	2,575
Pension wealth	1,239	1,407	649	970	947	1,254	943	1,181
Share of wealth	.220	.188	.114	.140	.308	.279	.346	.308
Retire	.570		.530		.587		.628	
Age of retirement	61.940	1.790	61.038	1.406	61.654	1.889	61.884	1.891
Eligible	.825		.660		.739		.808	
Age of eligibility	60.592	.556	60.456	.521	60.618	.921	60.676	.700
# Obs	578,298		578,298		150,674		229,511	
# Households	87,536		87,536		25,984		36,803	

Notes: The table reports data across all years, ranging from 1996-2008. "Eligible" and "Retire" refers to wheter the individual is eligible for early retirement by the age of 64 and whether the individual retire in the observed sample, respectively. Income and wealth is measured in 1,000 DKK 2008 prices.

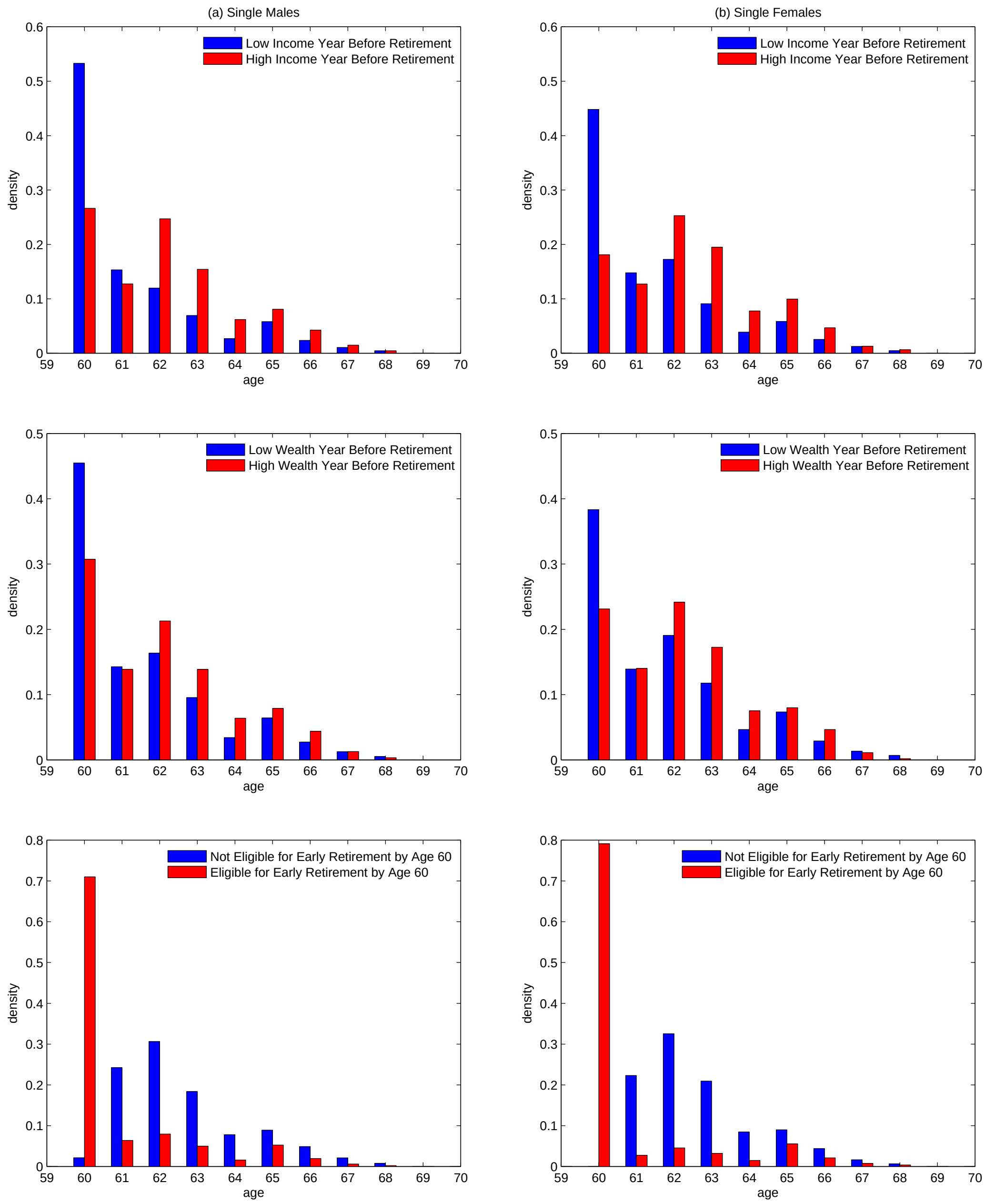

Figure A1 – Single's Retirement Age Across Income, Wealth and Eligibility for Early Retirement.

B Implemented Institutional Settings

B.1 Old Age Pension

Due to these simplifying assumptions mentioned in Section 3, the OA only depend upon individual income, potential spousal income and whether the spouse is retired,

$\mathbf{OA}(y^m, y^f, d)$, and can be formulated as

$$
\begin{aligned}
OA_B &= \mathbf{1}(y_i < \overline{y}_B)\max\{0, (B - \tau_B \max\{0, y_i - \overline{D}_B\})\}, \\
y_h &= y_i + y_s - .5\min\{\overline{D}_{y_s}, y_s\}\mathbf{1}(j = 3), \\
OA_A &= \mathbf{1}(y_h < \overline{y}_j)\max\{0, (A_j - \max\{0, \tau_j(y_h - \overline{D}_j))\}, \\
OA &= OA_B + OA_A,
\end{aligned}
$$

where

$$
j = \begin{cases} 1 & \text{if } single \neq 0, \\ 2 & \text{if } single = 0, d_t^s = 0, \\ 3 & \text{if } single = 0, d_t^s = 1, age_t^s \geq 65, \end{cases}
$$

with the parameters of the scheme given in Table A2. Figure A2 plot the OA for two levels of spousal income.

Table A2 – Old Age Pension Parameters, $\tau_{\mathbf{T}}$.

Symbol	Value in 2008	Description
y_i	-	Income of individual
OA_B	-	Old age pension, main part
B	$61,152 \approx \$10,700$	Base value of old age pension
$\overline{y}_B$	$463,500 \approx \$81,000$	Maximum annual income before loss of OA_B
τ_B	.3	Marginal reduction in deduction regarding income
$\overline{D}_B$	$259,700 \approx \$45,500$	Deduction regarding base value of OA
OA_A	-	Additional old age pension on top of base value
y_s	-	Spousal income
y_h	-	Household income to be tested
$\overline{D}_{y_s}$	$179,400 \approx \$31,500$	Maximum deduction in spousal income
A_j	$\begin{cases} 61,560 \approx \$10,800 \\ 28,752 \approx \$5,000 \\ 28,752 \approx \$5,000 \end{cases}$	Maximum OA_A, for $j = 1, 2, 3$.
$\overline{y}_j$	$\begin{cases} 262,500 \approx \$46,000 \\ 210,800 \approx \$37,000 \\ 306,600 \approx \$54,000 \end{cases}$	Maximum income before loss of OA_A, for $j = 1, 2, 3$.
τ_j	$\begin{cases} .30 \\ .15 \\ .30 \end{cases}$	Marginal reduction in OA_A, for $j = 1, 2, 3$.
$\overline{D}_j$	$\begin{cases} 57,300 \approx \$10,000 \\ 115,000 \approx \$20,000 \\ 115,000 \approx \$20,000 \end{cases}$	Maximum deduction regarding OA_A, for $j = 1, 2, 3$.

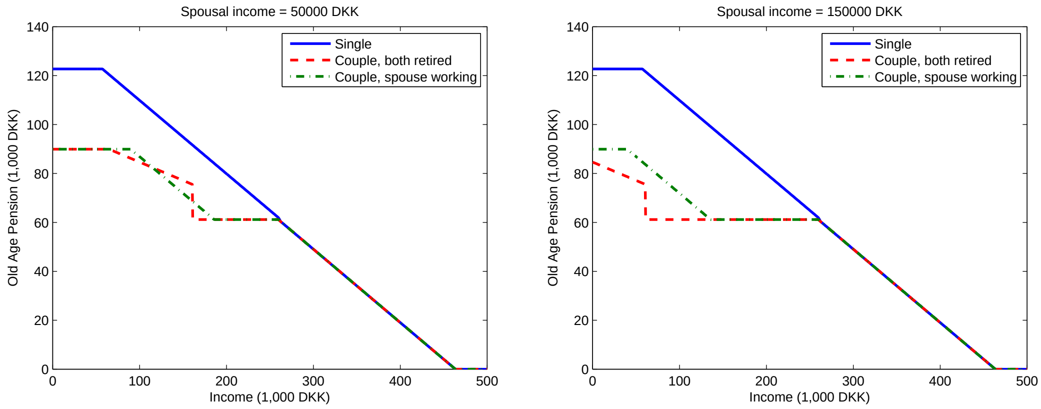

Figure A2 – Old Age Pension (OA) as a Function of Income.

B.2 Tax System

The after tax income can be calculated by applying the following formulas:

$$\begin{aligned}
\tau_{\max} &= \tau_l + \tau_m + \tau_u + \tau_c + \tau_h - \bar{\tau}, \\
\text{personal income} &= (1 - \tau_{LMC}) \cdot \text{income} - \text{pension fund contribution}, \\
\text{taxable income} &= \text{personal income} - \min\{WD \cdot \text{income}, \overline{WD}\}, \\
\mathrm{T}_c &= \max\{\tau_c \cdot (\text{taxable income} - \underline{y}_l), 0\}, \\
\mathrm{T}_h &= \max\{\tau_h \cdot (\text{taxable income} - \underline{y}_l), 0\}, \\
\mathrm{T}_l &= \max\{\tau_l \cdot (\text{personal income} - \underline{y}_l), 0\}, \\
\mathrm{T}_m &= \max\{\tau_m \cdot (\text{personal income} - \underline{y}_m), 0\}, \\
\mathrm{T}_u &= \max\{\min\{\tau_u, \tau_{\max}\} \cdot (\text{personal income} - \underline{y}_u), 0\}, \\
\text{after tax income} &= (1 - \tau_{LMC}) \cdot \text{income} - \mathrm{T}_c - \mathrm{T}_h - \mathrm{T}_l - \mathrm{T}_m - \mathrm{T}_u,
\end{aligned}$$

where the values from 2008 along with descriptions are given in Table A3 and Figure A3 plots the tax schedule dependence on income.

Table A3 – Tax System Parameters, $\tau_{\mathbf{Y}}$, in 2008.

Symbol	Value in 2008	Description
$\bar{\tau}$	.59	Maximum tax rate, »Skatteloft«
τ_{LMC}	.08	Labor Market Contribution, »Arbejdsmarkedsbidrag«
WD	.04	Working Deduction, »Beskæftigelsesfradrag«
$\overline{WD}$	$12,300 \approx \$2,200$	Maximum deduction possible
τ_c	.2554	Average county-specific tax rate (including .073 in church tax)
$\underline{y}_l$	$41,000 \approx \$7,500$	Amount deductible from all income
$\underline{y}_m$	$279,800 \approx \$50,800$	Amount deductible from middle tax bracket
$\underline{y}_u$	$335,800 \approx \$61,000$	Amount deductible from top tax bracket
τ_h	.08	Health contribution tax (in Danish »Sundhedsbidrag«)
τ_l	0.0548	Tax rate in lowest tax bracket
τ_m	0.06	Tax rate in middle tax bracket
τ_u	0.15	Tax rate in upper tax bracket

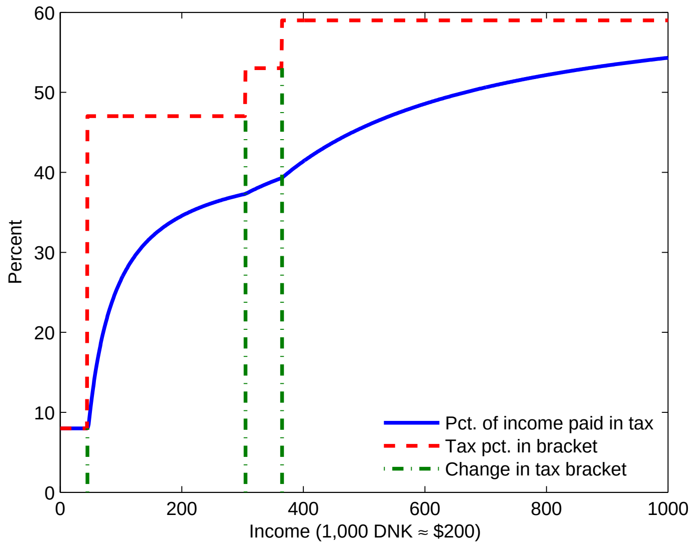

Figure A3 – Implemented Danish Tax System.

C Estimation of Death Probabilities

The data used for estimation are the time tables BEF5 and FOD207 supplied by Statistics Denmark. In these tables, only data up to age 98 is available. See Table A5 for a sample of the used data.

The fit of the "model" with a constant and age is surprisingly good, as can be seen in Table A4 and Figure A4. As expected, the death probability is always greater for males.

Table A4 – Death Probability Estimates, θ_π.

	Males	Females
	Estimate (SE)	Estimate (SE)
constant	-10.338 (.036)***	-11.142 (.039)***
age	.097 (.001)***	.103 (.001)***
$\bar{R}^2$	.996	.996
#Obs	245	245

Data is based on Statistics Denmark's series BEF5 and FOD207 for the years 2006-2010. Consult Table A5 for a sample of the used data. Robust standard errors reported. *: $p < .05$, **: $p < .001$, ***: $p < .0001$.

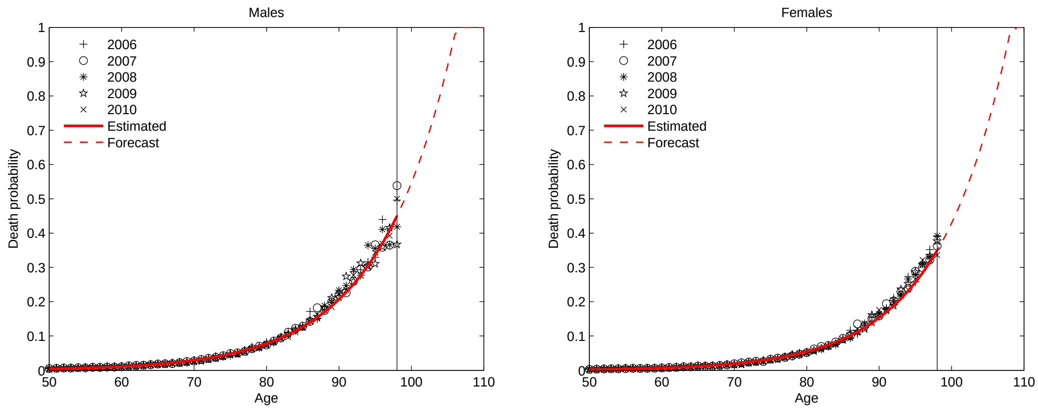

Figure A4 – Actual and Predicted Death Probabilities, 2006-2010.

The out-of sample predictions (individuals aged 99 or older) are in line with the actual probabilities of death since the oldest males in 2010 were 105 years old and the oldest females were 108 years old.

Table A5 – Death Probability Data, 2008.

	Alive		Deaths		Proportion Died	
age	Males	Females	Males	Females	Males	Females
50	36,930	36,283	172	106	0.47	0.29
51	37,129	36,546	181	120	0.49	0.33
52	36,677	35,973	204	124	0.56	0.34
53	35,542	35,510	214	139	0.60	0.39
54	36,198	35,969	239	143	0.66	0.40
55	35,297	35,264	248	170	0.70	0.48
56	34,825	34,416	271	178	0.78	0.52
57	35,340	35,754	252	178	0.71	0.50
58	35,056	35,434	321	216	0.92	0.61
59	36,890	36,960	326	221	0.88	0.60
60	38,982	39,133	415	269	1.06	0.69
61	40,313	39,960	456	269	1.13	0.67
62	38,560	38,451	513	325	1.33	0.85
63	36,117	36,486	516	346	1.43	0.95
64	32,600	33,689	536	325	1.64	0.96
65	30,543	31,314	562	356	1.84	1.14
66	26,640	27,887	537	354	2.02	1.27
67	25,473	26,960	510	345	2.00	1.28
68	23,993	25,371	578	360	2.41	1.42
69	23,211	25,086	510	400	2.20	1.59
70	21,586	24,185	582	392	2.70	1.62
71	20,516	22,785	642	451	3.13	1.98
72	18,944	21,551	628	517	3.32	2.40
73	17,834	20,746	643	490	3.61	2.36
74	16,450	19,430	655	518	3.98	2.67
75	15,393	19,153	663	611	4.31	3.19
76	14,537	18,218	720	624	4.95	3.43
77	13,773	17,726	787	699	5.71	3.94
78	12,901	16,838	803	676	6.22	4.01
79	12,298	16,659	814	748	6.62	4.49
80	10,884	15,634	847	804	7.78	5.14
81	10,338	15,169	838	886	8.11	5.84
82	9,235	14,585	866	882	9.38	6.05
83	8,427	13,860	865	960	10.26	6.93
84	7,159	12,905	844	967	11.79	7.49
85	6,235	11,419	788	1030	12.64	9.02
86	5,635	11,236	800	1047	14.20	9.32
87	4,794	10,232	721	1138	15.04	11.12
88	3,531	7,844	665	1071	18.83	13.65
89	3,037	7,001	602	988	19.82	14.11
90	2,248	5,869	525	964	23.35	16.43
91	1,877	4,986	463	884	24.67	17.73
92	1,362	3,960	400	808	29.37	20.40
93	1,085	3,393	300	755	27.65	22.25
94	757	2,580	276	687	36.46	26.63
95	569	1,994	202	555	35.50	27.83
96	370	1,404	152	434	41.08	30.91
97	243	1,069	89	355	36.63	33.21
98	141	699	59	274	41.84	39.20

Data is based on Statistics Denmark's series BEF1 (population 1st of January) and FOD207 (deaths) for the year 2008. The probability of death is calculated as the ratio "number of deaths during the year"/"number of individuals alive 1st of january that year"

D Private Pension Share of Wealth

To restrict the estimated fraction to be on the $[0, 1]$ domain, the parameters are estimated by OLS on the transformed response, $\tilde{\wp} = \log\left((\wp + 10^{-15})/(1 - \wp + 10^{-15})\right)$.[20]

Further, in order to ensure that individual private pension wealth of marrieds are consistent with total household wealth, I estimate the household fraction of private pension wealth to total wealth, $\wp^h$, and the male fraction of private pension wealth to total household pension wealth, $\wp_p^m$. Using the identity $\wp^h = \wp_p^m \wp^h + (1 - \wp_p^m)\wp^h$ each spouse's fraction of private pension wealth consistent with total household wealth can be calculated as $\wp^m = \wp_p^m \wp^h$ and $\wp^f = (1 - \wp_p^m)\wp^h$. Since estimation is carried out using the transformed response variables, $(\hat{\wp}^m, \hat{\wp}^f, \hat{\wp}^h) \in [0; 1]$.

The estimated parameters are presented in Table A6. Note, the parameters of the first two columns (couples) is not directly comparable to the estimated coefficients for singles, as discussed in Section 4.3. Figure A5 plots the approximation error. The fit of the model looks reasonable, albeit a slight tendency to underestimating the private pension shares for singles.

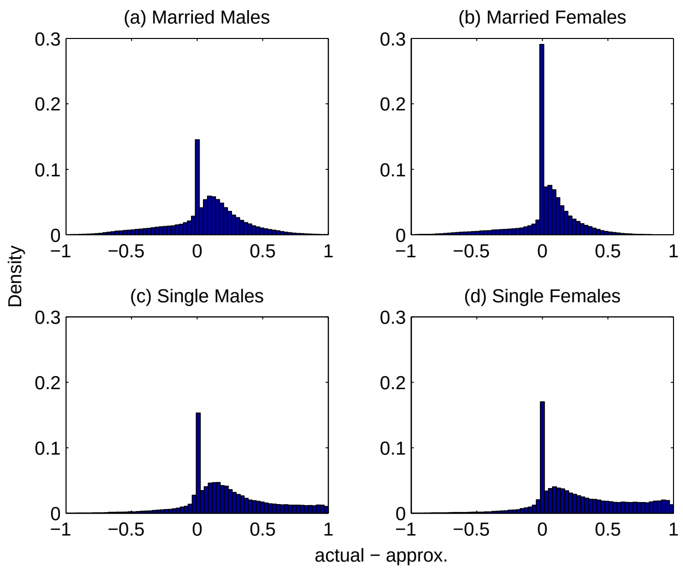

Figure A5 – Approximation Error, Share of Wealth in Private Pension.

[20] Alternatively, the fraction could be estimated in a double censored Tobit framework. See Appendix D.1 on page 40 for results using the double censored approach. Since the cumulative normal distribution, $\Phi(\cdot)$, does not have a closed form, numerical integration at each point in the state space would have to be applied when predicting the share of private pension wealth using the double censored regression approach. This is a rather costly operation and ultimately lead to the implementation of the transformation-approach.

Table A6 – Estimates of Private Pension Wealth Share of Total Wealth.

	Couples		Singles	
	Household Share, $\tilde{\wp}^h$	Males Share, $\tilde{\wp}_p^m$	Males, $\tilde{\wp}^m$	Females, $\tilde{\wp}^f$
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
$age_t^m = 60$	.329 (.004) ***	-.138 (.008)***	.400 (.012)***	
$age_t^m = 61$	.461 (.006) ***	-.513 (.014)***	.365 (.017)***	
$age_t^m = 62$	.473 (.006) ***	-.730 (.015)***	.301 (.018)***	
$age_t^m = 63$	.478 (.007) ***	-.970 (.017)***	.225 (.020)***	
$age_t^m = 64$	.495 (.007) ***	-1.259 (.018)***	.123 (.021)***	
$age_t^m = 65$	.404 (.007) ***	-1.485 (.018)***		
$age_t^m > 65$	.378 (.007) ***	-2.141 (.021)***		
$d_t^m = 1$	.072 (.003) ***	.022 (.009)*	-.035 (.012)*	
y_t^m	-.028 (.001) ***	.023 (.002)***	-.006 (.003)*	
$e_t^m > 0$	.673 (.008) ***	-.323 (.016)***	.246 (.019)***	
$e_t^m = 2$	.016 (.003) ***	.009 (.010)	.015 (.014)	
$age_t^f = 60$	.036 (.003) ***	-.598 (.008)***		.456 (.009)***
$age_t^f = 61$	.010 (.005) *	-.369 (.012)***		.428 (.016)***
$age_t^f = 62$	-.039 (.005) ***	-.170 (.013)***		.366 (.016)***
$age_t^f = 63$	-.088 (.006) ***	.049 (.016)*		.297 (.018)***
$age_t^f = 64$	-.144 (.007) ***	.279 (.018)***		.213 (.019)***
$age_t^f = 65$	-.243 (.008) ***	.994 (.025)***		
$age_t^f > 65$	-.266 (.008) ***	1.538 (.035)***		
$d_t^f = 1$	-.023 (.004) ***	-.079 (.009)***		-.078 (.010)***
y_t^f	.007 (.001) ***	.033 (.002)***		.036 (.003)***
$e_t^f > 0$	.775 (.009) ***	-.569 (.019)***		.252 (.016)***
$e_t^f = 2$	-.002 (.004)	-.017 (.014)		.010 (.012)
$e_t^m, e_t^f > 0$	-.904 (.009) ***	.631 (.021)***		
$e_t^m = e_t^f = 2$	-.008 (.009)	.108 (.021)***		
$wealth_t, e_t^m > 0$	-1.201 (.008) ***	.826 (.019)***	-.682 (.023)***	
$wealth_t, e_t^f > 0$	-1.185 (.009) ***	1.007 (.020)***		-.798 (.020)***
$wealth_t, e_t^m, e_t^f > 0$	1.304 (.011) ***	-1.029 (.026)***		
$age_t^m > age_t^f$	-.025 (.005) ***	.715 (.006)***		
$age_t^m < age_t^f$	.225 (.005) ***	-.088 (.008)***		
$wealth_t$	2.786 (.014) ***	-.663 (.023)***	2.600 (.035)***	2.618 (.030)***
$wealth_t^2$	-.978 (.008) ***	.036 (.014)*	-1.164 (.025)***	-1.206 (.023)***
$Constant$	-1.645 (.008) ***	.720 (.014)***	-1.314 (.014)***	-1.413 (.012)***
$\bar{R}^2$	.227	.165	.102	.099
$\#Obs$	517,298	469,162	113,466	176,196

Notes: Estimates based on individuals aged under 65 who are eligible for early retirement by the age of 64 or earlier. For couples, one spouse has to meet these criteria to be in the used subsample and the male fraction of household private pension wealth (column two) is based only on households who has private pension wealth. Household wealth is measured in 10,000,000 DKK and income in 100,000 DKK. *: $p < .05$, **: $p < .001$, ***: $p < .0001$.

D.1 Alternative Double Censored Approach

Here, the fraction of private pension wealth to total wealth estimated in Section 12 by a transformation of the response variable, is estimated by a double censored Tobit regression model. This is done to illustrate, the performance of this model relative to the one used in the tructural model. Even though the double censored model does seem to predict the actual shares better (see Figure 41) for fractions close to one, the used transformation approach is applied to avoid costly numerical integration of the cummulative normal density function, $\Phi(\cdot)$, in the equation below.

When predicting the fractions in the double censored regression model, the double truncated normal distribution yeilds the formula:

$$\hat{\wp} \equiv \mathbb{E}[\wp] = \left(\mathbf{x}\hat{\beta} + \hat{\sigma}\Lambda(\mathbf{x}\hat{\beta}/\hat{\sigma})\right)\left(1 - \Phi\left((1-\mathbf{x}\hat{\beta})/\hat{\sigma}\right) - \Phi\left(\mathbf{x}\hat{\beta}/\hat{\sigma}\right)\right) + 1 - \Phi\left((1-\mathbf{x}\hat{\beta})/\hat{\sigma}\right),$$

where

$$\Lambda(\mathbf{x}\hat{\beta}/\hat{\sigma}) = \frac{\phi\left((1-\mathbf{x}\hat{\beta})/\hat{\sigma}\right) - \phi\left(\mathbf{x}\hat{\beta}/\hat{\sigma}\right)}{1 - \Phi\left((1-\mathbf{x}\hat{\beta})/\hat{\sigma}\right) - \Phi\left(\mathbf{x}\hat{\beta}/\hat{\sigma}\right)}.$$

Figure A6 – Error in Predicting Share of Wealth in Private Pension.

Table A7 – Tobit Estimates of Private Pension Wealth Share of Total Wealth.

	Couples		Singles	
	Household Share, $\wp^c$	Males Share, $\wp_p^m$	Males, $\wp^m$	Females, $\wp^f$
	Estimate (SE)	Estimate (SE)	Estimate (SE)	Estimate (SE)
$age_t^m = 60$	.053 (.001) ***	-.013 (.002)***	.116 (.003)***	
$age_t^m = 61$	.095 (.002) ***	-.081 (.003)***	.195 (.005)***	
$age_t^m = 62$	.093 (.002) ***	-.115 (.003)***	.192 (.005)***	
$age_t^m = 63$	.091 (.002) ***	-.146 (.004)***	.193 (.006)***	
$age_t^m = 64$	.092 (.003) ***	-.193 (.004)***	.179 (.007)***	
$age_t^m = 65$	.082 (.003) ***	-.290 (.005)***		
$age_t^m > 65$	.047 (.003) ***	-.445 (.005)***		
$d_{t-1}^m = 1$	-.022 (.001) ***	.000 (.002)	-.026 (.004)***	
y_{t-1}^m	.012 (.000) ***	.029 (.000)***	.023 (.001)***	
$e_t^m > 0$	-.022 (.002) ***	-.091 (.004)***	-.132 (.005)***	
$e_t^m = 2$	-.008 (.002) ***	-.034 (.003)***	-.030 (.005)***	
$age_t^f = 60$	.029 (.001) ***	-.117 (.002)***		.113 (.003)***
$age_t^f = 61$	.026 (.002) ***	-.090 (.003)***		.161 (.005)***
$age_t^f = 62$	.025 (.002) ***	-.076 (.003)***		.162 (.005)***
$age_t^f = 63$	.019 (.003) ***	-.042 (.004)***		.156 (.006)***
$age_t^f = 64$	.014 (.003) ***	-.013 (.005)*		.155 (.006)***
$age_t^f = 65$	-.013 (.004) **	.120 (.006)***		
$age_t^f > 65$	-.033 (.005) ***	.210 (.009)***		
$d_{t-1}^f = 1$	-.034 (.001) ***	.038 (.002)***		-.063 (.004)***
y_{t-1}^f	.024 (.000) ***	-.059 (.001)***		.049 (.001)***
$e_t^f > 0$	.031 (.003) ***	.072 (.004)***		-.040 (.005)***
$e_t^f = 2$	-.022 (.002) ***	.017 (.004)***		-.024 (.005)***
$e_t^m, e_t^f > 0$	-.022 (.003) ***	-.024 (.005)***		
$e_t^m = e_t^f = 2$	.006 (.004)	.045 (.006)***		
$wealth_{t-1}, e_t^m > 0$	-.091 (.003) ***	.225 (.005)***	-.066 (.008)***	
$wealth_{t-1}, e_t^f > 0$	-.145 (.004) ***	-.026 (.006)***		-.184 (.007)***
$wealth_{t-1}, e_t^m, e_t^f > 0$	.089 (.005) ***	.007 (.007)		
$age_t^m > age_t^f$	-.026 (.001) ***	.116 (.002)***		
$age_t^m < age_t^f$	.036 (.001) ***	-.026 (.002)***		
$wealth_{t-1}$	.495 (.004) ***	.043 (.006)***	.035 (.011)*	-.034 (.010)**
$wealth_{t-1}^2$	-.173 (.002) ***	-.084 (.004)***	.040 (.009)***	.074 (.008)***
$Constant$	.123 (.002) ***	.686 (.004)***	.227 (.005)***	.226 (.004)***
σ	.227 (.000) ***	.347 (.000)***	.310 (.001)***	.348 (.001)***
$\bar{R}^2$	.679	.132	.057	.044
$\#Obs$	491,757	446,364	99,515	161,152

Notes: Estimates based on individuals aged under 65 who are eligible for early retirement by the age of 64 or earlier. For couples, one spouse has to meet these criteria to be in the used subsample and the male fraction of household private pension wealth (column two) is based only on households who has private pension wealth. Household wealth is measured in 10,000,000 DKK and income in 100,000 DKK. *: $p < .05$, **: $p < .001$, ***: $p < .0001$.

E Solving the Model by EGM

In order to solve the model, I assume that both spouses die with probability one when the male is 100 years old. Furthermore, to speed up the solution algorithm, I assume forced retirement at the age of 70 years old.

Interpolation of consumption- and value-functions are by linear spline. Since linear extrapolation of value functions can result in serious approximation errors, the linear spline is applied to a transformed value function, following the ideas in Carroll (2011). Since the value function "inherits" the curvature from the utility function (see Carroll and Kimball, 1996) I interpolate $\tilde{\mathbf{v}} = (\mathbf{v}(1-\rho))^{1/(1-\rho)}$ and re-transform the resulting interpolated data, such that $\check{\mathbf{v}} = (\check{\tilde{\mathbf{v}}})^{(1-\rho)}/(1-\rho)$, where $\check{}$ indicates a linear interpolation function.[21]

Since the solution method is similar for singles and couples, the following will focus on inplementation of EGM for the model of couples. For the ease of exposition, the notation in the following is going to leave out all other state variables than cash-on-hand, m_t, and the labor market status this period, d_t. Therefore, it is convenient to bear in mind the budget and the relationships between the different elements stated in equations (4.4)-(4.6) on page 11 as well as d_{t+1} is the discrete choice at time t.[22]

Solution at time T

In the last period of life households know they will both be dead with probability $(1-\pi^f_{T+1})(1-\pi^m_{T+1}) \equiv 1$ in the next period and since agents are forced to retire at $t \geq T_r$ they only chose the optimal consumption in the last period of life. Therefore, the value function in the last period can be formulated as

$$\mathbf{V}_T(m_T, d_T) = \max_{0 \leq c_T \leq m_T} \left\{ \mathbf{U}(c_T, \mathbf{0}, \mathbf{0}) + \beta \mathbf{B}((1+r)(m_T - c_T)) \right\},$$

where the first order condition is given by

$$\mathbf{U}'(c_T, \mathbf{0}, \mathbf{0}) = (1+r)\beta \mathbf{B}'((1+r)(m_T - c_T)). \tag{E.1}$$

Inserting the partial derivative of (4.1) and (4.8) in (E.1) and defining $\ell_T \equiv \lambda l^m(0)^{(1-\eta)(1-\rho)} e^{\alpha' \mathbf{x}^m_T} + (1-\lambda) l^f(0)^{(1-\eta)(1-\rho)} e^{\alpha' \mathbf{x}^f_T}$ yields the closed form solution to the last period problem as

$$c^*_T(m_T) = \left(1 + (1+r)\left(\frac{(1+r)\beta\gamma}{\ell_T}\right)^{\frac{1}{\eta(1-\rho)-1}}\right)^{-1} \left(\frac{(1+r)\beta\gamma}{\ell_T}\right)^{\frac{1}{\eta(1-\rho)-1}} ((1+r)m_T + \kappa). \tag{E.2}$$

Note, the level of cash-on-hand given by $m^{s_T=0}_T = \kappa \left(\frac{(1+r)\beta\gamma}{\ell_T}\right)^{\frac{1}{\eta(1-\rho)-1}}$ is consistent with

[21] Alternatively, the shape preserving piecewise cubic spline proposed by Schumaker (1983) was implemented without any noticeable difference on the results while slowing the solution algorithm significantly down. Further, to increase the accuracy of the approximated curvature of the consumption- and value function, the wealth and income grids used when solving the model is unequally spaced, with more points at the lower end of the distributions.

[22] In the following couples will be assumed to be of the same age. This is only for readability, since keeping track of agedifferences does not add any intuition to the solution method.

no savings, i.e., $c_T = m_T$ iff $m_T \leq m_T^{s_T=0}$.

Solution at time $T_r - 1 \leq t < T$

Since agents are forced into retirement in the considered periods, the household only chose the level of consumption.[23] Therefore, the value function in these periods can be formulated as

$$\begin{aligned} \mathbf{V}_t(m_t, d_t) &= \max_{0 \leq c_t \leq m_t} \Big\{ \mathbf{U}(c_t, d_t, \mathbf{0}) + \beta \mathbb{E}_t \Big[\pi^f_{t+1}\pi^m_{t+1}\mathbf{V}_{t+1}(m_{t+1}, d_{t+1}) + (1-\pi^f_{t+1})(1-\pi^m_{t+1})\mathbf{B}(a_{t+1}) \\ &\qquad + \pi^f_{t+1}(1-\pi^m_{t+1})\mathbf{V}^f_{t+1}(m^f_{t+1}, d^f_{t+1}) + (1-\pi^f_{t+1})\pi^m_{t+1}\mathbf{V}^m_{t+1}(m^m_{t+1}, d^m_{t+1}) \Big] \Big\} \qquad \text{(E.3)} \\ \text{s.t.} &\qquad a_{t+1} = (1+r)(m_t - c_t) \geq 0, \end{aligned}$$

where m^j_t is the cash-on-hand for spouse j if single in period t. This distinction is necessary, since the cash-on-hand available for consumption next period depend on whether the household consist of one or two people. Note, however, household assets are passed on to the widowed spouse without any costs. The consumption function is found in a similar way as for time periods prior to $T_r - 1$, by inserting $d_{t+1} = (0,0)$ in equation (E.8) below.

Solution at time $t < T_r - 1$

Prior to forced retirement, the household is choosing the optimal household consumption, c_t, and labor choice of each spouse *next* period, $d_{t+1} = (d^m_{t+1}, d^f_{t+1})$. Using the value function in (5.2), the problem can be reformulated using notation inspired by Carroll (2006) as

$$\mathbf{V}_t(m_t, d_t, \varepsilon_t) = \max_{\substack{0 \leq c_t \leq m(\mathbf{z}_t) \\ d_{t+1} \in \{1,2,3,4\}}} \{\mathbf{U}(c_t, d_t, d_{t+1}) + \varepsilon_t(d_{t+1}) + \mathfrak{v}_t(s_t, d_{t+1})\} \qquad \text{(E.4)}$$

where the *expected marginal utility from savings* is

$$\begin{aligned} \mathfrak{v}_t(s_t, d_{t+1}) &\equiv \beta \mathbb{E}_t \Big[\pi^f_{t+1}\pi^m_{t+1}\mathbf{V}_{t+1}(m_{t+1}, d_{t+1}, \varepsilon_{t+1}) + \pi^f_{t+1}(1-\pi^m_{t+1})\mathbf{V}^f_{t+1}(m^f_{t+1}, d^f_{t+1}, \varepsilon_{t+1}) \\ &\qquad + (1-\pi^f_{t+1})\pi^m_{t+1}\mathbf{V}^m_{t+1}(m^m_{t+1}, d^m_{t+1}, \varepsilon_{t+1}) + (1-\pi^f_{t+1})(1-\pi^m_{t+1})\mathbf{B}(a_{t+1}) \Big] \qquad \text{(E.5)} \end{aligned}$$

with the derivative w.r.t. s_t given by

$$\mathfrak{v}'_t(s_t, d_{t+1}) = \beta \mathbb{E}_t \Big[\mathbf{r}\mathbf{V}'_{t+1}(m_{t+1}, d_{t+1}) + \mathbf{r}^m \mathbf{V}'^m_{t+1}(m^m_{t+1}, d^m_{t+1}) + \mathbf{r}^f \mathbf{V}'^f_{t+1}(m^f_{t+1}, d^f_{t+1}) + \mathbf{r}^b \mathbf{B}'(a_{t+1}) \Big]. \qquad \text{(E.6)}$$

[23] Note, due to the timing of this model, people are only choosing consumption at time $T_r - 1$, since their choice over labor market status is d_{T_r} and is forced to retirement.

The transfer- and mortality- adjusted interest rates are defined as

$$\begin{aligned}
\mathbf{r} &\equiv (1+r)\left(1+\mathbf{T}'(\mathbf{z}_{t+1})\right)\pi^f_{t+1}\pi^m_{t+1},\\
\mathbf{r}^m &\equiv (1+r)\left(1+\mathbf{T}'(\mathbf{z}^m_{t+1})\right)(1-\pi^f_{t+1})\pi^m_{t+1},\\
\mathbf{r}^f &\equiv (1+r)\left(1+\mathbf{T}'(\mathbf{z}^f_{t+1})\right)(1-\pi^m_{t+1})\pi^f_{t+1},\\
\mathbf{r}^b &\equiv (1+r)(1-\pi^f_{t+1})(1-\pi^m_{t+1}),
\end{aligned}$$

where π^j_{t+1} is based on the estimated survival probabilities in Appendix C, and $\mathbf{T}'(\mathbf{z}_{t+1}) = \partial\mathbf{T}(\mathbf{z}_{t+1})/\partial s_t$.

Returning to the value function in (E.4), the first order condition is given by $\mathbf{U}'(c_t, d_t) = \mathfrak{v}'_t(s_t, d_{t+1})$ and the envelope theorem yields

$$\begin{aligned}
\frac{\partial \mathbf{V}_t(m_t, d_t)}{\partial m_t} &= \beta\mathbb{E}_t\left[\mathbf{r}\mathbf{V}'_{t+1}(m_{t+1}, d_{t+1}) + \mathbf{r}^m\mathbf{V}'^m_{t+1}(m^m_{t+1}, d^m_{t+1}) + \mathbf{r}^f\mathbf{V}'^f_{t+1}(m^f_{t+1}, d^f_{t+1}) + \mathbf{r}^b\mathbf{B}'(a_{t+1})\right]\\
&= \mathfrak{v}'_t(s_t, d_{t+1}),
\end{aligned}$$

such that we must have $\mathbf{U}'(c_{t+1}, d_{t+1}) = \mathbf{V}'_{t+1}(m_{t+1}, d_{t+1})$. Hence, the Euler equation w.r.t. consumption is given by

$$\begin{aligned}
\mathbf{U}'(c_t, d_t) &= \mathfrak{v}'_t(s_t, d_{t+1})\\
&= \beta\mathbb{E}_t\left[\mathbf{r}\mathbf{V}'_{t+1}(m_{t+1}, d_{t+1}) + \mathbf{r}^m\mathbf{V}'^m_{t+1}(m^m_{t+1}, d^m_{t+1}) + \mathbf{r}^f\mathbf{V}'^f_{t+1}(m^f_{t+1}, d^f_{t+1}) + \mathbf{r}^b\mathbf{B}'(a_{t+1})\right]\\
&= \beta\mathbb{E}_t\left[\mathbf{r}\mathbf{U}'(c_{t+1}, d_{t+1}) + \mathbf{r}^m\mathbf{U}'^m(c^m_{t+1}, d^m_{t+1}) + \mathbf{r}^f\mathbf{U}'^f(c^f_{t+1}, d^f_{t+1}) + \mathbf{r}^b\mathbf{B}'(a_{t+1})\right]. \qquad \text{(E.7)}
\end{aligned}$$

In stead of solving the nonlinear Euler equation by numerical root finding routines over a grid of c_t (or s_{t-1}), Carroll (2006) suggests defining a grid over s_t and simply calculate the consumption level correspondent to the level of savings. Hence, the optimal consumption can be represented as a function of savings (and labor market choice) as the inverse of the partial derivative of the household utility function, refered to as *the inverse Euler equation*:

$$\hat{c}_t(\hat{s}, d_t, d_{t+1}) = \left(\frac{\mathfrak{v}'_t(\hat{s}, d_{t+1})}{\eta\left(\lambda l_m(d_t)^{(1-\eta)(1-\rho)}e^{\alpha'\mathbf{x}^m_t} + (1-\lambda)l_f(d_t)^{(1-\eta)(1-\rho)}e^{\alpha'\mathbf{x}^f_t}\right)}\right)^{\frac{1}{\eta(1-\rho)-1}} \qquad \text{(E.8)}$$

Since $\mathfrak{v}'_t(\hat{s}, d_{t+1})$ is not know, the marginal utility of savings are approximated as

$$\begin{aligned}
\mathfrak{v}'_t(\hat{s}, d_{t+1}) \;\approx\; & \beta \Bigg[\mathbf{r}^b \mathbf{B}'((1+r)\hat{s}) + \sum_{e^m_{t+1}} P^m_{e=e^m_{t+1}} \sum_{e^f_{t+1}} P^f_{e=e^f_{t+1}} \sum_{y^m_{t+1}} \sum_{y^f_{t+1}} P_{y^m_{t+1}, y^f_{t+1}} \sum_{d^m_{t+2}} \sum_{d^f_{t+2}} \check{P}(d^m_{t+2}, d^f_{t+2} | \mathbf{z}_{t+1}) \\
& \times \mathbf{r} \mathbf{U}'(\check{c}_{t+1}(\mathbf{z}_{t+1}, d^m_{t+2}, d^f_{t+2}, \hat{s}), d_{t+1}) \\
& + \sum_{e^m_{t+1}} P^m_{e=e^m_{t+1}} \sum_{y^m_{t+1}} P_{y^m_{t+1}} \sum_{d^m_{t+2}} \check{P}(d^m_{t+2} | \mathbf{z}^m_{t+1}) \mathbf{r}^m \mathbf{U}'(\check{c}^m_{t+1}(\mathbf{z}^m_{t+1}, d^m_{t+2}, \hat{s}), d^m_{t+1}) \qquad \text{(E.9)} \\
& + \sum_{e^f_{t+1}} P^f_{e=e^f_{t+1}} \sum_{y^f_{t+1}} P_{y^f_{t+1}} \sum_{d^f_{t+2}} \check{P}(d^f_{t+2} | \mathbf{z}^f_{t+1}) \mathbf{r}^f \mathbf{U}'(\check{c}^f_{t+1}(\mathbf{z}^f_{t+1}, d^f_{t+2}, \hat{s}), d^f_{t+1}) \Bigg], \qquad \text{(E.10)}
\end{aligned}$$

where $\check{c}_{t+1}(\cdot)$ is the interpolated consumption function as a function of state variables, $P^j_{e=e^j}$ is the estimated probability of eligibility of individual j being e^j from Section 4.5.2, $P_{y^m_{t+1}, y^f_{t+1}}$ is the estimated income transition probability from Section 4.5.1 (conditional on all state variables and e_{t+1}), and

$$\check{P}(d^m_{t+2}, d^f_{t+2} | \mathbf{z}_{t+1}) \equiv \frac{\exp(\check{\mathbf{v}}_{t+1}(\mathbf{z}_{t+1}, d_{t+2}))}{\sum_{k \in \mathcal{D}(\mathbf{z}_{t+1})} \exp(\check{\mathbf{v}}_{t+1}(\mathbf{z}_{t+1}, d_{t+2} = k))}$$

is the interpolated conditional choice probability of choosing d_{t+2} using the solution from the previous iteration.

Defining the grid on savings, $\hat{s}$, the grid on cash-on-hand is determined "endogenously" by the inverse Euler equation (E.8) and the budget constraint (4.4):

$$\hat{m}(\hat{s}, d_t, d_{t+1}) = \hat{c}(\hat{s}, d_t, d_{t+1}) + \hat{s},$$

yielding the name.

F Maximum Likelihood Estimation of the Model Parameters

In order to derive the log likelihood function, assume that all variables are observed, i.e., ε is also known to the researcher. The joint distribution of c, d, $\mathbf{z}$ and ε can be written, as[24]

$$
\begin{aligned}
F(c,d,\mathbf{z},\varepsilon) &\overset{(1)}{=} \prod_{i=1}^{N} F(c_{i1},\ldots,c_{iT},d_{i1},\ldots,d_{iT},\mathbf{z}_{i1},\ldots,\mathbf{z}_{iT},\varepsilon_{i1},\ldots,\varepsilon_{iT}) \\
&\overset{(2)}{=} \prod_{i=1}^{N}\prod_{t=1}^{T_i} F(c_{it},d_{it+1},\mathbf{z}_{it},\varepsilon_{it}|c_{it-1},d_{it},\mathbf{z}_{it-1},\varepsilon_{it-1}) \\
&\overset{(3)}{=} \prod_{i=1}^{N}\prod_{t=1}^{T_i} F(c_{it}|c_{it-1},d_{it+1},d_{it},\mathbf{z}_{it},\mathbf{z}_{it-1},\varepsilon_{it},\varepsilon_{it-1}) \\
&\qquad \times F(d_{it+1}|c_{it},\mathbf{z}_{it},\mathbf{z}_{it-1},\varepsilon_{it},\varepsilon_{it-1}) \\
&\qquad \times F(\varepsilon_{it}|c_{it-1},\mathbf{z}_{it},\mathbf{z}_{it-1},\varepsilon_{it-1}) \\
&\qquad \times F(\mathbf{z}_{it}|c_{it-1},\mathbf{z}_{it-1},\varepsilon_{it-1}) \\
&\overset{(4)}{=} \prod_{i=1}^{N}\prod_{t=1}^{T_i} F(c_{it}|d_{it+1},\mathbf{z}_{it}) \\
&\qquad \times F(d_{it+1}|\mathbf{z}_{it},\varepsilon_{it}) \\
&\qquad \times F(\varepsilon_{it}|\mathbf{z}_{it}) \\
&\qquad \times F(\mathbf{z}_{it}|\mathbf{z}_{it-1})
\end{aligned} \tag{F.1}
$$

where (1) is due to the assumption of independence across households, (2) is a Markov assumption along with the fact that the left hand side is the joint distribution conditioned on initial values, (3) follows from Bayes formula, and (4) is due to the extended conditional independence (CI) assumption, stated in equation (6.1) in Section 6.

Since we actually do not observe ε the likelihood function can be found by integrating over the unobserved state in (F.1):

$$
F(c,d,\mathbf{z};\Theta) = \prod_{i=1}^{N}\prod_{t=1}^{T_i} F(\mathbf{z}_{it}|\mathbf{z}_{it-1};\Theta_1)\underbrace{F(c_{it}|d_{it+1},\mathbf{z}_{it};\Theta)\overbrace{\int_{\varepsilon} F(d_{it+1}|\mathbf{z}_{it},\varepsilon_{it};\Theta)F(\mathrm{d}\varepsilon_{it}|\mathbf{z}_{it};\Theta)}^{F(d_{it+1}|\mathbf{z}_{it};\Theta)}}_{F(c_{it},d_{it+1}|\mathbf{z}_{it};\Theta)}, \tag{F.2}
$$

where the transition of the states, $F(\mathbf{z}_{it}|\mathbf{z}_{it-1};\Theta_1)$, are discussed and estimated in section 4.5 on page 12.

As mentioned in Section 5 the probability of household i choosing labor status j at time $t+1$ is given by the Multinomial Logit (MNL) formula,

$$
F(d_{it+1}=j|\mathbf{z}_{it};\Theta) = \frac{\exp(\mathbf{v}_t(\mathbf{z}_{it},d_{it+1}=j))}{\sum_{k\in\mathcal{D}(\mathbf{z}_{it})}\exp(\mathbf{v}_t(\mathbf{z}_{it},d_{it+1}=k))}. \tag{F.3}
$$

[24] Since the model is dynamic, the distribution of the initial observations has to be specified or conditioned upon. I condition on the initial values in every distribution, but for notational reasons I do not explicitly state that conditioning. For example, the joint distribution is $F(c,d,\mathbf{z},\varepsilon|c_0,d_0,\mathbf{z}_0,\varepsilon_0)$ but I simply write $F(c,d,\mathbf{z},\varepsilon)$.

In stead of maximizating (F.2), the estimation procedure applied to uncover the parameters of the model is asymptotic equivalent to Full Information Maximum Likelihood (FIML). Since the number of parameters in the model is enormous, the procedure follows the one proposed by Rust (1994): First, the parameters in the transition probabilities of the observed state variables, summarized in $F(\mathbf{z}_{it}|d_{it-1}, \mathbf{z}_{it-1}; \Theta_1)$, are estimated using partial MLE:

$$\hat{\Theta}_1 = \underset{\Theta_1}{\operatorname{argmax}} \mathcal{L}_1(\Theta_1) \equiv \sum_{i=1}^{N} \sum_{t=1}^{T_i} \log(F(\mathbf{z}_{it}|\mathbf{z}_{it-1}; \Theta_1)). \tag{F.4}$$

Secondly, the structural parameters, summarized in $F(d_{it+1}|\mathbf{z}_{it}; \Theta)$, are estimated also using partial MLE conditional on the first step estimates:

$$\hat{\Theta}_2 = \underset{\Theta_2}{\operatorname{argmax}} \mathcal{L}_2(\Theta_2|\hat{\Theta}_1) \equiv \sum_{i=1}^{N} \sum_{t=1}^{T_i} \log(F(d_{it+1}|\mathbf{z}_{it}; \hat{\Theta}_1, \Theta_2)). \tag{F.5}$$

The likelihood function used to estimate the preference parameters in the second step is given by

$$\begin{aligned} \mathcal{L}_2(\Theta_2|\hat{\Theta}_1) &= \log\left(\prod_{i=1}^{N} \prod_{t=1}^{T_i} \prod_{j \in \mathcal{D}(\mathbf{z}_{it})} F(d_{it+1}|\mathbf{z}_{it})\right) \\ &= \log \prod_{i=1}^{N} \prod_{t=1}^{T_i} \prod_{j \in \mathcal{D}(\mathbf{z}_{it})} \left(\frac{e^{\mathbf{v}_t(\mathbf{z}_{it}, d_{it+1}=j)}}{\sum_{k=1}^{K_{it}} e^{\mathbf{v}_t(\mathbf{z}_{it}, d_{it+1}=k)}}\right)^{\mathbf{1}(d_{it+1}=j)} \\ &= \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left[\sum_{j \in \mathcal{D}(\mathbf{z}_{it})} \mathbf{1}\,(d_{it+1}=j)\, \mathbf{v}_t(\mathbf{z}_{it}, d_{it+1}=j) - \underbrace{\log\left(\sum_{k=1}^{K_{it+1}} e^{\mathbf{v}_t(\mathbf{z}_{it}, d_{it+1}=k)}\right)}_{=EV_t(\mathbf{z}_{it})}\right]. \end{aligned}$$