Zhiyang Jia

Retirement Behavior of Working Couples in Norway: A Dynamic Programming Approach

Abstract: This paper aims to provide an empirical analysis of the joint retirement behavior of working couples in Norway. A dynamic programming model is specified and estimated on micro data. The estimation results show that a model, which uses only measures of economic incentives: wages and pension benefits gives a satisfactory fit to the observed retirement pattern. The results also indicate that husbands have higher bargaining power within the household. A hypothetical policy simulation shows that by taxing pension benefits as wage income, the labor market participation of both husbands and wives will increase around 4 percentage points at age 65.

Keywords: Household Retirement, Dynamic Programming

JEL classification: J26

Acknowledgement: I thank John Dagsvik, Erik Hernæs and Steinar Strøm for helpful comments. Financial support from the Research Council of Norway (project 140082/530) is gratefully acknowledged.

Address: Zhiyang Jia, Statistics Norway, Research Department. E-mail: Zhiyang.jia@ssb.no
Discussion Papers comprise research papers intended for international journals or books. A preprint of a Discussion Paper may be longer and more elaborate than a standard journal article, as it may include intermediate calculations and background material etc.

Abstracts with downloadable Discussion Papers in PDF are available on the Internet:
http://www.ssb.no
http://ideas.repec.org/s/ssb/dispap.html

For printed Discussion Papers contact:

Statistics Norway
Sales- and subscription service
NO-2225 Kongsvinger

Telephone: +47 62 88 55 00
Telefax: +47 62 88 55 95
E-mail: Salg-abonnement@ssb.no
1 Introduction

In this paper, we develop and estimate a dynamic programming model that allows structural empirical analysis of joint labor market behavior of working couples in Norway. The empirical analysis is based on working couples in which husbands are qualified for a subsidized early retirement scheme, AFP\textsuperscript{1}, in 1997 or 1998.

This study tries to join together two branches of the retirement literature. In one, various attempts are made to explicitly model the joint retirement decision within a household. In the other, structural dynamic models of retirement are specified and estimated for individuals, while ignoring the retirement decisions and retirement status of their spouse.

Recent studies of retirement behavior have recognized the fact that labor force status and transitions of older married couples are correlated. An older individual is more likely to retire if the individual’s spouse is retired than if the spouse is not retired. Similar patterns have been documented in a number of countries and time periods (Hurd (1990), Coile (2003), Jimenez-Martin, Labeeaga, and Granado (1999)). Spouses are likely to coordinate their exits from the labor market for several reasons. Firstly, the labor market choice of one member may affect the financial rewards of the other member through specific tax and social security rules. Secondly, the preference of one member may be different with different labor market status of the other member, for example, due to the complementarity of leisure between spouses. Thirdly, there may be correlation across spouses in unobserved tastes. It has been argued that it is not possible to understand the retirement decision of one spouse without considering the behavior of the other (for example Blau (1997), Gustman and Steinmeier (2001) among others). As a result of these concerns, a growing literature which explicitly models the joint labor force behavior of older couples has emerged. See for example, Blau and Rihahn (1999), Gustman and Steinmeier (2000) and An, Christensen, and Gupta (1999).

However, most of these studies are based on either reduced-form models or static structural models. As Rust, Buchinsky, and Benitez-Silva (2003) point out, these models suffer from two major shortcomings. First, reduced-form models cannot be used to predict behavioral responses to pol-

\textsuperscript{1} AFP is a Norwegian notation for Avtalefestet Pensjonsordning
icy changes. Second, static models cannot capture the important dynamic
elements of the retirement decision process. Theoretical life cycle model
framework suggests that the retirement decision making cannot be fully
explained by current income levels (see for example: Feldstein (1974)). Ret-
tirement is an intertemporal life cycle decision problem. Uncertainty and
re-optimization based on updated information period by period should be a
natural component of the retirement model. Dynamic models account better
for the sequential nature of retirement process in which the decision makers
adjust their behavior as events unfold. Two strands of dynamic retirement
models can be found in the literature, the dynamic programming model
suggested by Rust (1989) and the ‘option value model’ suggested by Stock
and Wise (1990). In general, the ‘option value model’ is less computa-
tionally demanding, but may ‘result(s) in a temporarily inconsistent decision
in which the worker ignores the fact as that new information arrives he
will be continually revising his estimate of the optimal departure date $t^*$’
(Rust (1994)). With the recent progress of simulation methods and com-
puter hardware, more and more empirically tractable dynamic programming
retirement models are specified and estimated, for example, French (2001),
However, the focus of those papers has been single agent models.

Empirical studies of joint retirement behavior in a structural dynamic
framework are rare. Christensen and Gupta (1994) introduce a dynamic pro-
gramming model of couples’ joint retirement decisions. Couples are assumed
to maximize a joint household utility function with respect to their retire-
ment decisions. Husbands are found to have at least as strong preferences
for leisure as wives.

In this paper, we model the working couple’s joint retirement behavior
as a discrete time, discrete choice dynamic programming problem. Each
member of the household is assumed to choose annually between two op-
tions, to continue working or to quit working (retirement, if eligible). Our
approach differs from Christensen and Gupta (1994) in three aspects: First,
instead of simply specifying a household utility function, we assume that
the household retirement behavior is an outcome of a cooperative bargain-
ing process between husband and wife. Second, detailed pension benefits
and tax rules are included in the model. This not only makes hypothetical
policy simulations possible, but also helps to accurately describe the budget
sets corresponding to different choice behavior. In addition, we also take into account the dissolution possibility of the decision unit due to either mortality or divorce.

The model is estimated on Norwegian register data. Our sample consists of working couples in which the husband qualified for AFP and wife are 50 years or older in 1997 or 1998. The retirement behavior of the couples is observed annually up to four years. Similar to Karlstrom, Palme, and Svensson (2003), we have excluded those couples where the wife is observed taking out disability insurance, mainly due to the difficulties of modeling the eligibility condition of the disability insurance scheme.

The estimation results demonstrate that the behavior is rather well explained by a model with only measures of economic incentives: wages and pension benefits. The parameters corresponding to the wife are found to be smaller than their counterparts for the husband in the joint utility function. It indicates that husbands have higher bargaining power within the household. In contrast to Christensen and Gupta (1994), we find that wives’ leisure is valued more than their husbands’ leisure. A possible explanation is that wives may be more efficient in household production than husbands.

The next section describes the Norwegian institutional settings. Section 3 introduces the data and looks at the dynamic pattern of the retirement behavior for our sample. Section 4 presents our dynamic programming model. The empirical setting and solution method are discussed in section 5. Section 6 reports the estimation results and performs a hypothetical policy simulation. Section 7 concludes.

2 Institutional Settings

In this section, we provide a short description of the Norwegian Pension System which is relevant to our study, namely the public old age pension system and the early retirement (AFP) system.

2.1 Public Old Age Pension

The backbone of the retirement system in Norway is a mandatory, defined benefit public pension system, covering all permanent residents, established in its current form in 1967. The standard retirement age is 67 under this scheme.
A crucial parameter in the system, used for defining contributions as well as benefits, is the basic amount which is referred to as $G$. This amount is adjusted by the Parliament once or more times each year, in accordance with changes in the general income level.

The benefits consist of two main components. One component is a minimum pension, paid to all persons who are permanently residing in the country. The minimum pension is the basic amount plus a special supplement pension $ST$, which is determined by the Parliament together with the basic pension $G$. The other component is an earnings based pension. The level depends on the number of pension earning years and the yearly pension points. A full earnings based pension requires as a general rule 40 years with income above $1G$. In the case of less than 40 pension-earning years, the pension is reduced proportionally. Pension points are computed for each calendar year. Each year, earnings exceeding the basic pension is divided by the basic pension to give pension 'points' for that year. Earnings above 12 times the basic pension do not give points, and earnings between 6 and 12 times the basic pension (8 and 12 times before 1992) are reduced to one third before calculating points. For earnings from year 1992 on, the maximum pension point is 7. The average yearly points over the 20 best years are calculated as the FPP (final pension point). A full earnings based pension is 42 percent of the amount which appears when the basic pension $G$ is multiplied by FPP. For years prior to 1992 the pension percentage is 45. However, many elderly people have had no possibilities of earning a full supplementary pension. In consequence, special transitional provisions have been introduced regarding people born before 1937. The detailed rules can be found in Haugen (2000).

In short, the old age pension is calculated using the following formula:

$$Y = bG + \max(G \cdot FPP \cdot \frac{(0.45T_1 + 0.42 \min(T_2, PY - T_1))}{PY}, ST). \quad (1)$$

where $T_1$ is the number of years with pension points greater than 0 before 1992, and $T_2$ is the number of years with pension point greater than 0 after 1992. $b = 1$ if spouse’s income is less than $2G$, $b = 0.75$ otherwise.

$$PY = \begin{cases} 
20 & \text{if born before 1918}, \\
20 + \text{birth year} - 1917 & \text{if born between 1918-1936}, \\
40 & \text{if born after 1936}.
\end{cases}$$
2.2 Early Retirement (AFP)

An early retirement scheme was introduced in 1989 as a result of the wage negotiations between trade unions and major employers in 1988. People covered by it have an opportunity to retire earlier than 67 with a pension as if they continued working with their ‘normal’ earnings up to age 67.

The scheme covers the whole public sector and part of the private sector. In order to be eligible an individual must be employed in a company covered by the scheme and meet certain individual requirements. Now the AFP scheme covers about 65-70% of the labor force.

From January 1 1989, the AFP age was 66. It was lowered to 65 from January 1 1990, to 64 from October 1 1994, to 63 on from October 1 1997 and to 62 from March 1 1998.

The pension level calculations under AFP scheme are aimed to provide the same pension benefits as if pensioner would continue to work until the ordinary retirement age instead of retiring early. The AFP pension is the sum of two parts. The first part is main component of AFP benefit, and is calculated using the same formula (1) as if AFP retiree had worked until age 67. This implies that the pension points in the years between the AFP eligibility age and 67 should be forecasted with some mechanism. The agreement is to use the maximum between the average of the last three earned points and FPP to substitute the unrealized points from the ‘future’ years. The second part is a supplement lump sum amount which is decided annually similar to the basic pension $G$.

3 Data

The analysis draws on data at the Frisch Centre, which are merged administrative registers. The original data have been received from Statistics Norway, and are held by the Frisch Centre with permission for research use. The data give an account of the main labor market activity for virtually the whole Norwegian adult population. Detailed income information is available from 1993, while the accumulated pension rights (annual pension points) are available back to 1967 from the pension register.

We concentrate our study on working couples, i.e. the couples where

\[2\text{The detailed requirements can be found in Haugen (2000)}\]
both husband and wife are active in the labor market prior to the initial time period. There are two reasons for this. On one hand, the share of dual earner households is increasing, mostly due to the fact that the labor force participation for elderly women has increased dramatically in most western countries. In Norway, the participation rate for women aged 55-66 rose from 40% in 1972 to 54% in 1997 (Dahl, Nilsen, and Vaage (2003)). On the other hand, our study treats the couple’s retirement decision as an optimization problem, which begins at a point in the middle of their life cycle. Restricting our sample to couples who are both working prior to the initial time period helps us to eliminate the variation of the initial conditions.

Our sample contains all working couples which satisfy three criteria: (i) the husband is qualified for AFP in 1997 or 1998. (ii) the wife is at least 50 years old at the year husband is qualified for AFP. (iii) none of members is known to take out disability pension during our observation period. When we impose all these criteria, we are left with a sample of 2081 households. We track the retirement behavior for all individuals in the data set on an annual basis for 4 years.

We are interested in the transition patterns into retirement by age. Fig-
Figure 2: Observed Retirement Hazard and Survival Rate by Age: Wives

Figure (1) and (2) give pictures of transition pattern in the form of retirement hazard and survival rate for the relevant age interval for husband and wife separately. Limited information can be seen from the pattern for the husbands, mainly due to the limited age interval. However, the age retirement hazard for wives is double peaked at 62 and 67. These peaks are corresponding to the early and normal retirement ages respectively.

4 A Dynamic Programming Model of Joint retirement

In this section, we develop a dynamic programming model for couples’ joint retirement decisions.

Let $t$ be the discrete time index, with $t = 1$ as the initial time period in our analysis, and $t = T$ is upper bound of the planning horizon, which is defined in section 4.2.
4.1 Decision variables

In each period, the couples choose whether or not to stay in the labor force. Let \( d^m_t = 1 \) if the husband chooses to take retirement; and \( d^m_t = 0 \) otherwise. Similarly, we let \( d^f_t = 1 \) if the wife chooses not to work. The decision variables will be \( d^t = \{d^m_t, d^f_t\} \).

An individual who takes out retirement (both early and old age) is not allowed to return to labor force. For the wife who is not qualified to either old age or early retirement pension, she can still choose not to work. In this situation, she will enjoy the same leisure as retirement but with pension income as 0 until she reaches the old age pension eligibility age 67. Moreover, we assume that she is not allowed to come back to labor force once she decides to stop working. This assumption may seem to be restrictive. But when we look at the data, we see that none of wives who quit their jobs without eligibility to any pension benefits returns to the labor force later.

Under this assumption, the joint retirement decision problem is essentially a two-dimensional optimal stopping problem.

4.2 Mortality Risk, Divorce and Planning Horizon

Our decision unit, the two-member household, may dissolve due to either the death of any member or a divorce/separation.

We do not treat divorce as a choice, but rather as a result of an exogenous stochastic process just like mortality. The divorce probability is assumed to be the same as the national average divorce rate of wife’s age group\(^3\). The mortality rate is also considered to be exogenous and equal across sample given gender and age. A better treatment will be to link the health condition to the mortality rate, as Rust and Phelan (1997) did. However, in our study, we have only incomplete data on the health status, which makes the estimation of a health related mortality risk system infeasible.

The probability that the household does not dissolve at period \( t \) can then be calculated as \( \pi_t = (1 - M^f_t)(1 - M^m_t)(1 - \delta_t) \) where \( M^k_t \) (\( k = m, f \)) is the gender age specific mortality risk, and \( \delta_t \) is the divorce rate. In fact, \( 1 - \pi_t \) can be seen as the hazard of household dissolution, and it is a function of both husband’s and wife’s age. Figure (3) illustrates the

\(^3\) In year 1997, the divorce rate for women of age 55-59 is 0.4%, and 0.25% for those of age 60-64. (Source, Population statistics. Divorces and separations, SSB (2002))
household dissolution hazard and survival rate for the household where both members are of age 62 at the initial time period.

Since it is difficult to model the behavior of the household after it dissolves, we assign a constant terminal value to these cases as Blau and Gilleskie (2001) did. We denote it as $V_a$. In our model the choice of labor market does not influence the mortality rate and probability of a divorce, so the magnitude of the terminal value does not play any role in our analysis. However, we cannot drop the mortality risk and divorce probability in our model, because they do have an effect on the choice probability in combination with the discount factor $\rho$. The detailed proof of this point is given in section 5.3.

The individuals are assumed to die with probability one at age 90. It implies that the planning horizon $T$ will be defined as $T = 91 - \max(\text{age}_m^1, \text{age}_f^1)$, where $\text{age}_m^1$ and $\text{age}_f^1$ is the age for husband and wife at the initial period $t = 1$. 

Figure 3: household dissolution hazard and survival rate
4.3 The Preferences

The preferences for husband and wife at time \( t \) are specified as \( U^t_m = U_m(s^t, d^t; \theta) \) and \( U^t_f = U_f(s^t, d^t; \theta) \) respectively, where \( s^t \) is a vector of household state variables at period \( t \), \( d^t \) is the couple’s decision variables defined above, and \( \theta \) is a vector of preference parameters to be estimated, but known to the individual.

Denote the discount factor as \( \rho \), we define \( LU^t_m \) and \( LU^t_f \) as the discounted ‘remaining life time’ utilities of husband and wife respectively:

\[
LU^t_m = U^t_m + E(\sum_{t=1}^{T} \rho^{t-t}U^t_m), \\
LU^t_f = U^t_f + E(\sum_{t=1}^{T} \rho^{t-t}U^t_f).
\] (2)

Similar to Mastrogicacomo, Alessie, and Lindeboom (2002) and Maestas (2001), we model the couple’s retirement decision as an outcome of a cooperative bargaining process. At each period \( t \in \{1, 2, \cdots, T - 1\} \), the couple is assumed to maximize a collective household utility function \( LU^t \) with respect to the decision variable \( d^t \). The collective utility function is a weighted sum of the ‘remaining life time’ utilities of the husband and the wife:

\[
LU^t = \lambda LU^t_m + (1 - \lambda) LU^t_f. \tag{3}
\]

Following Maestas (2001), Browning and Chiappori (1998), \( \lambda \) measures the husband’s decision-making control in the household. If \( \lambda = 1 \), the household behaves as if the husband has exclusive decision-making control, whereas \( \lambda = 0 \) implies that the wife has exclusive control. In the present study, \( \lambda \) is assumed to be constant over time for the same couple, and is not affected by the retirement decision.

Let \( U^t \) be the weighted sum of the single period utility function of the husband and the wife with weight as \( \lambda \) and \( 1 - \lambda \):

\[
U^t(s^t, d^t; \theta, \lambda) = \lambda U^t_m(s^t, d^t; \theta) + (1 - \lambda) U^t_f(s^t, d^t; \theta). \tag{4}
\]

The optimization facing the couples can then be written as

\[
\max_{d^t}(\lambda LU^t_m + (1 - \lambda) LU^t_f) \tag{5}
\]

\[
= \max_{d^t}(U^t + E(\sum_{s=t+1}^{T} \rho^{s-t}U^t)). \tag{6}
\]
It shows that if we assume that the members’ expectations on future events are the same, we could treat the couple as the decision unit with a single period joint utility function $U^t$. This enables us to use the single agent dynamic programming framework developed by Rust (1989).

4.4 Choice Probabilities and Likelihood Function

As econometricians, we are not able to observe the full set of state variables. Following Rust (1989), the state variables are partitioned into two components, $s^t = (x^t, \varepsilon^t)$. At time $t$, $x^t$ can be observed by both the econometrician and the decision maker, while $\varepsilon^t$ is only observed by the decision maker. This assumption leads to a random utility framework. The single period joint utility function for the couple can then be written $U^t(s^t, d^t; \theta, \lambda) = u^t(x^t, d^t; \theta, \lambda) + \varepsilon^t(d^t)$.

Let $p(x^{t+1}|x^t, d^t; \theta_p)$ represent the couple’s subjective belief of future events, which is a probability distribution of the state variables at $t+1$ given their observed values and current decision at period $t$. $\theta_p$ is a vector of parameters related to the subjective belief, which are known to the couples. For a state variable which follows a deterministic dynamic process, such as the husband’s age, the corresponding distribution degenerates to a mass point.

Although from the decision maker’s point of view the optimal decision rule is deterministic, it is random for econometricians. We can at most only obtain a conditional choice probability $Pr_t(d^t|x^t; \theta, \lambda, \rho)$ as in the static discrete choice framework.

Assume that error term $\varepsilon^t$ is i.i.d. extreme value distributed across choices and periods, Rust (1994) shows that, for $t < T$:

$$Pr_t(d^t|x^t; \theta, \lambda, \theta_p, \rho) = \frac{\exp(v_t(d^t, x^t; \theta, \lambda, \theta_p, \rho))}{\sum_{d^t \in D(x^t)} \exp(v_t(d^t, x^t; \theta, \lambda, \theta_p, \rho))}. \quad (7)$$

where $v_t(d^t, x^t; \theta, \lambda, \theta_p, \rho)$ is the expected valuation function, which is
defined as
\[ v_t(d', x^t; \theta, \lambda, \theta_p, \rho) = u_t(d', x^t; \theta, \lambda) \]
\[ + \rho \pi_t \int \log [ \sum_{d' \in D(x^t)} \exp(v_{t+1}(d', x^{t+1}; \theta, \lambda, \theta_p, \rho))dp(x^{t+1}|x^t, d^t; \theta_p)] \]
\[ + \rho(1 - \pi_t)V_a. \] (8)

Here \( D(x^t) \) is the available choice set at \( t+1 \) given current states variable \( x_t \).

To sum up, the couple’s behavior depends on the discount factor \( \rho \), the joint preference \( U^t \) and the subjective transition probability \( p(x^{t+1}|x^t, d^t; \theta_p) \).

If we observe panel data \( \{x^t_i, d^t_i\} \) for \( i = 1, 2, \ldots N, t = 1, 2, \ldots, T_i \), we can estimate the model using a two-stage procedure which is often followed in the literature, see for example Rust and Phelan (1997) and Karlstrom, Palme, and Svensson (2003). In the first step, the transition probability \( p(x^{t+1}|x^t, d^t; \theta_p) \) is estimated using available data. Then we solve (8) by backward induction from terminal period \( T \), calculate the choice probability (7) accordingly, and construct the likelihood function as follows:
\[ L(\theta, \lambda, \rho) = \prod_{t=1}^{T_i} P_t(d^t_i|x^t_i; \theta, \lambda, \rho). \] (9)

5 Empirical Specifications

In the current analysis, we make the assumption that there is no heterogeneity across the households. In other words, we assume that the parameters in the individual utility functions, the bargaining parameter \( \lambda \) and the discount parameter \( \rho \) are constant across different households.

5.1 State Variables

There are totally eleven state variables for each couple. Among them, five for husband \( X^t_m = (age^t_m, wage^t_m, benefit^t_m, Z^t_m, L^t_m) \), and six state variables for wife \( X^t_f = (age^t_f, wage^t_f, benefit^t_f, Z^t_f, L^t_f, e^t) \).

\( age^t_k \) denotes the age of member \( k \) at time \( t \). \( wage^t_k \) is the wage earnings if member \( k \) is working at time \( t \). \( benefit^t_k \) is the (potential) pension income.

Similar to Christensen and Gupta (1994), we use a variable \( Z^t_k \) denotes
the retirement date, which is defined as the following:

$$Z^t_k = \begin{cases} \text{date of retirement} & \text{if member } k \text{ retires at } t \text{ or earlier, and} \\
w & \text{if member } k \text{ continues to work at } t. \end{cases}$$

$Z^t_k$ takes value from the set $\{1, 2, \cdots, t, w\}$. This state variable summarizes the whole decision sequence up to $t$. For example, $Z^3_m = 2$ means that the husband retires at period 2, and corresponds to a decision sequence $\{d^1_m = 0, d^2_m = d^3_m = 0\}$. Instead $Z^3_m = w$ means that the husband is still working at $t = 3$, and corresponds to a decision sequence $\{d^1_m = d^2_m = d^3_m = 0\}$.

$L^t_k$ is the leisure enjoyed by the member $k$, $L^t_k = 1$ if and only if $Z^t_k \leq t$.

$e^t$ is the retirement eligibility indicator for wife, $e_t = 1$ if she is eligible for AFP or ordinary retirement at $t$, and $e_t = 0$ otherwise.

Health is considered to be a very important factor for retirement decision. Unfortunately, it is a variable which is difficult to measure. Unlike the HRS data for United States, there is no variable representing the individual health information. The closest thing we can get is the sick leave days reported to the authority. Naturally, these data are not available to those who are retired. So we are not able to observe the dynamic of the health status, neither can we make any estimation of it. Thus, the health condition is not included in our analysis.

5.1.1 The Dynamics of the State Variables

Some of the state variables, in their nature, are deterministic and follow a predetermined path, such as age at period $t$, $\text{age}^t_k$.

For the labor market status, we assume that the individuals have perfect control. The decision variable sequence $\{d^t_k, \tau = 1, \cdots, t\}$ completely defines the retirement dates for the couple $Z^t_k$. Once $Z^t_k$ is determined, the leisure enjoyed by member $k$, $L^t_k$ follows immediately by definition.

For the early retirement, old age pension rules, we assume that the individuals correctly anticipate the changes of the social security rules over the whole period 1997-2001, but maintain the static expectation that no further changes will occur thereafter. This assumption is similar to the ‘semi-rational’ expectation assumption in Rust (1989). The same assumption is

---

4 Note that these two variables $Z^t_m$ and $Z^t_f$ are well defined only when the retirement states are absorbing.
made for the tax rules.

The social security eligibility variable for wife, \( e_t \), can not be treated as exogenous for the reason that AFP eligibility is also conditional on past labor market behavior. However, once the past decisions are given, it can be calculated using the social security rules with certainty.

As we see in section 2, retirement benefit \( benefit_t^k \) is closely related to the past wage earnings. There will be no uncertainty in expected early retirement and old age pension incomes once the wage income history is given. These incomes can be imputed in detail by the existing rules.

Thus, accurate modeling of wage dynamic is crucial for us to obtain successful estimates for the dynamic programming model. Similar to Rust, Buchinsky, and Benitez-Silva (2003), Knaus (2002) and Karlstrom, Palme, and Svensson (2003), we specify a ‘mis-specified’ log normal regression of the individuals’ annual wage income as the following:

\[
\log(wage_t+1) = \omega_1 + \omega_2 \log(wage_t) + \omega_3 age_t + \omega_4 (age_t)^2 + \xi_t, \tag{10}
\]

where \( age_t \) denotes age, and \( \xi_t \) are \( i.i.d \) normal distributed with mean 0 and variance \( \sigma^2_\xi \). The quadratic specification allows for an age income profile.

The regression is done separately for husbands and wives. Table (1) shows the results from the estimation of this model.

While these regressions need not correspond to the true process governing the wage dynamic, the estimated regressions for both husband and wife have quite high \( R^2 \), and the estimated variances of the error terms are very small, as seen in table (1). This indicates the low variability of the

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
<th>Husbands</th>
<th>Wives</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
<td>( \omega_1 )</td>
<td>3.516</td>
<td>5.03</td>
</tr>
<tr>
<td>( \log(wage_t) )</td>
<td>( \omega_2 )</td>
<td>0.951</td>
<td>0.005</td>
</tr>
<tr>
<td>age</td>
<td>( \omega_3 )</td>
<td>-0.088</td>
<td>0.15</td>
</tr>
<tr>
<td>( age^2 )</td>
<td>( \omega_4 )</td>
<td>6.8e-4</td>
<td>1.2e-3</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>0.93</td>
<td>0.85</td>
<td></td>
</tr>
<tr>
<td>( \sigma^2_\xi )</td>
<td>9.2e-4</td>
<td>9.8e-3</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Estimation results for wage regression equations, husbands and wives
\{\log(wage^t)\} sequences. Thus, we are able to use a deterministic Markovian updating formula with quite high precision to impute the wage income at period \(t + 1\) based on the wage income of period \(t\):

\[
\text{wage}^{t+1} = \exp(\hat{\omega}_1 + \hat{\omega}_2 \log(\text{wage}^t) + \hat{\omega}_3 \text{age}^t + \hat{\omega}_4 (\text{age}^t)^2 + \hat{\sigma}^2_{\epsilon}/2).
\] (11)

This finding is highly encouraging. First, a deterministic wage dynamic will allow us to carry the whole history of wages and give a precise pension calculation without incurring the problem of ‘curse of dimensionality’. More important, a deterministic wage dynamic essentially makes the whole subjective belief system of the state variables deterministic, which greatly reduces the numerical complexity of the model implementation.

5.2 Utility Function

Disposable income for spouse \(k\), \(y_k^t\) is defined as

\[ y_k^t = \text{wage}^t_k \cdot (1 - L^t_k) + L^t_k \cdot \text{benefit}^t_k - \text{Tax}^t_k, \]

where \(\text{wage}^t_k\) is the annual wage income, \(\text{benefit}^t_k\) is the annual pension income respectively for spouse \(k\) at time \(t\). \(\text{Tax}^t_k\) is the tax paid by spouse \(k\) at time \(t\).\(^5\)

The utility functions are specified as follows:

\[
\begin{align*}
U^t_m &= u^t_m + \varepsilon^t_m = \alpha_{m1} \log(y^t_m) + \alpha_{m2} \log(y^t_f) + \beta_{m1} L^t_m + \beta_{m2} L^t_f + \varepsilon^t_m, \\
U^t_f &= u^t_f + \varepsilon^t_f = \alpha_{f1} \log(y^t_f) + \alpha_{f2} \log(y^t_m) + \beta_{f1} L^t_f + \beta_{f2} L^t_m + \varepsilon^t_f.
\end{align*}
\] (12)

According to this specification, \(\alpha_{m1}, \alpha_{f1}\) measure the contribution of economic incentives to the utility. Income or consumption sharing is captured by the effect that \(y^t = y^t_m + y^t_f\). The \(\beta\) parameters measure the utility of leisure. By letting the spouse’s leisure to enter his/her utility function, we allow for the individual’s preference to be ‘altruistic’ and the existence of caring within the household.

We expect that all these parameters to be non-negative, for the utility\(^6\)

\(^5\)For wives who are not eligible to any pension scheme, we intend to set the benefit as 0. However, due to numerical concerns, a very small number is given instead.

\(^6\)Detailed Norwegian tax rules are used when calculating the tax. The unit of tax calculation is the couple, not the individual, which means that the taxes paid by the couple depend on the labor market status of both members of the household.
function should generally be a increasing function of both income and leisure.

However, given the setting of the model, similar to Mastrogiacomo, Alessie, and Lindeboom (2002), we are not able to identify the bargaining parameter separately from the other parameters, since we can only estimate the parameters in the joint weight utility \( U_t = \lambda U_t^m + (1 - \lambda)U_t^f \). In fact, using (12), we see that

\[
U_t = \lambda \alpha_{m1} \log(y_m^t) + (1 - \lambda)\alpha_{f1} \log(y_f^t) + (\lambda \alpha_{m2} + (1 - \lambda)\alpha_{f2}) \log(y^t) + (\lambda \beta_{m1} + (1 - \lambda)\beta_{f2})L_m^t + ((1 - \lambda)\beta_{f1} + \lambda \beta_{m2})L_f^t + (\lambda \varepsilon_m^t + (1 - \lambda)\varepsilon_f^t)
\]

\[
= \kappa_m \log(y_m^t) + \kappa_f \log(y_f^t) + \kappa \log(y^t) + \mu_m L_m^t + \mu_f L_f^t + \varepsilon_t^*
\]

\[
= u^t(y_m^t, L_m^t, y_f^t, L_f^t) + \varepsilon_t^*.
\]

We can only identify \( \kappa_m, \kappa_f, \kappa, \mu_m, \mu_f \). Similar to the individual utility case, we expect that all these parameters to be non-negative. The parameters for leisure might be related to the age of the member. It has been found in several studies that it is appropriate to use an age-leisure preference profile that is rather fast increasing in certain age interval (Karlstrom, Palme, and Svensson (2003) and Heyma (2001), among others). So we let

\[
\mu_k = \exp(\mu_{k,1}) + \exp(\mu_{k,2}) \frac{\exp(\frac{\text{age}^t - \mu_{k,3}}{\mu_{k,4}})}{1 + \exp(\frac{\text{age}^t - \mu_{k,3}}{\mu_{k,4}})}.
\]

The joint error term \( \varepsilon_t^* = \lambda \varepsilon_m^t + (1 - \lambda)\varepsilon_f^t \) is assumed to be i.i.d. extreme value distributed.

Note that from the above section 5.1.1, we see that for any given couple, both leisure variables \( L_k^t \) and disposable incomes \( (y_m^t, y_f^t) \) depend only on possible retirement dates \( (Z_m^t, Z_f^t) \). So the deterministic part of the single period joint utility function \( u^t(y_m^t, L_m^t, y_f^t, L_f^t) \) in (13) can be written as \( u^t(Z_m^t, Z_f^t) \).

---

7 An elaborately designed non-linear relationship between the parameters and couple specific variables such as age, wealth etc might ensure the identification of the bargaining power parameter \( \lambda \). However, this identification will fully hinge on the functional form of these relationships, which might be wrong. This is also the reason why we specify the age varying profile of leisure preference in the joint utility function (13) but not in the individual utility function (12).
5.3 Solving the Dynamic Programming Model

5.3.1 Expected Valuation and Choice Probability

As we discussed in last section, we can denote the single period joint utility as \( u_t(Z_t^m, Z_t^f) \). Using (8), we can write the expect value function \( v_t(d^t, x^t; \theta, \lambda, \rho) \) as \( v_t(Z_t^m, Z_t^f) \).

The expected valuation function \( v_t(Z_t^m, Z_t^f) \) can be calculated using the following formula.

Note that \( Z_t^k \) can take value from the set \( \{1,2,\cdots,t,w\} \), then we have for \( h,j \in \{1,2,\cdots,t\} \),

\[
v_t(h, j) = u_t(h, j) + \rho \pi_{t+1} v_{t+1}(h, j) + \rho(1 - \pi_{t+1}) V_a , \tag{15}
\]

\[
v_t(h, w) = u_t(h, w) + \rho(1 - \pi_{t+1}) V_a + \rho \pi_{t+1} \ln[\exp(v_{t+1}(h, w)) + \exp(v_{t+1}(h, t + 1))], \tag{16}
\]

\[
v_t(w, j) = u_t(w, j) + \rho(1 - \pi_{t+1}) V_a + \rho \pi_{t+1} \ln[\exp(v_{t+1}(w, j)) + \exp(v_{t+1}(t + 1, j))], \tag{17}
\]

\[
v_t(w, w) = u_t(w, w) + \rho(1 - \pi_{t+1}) V_a + \rho \pi_{t+1} \ln[\sum_{(h',j') \in \{w,t\} \times \{w,t\}} \exp(v_{t+1}(h',j'))]. \tag{18}
\]

Moreover, the available choice set at current period also depends solely on the retirement status at last period \( (Z_{m}^{t-1}, Z_{f}^{t-1}) \). It can be written as

\[
D(x^{t-1}) = \begin{cases} 
\{w, t\} \times \{w, t\} & \text{if } Z_{m}^{t-1} = Z_{f}^{t-1} = w, \\
\{w, t\} \times \{Z_{f}^{t-1}\} & \text{if } Z_{m}^{t-1} = w, Z_{f}^{t-1} \leq t - 1, \\
\{Z_{m}^{t-1}\} \times \{w, t\} & \text{if } Z_{m}^{t-1} \leq t - 1, Z_{f}^{t-1} = w, \\
\{(Z_{m}^{t-1}, Z_{f}^{t-1})\} & \text{if } Z_{m}^{t-1} \leq t - 1, Z_{f}^{t-1} \leq t - 1.
\end{cases}
\]

Under the assumption that \( \varepsilon^{t}_i \) is i.i.d. extreme value distributed, we can rewrite the choice probability (7) at time \( t \) in detail as follows:

If \( Z_{m}^{t-1} \leq t - 1, Z_{f}^{t-1} \leq t - 1 \), both members have retired when entering time \( t \), then we have

\[
\Pr_t(Z_{m}^{t-1}, Z_{f}^{t-1}) = 1. \tag{19}
\]

If \( Z_{m}^{t-1} \leq t - 1, Z_{f}^{t-1} = w \), the husband has already retired, but the wife
is still working when entering time $t$, then

$$
\Pr_t(Z_{m}^{t-1}, w) = \frac{\exp(v_t(Z_{m}^{t-1}, w))}{\exp(v_t(Z_{m}^{t-1}, w)) + \exp(v_t(Z_{m}^{t-1}, t))},
$$

$$
\Pr_t(Z_{m}^{t-1}, t) = \frac{\exp(v_t(Z_{m}^{t-1}, t))}{\exp(v_t(Z_{m}^{t-1}, w)) + \exp(v_t(Z_{m}^{t-1}, t))}.
$$

If $Z_{m}^{t-1} = w, Z_{f}^{t-1} \leq t - 1$, the husband is working, but the wife has already retired when entering time $t$, we have

$$
\Pr_t(t, Z_{f}^{t-1}) = \frac{\exp(v_t(t, Z_{f}^{t-1}))}{\exp(v_t(w, Z_{f}^{t-1})) + \exp(v_t(t, Z_{f}^{t-1}))},
$$

$$
\Pr_t(w, Z_{f}^{t-1}) = \frac{\exp(v_t(w, Z_{f}^{t-1}))}{\exp(v_t(w, Z_{f}^{t-1})) + \exp(v_t(t, Z_{f}^{t-1}))}.
$$

If $Z_{m}^{t-1} = w, Z_{f}^{t-1} = w$, both husband and wife are working when entering time $t$, we have for all $(h, j) \in \{w, t\} \times \{w, t\}$

$$
\Pr_t(h, j) = \frac{\exp(v_t(h, j))}{\sum_{(h', j') \in \{w, t\} \times \{w, t\}} \exp(v_t(h', j'))}.
$$

5.3.2 The Solution Method

The model is solved using backward induction.

Given the fact that age 70 is the mandatory retirement age in Norway, no choice on labor market will be available after the younger member of the household turns into 70. Denote this period as $t = T^*$. The solution of the dynamic programming problem after $T^*$ is trivial, since there will be no choice available for the couples.

We start our backward induction at $T^*$. For any possible retirement dates for the couple $(h, j)$, the expected valuation function at $T^*$ is then:

$$
v_{T^*}(h, j) = \sum_{t=T^*}^{T} \rho^{t-t} \prod_{k=1}^{t-t} \pi_{t+k} \cdot u^{T^*}(h, j) + AV_a,
$$

where $A$ is a constant which depends on discount factor $\rho$ and the period specific household dissolution probabilities $1 - \pi_t$. $V_a$ is the terminal value when household dissolves.

It follows immediately from (15) to (23) that the expected value functions can be calculated as four one-dimensional backward recursions. The
Another insight we can gain from (15) to (23) is that the terminal value $V_a$ will cancel out when we calculate the choice probability so we could simply normalize it to zero. Both the mortality and the divorce probability enter into the model in conjunction with the discounting factor $\rho$.

6 Estimation results and Policy Simulation

6.1 A Technical Note on Estimation

The numerical optimization of the logarithm of the likelihood function (9) is not straight forward. To cope with this complex problem, a group of gradient and non-gradient optimization methods are used. First, a genetic algorithm procedure is applied to obtain the initial starting values. Then we use the iterated downhill simplex method. After the simplex method converges, a gradient method is implemented by using the estimates from the simplex method as starting values.

6.2 Estimation Results

We estimate all parameters in (13) and (14), namely the parameters related to incomes $\kappa_m, \kappa_f, \kappa$, and the age-leisure profile parameter $\mu_{k,1}, \mu_{k,2}, \mu_{k,3}, \mu_{k,4}$.
The first column of table (2) shows the estimation results of all 11 parameters. All estimates have the expected sign. However, a very interesting point to note is that, for both husband and wife, the resulting function of the preference for leisure (14) is a constant for all relevant age interval. The estimates suggest that the changes in preferences by age are small. Thus a model without allowing for preferences to change over time is sufficient for our data. The second column in table (2) shows the estimates of such a model, and the standard errors of these estimates are reported in column 3.

From the estimates, we see that both parameters corresponding to the wife have much smaller magnitude when compared with their counterparts for the husband. Namely $\kappa_f < \kappa_m$ and $\mu_{f,1} < \mu_{m,1}$. It could be a sign which indicates that the wives have lower bargaining power than their husbands, which is not surprising for the cohorts studied in our analysis. Interestingly,

---

8It is neither possible nor necessary to report the standard error for estimates in column 1.
in contrast to Christensen and Gupta (1994), our estimates suggest that wives’ leisure is valued much higher in the household than their husbands’ leisure. We think this is consistent to the view that wives are more efficient in household production than husbands. On the other hand, we notice that the parameter for the joint income $\kappa$ is higher than the sum of $\kappa_f$ and $\kappa_m$, which shows that the joint income is valued higher than individual income for both husband and wife.

The fitting of the model is satisfactory. The McFadden $R^2$ is 37.3 per cent, which is fairly high in non-transportation literature. Figure (4) and (5) show the observed and predicted cumulative probability of remaining in labor force (the survival rate) by age for husbands and wives respectively. Table (3) lists the retirement hazard rates by age for both husbands and wives.

As we can see from figure (4) and (5) and table (3), our model replicates the retirement pattern fairly well. The fitting for wives seem to be better than that for husbands. Our suspicion is that it might have something to do with the fact that we have rather limited age interval for husbands in our
Table 3: Retirement Hazards by Age

<table>
<thead>
<tr>
<th>age</th>
<th>wife observed</th>
<th>predicted</th>
<th>simulation</th>
<th>husband observed</th>
<th>predicted</th>
<th>simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>52</td>
<td>0.008</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>0.010</td>
<td>0.003</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>0.007</td>
<td>0.002</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>0.004</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>0.005</td>
<td>0.001</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>0.003</td>
<td>0.002</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>0.003</td>
<td>0.004</td>
<td>0.003</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>0.004</td>
<td>0.013</td>
<td>0.009</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>0.019</td>
<td>0.050</td>
<td>0.038</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>0.264</td>
<td>0.215</td>
<td>0.193</td>
<td>0.423</td>
<td>0.352</td>
<td>0.320</td>
</tr>
<tr>
<td>63</td>
<td>0.193</td>
<td>0.181</td>
<td>0.166</td>
<td>0.332</td>
<td>0.279</td>
<td>0.259</td>
</tr>
<tr>
<td>64</td>
<td>0.107</td>
<td>0.148</td>
<td>0.139</td>
<td>0.125</td>
<td>0.170</td>
<td>0.166</td>
</tr>
<tr>
<td>65</td>
<td>0.099</td>
<td>0.128</td>
<td>0.122</td>
<td>0.069</td>
<td>0.105</td>
<td>0.108</td>
</tr>
<tr>
<td>66</td>
<td>0.104</td>
<td>0.142</td>
<td>0.136</td>
<td>0.004</td>
<td>0.093</td>
<td>0.103</td>
</tr>
<tr>
<td>67</td>
<td>0.360</td>
<td>0.185</td>
<td>0.183</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>0.220</td>
<td>0.200</td>
<td>0.204</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

analysis. An important criterion for evaluation of retirement models is the ability of replicating the hazard spikes. In figure (2), we see there are two retirement peaks for wives, which are at age 62 and 67 respectively. From table (3) we see that our model successfully predicts that first retirement peak for wives at age 62 but misses the second peak at 67. The missing of the second peak might be due to the fact that there are quite few observations in our sample with wives of age 67 or older so that the maximum likelihood estimator attaches little weight to accurately predict the retirement behavior of this age group.

6.3 Policy Simulation

In order to illustrate the magnitude of the estimated relationship and the corresponding impact of potential policy changes, we have simulated the effect of a hypothetical policy reform based on the estimated model. The tax system in Norway strongly favors retirement (Hernæs, Sollie, and Strøm (2000)). Figure (6) shows the amount of tax to pay for a married individual.

---

9 This point is taken from Karlstrom, Palme, and Svensson (2003).
with a working spouse in 1998 from different income sources. For yearly earnings around 160,000 NOK, the tax for AFP pension income is around 20 per cent (8,000 NOK) lower than that for labor income. In the policy simulation, we make the tax system less generous to the pension benefits — pension benefits will be taxed the same way as labor earnings.

The policy simulation results are shown in figure (4), (5) and table (3). As we expected, this hypothetical policy clearly reduces the retirement hazard before age 65 for both husbands and wives. The model predicts a reduction of the hazard rate at the AFP eligibility ages (62 and 63 in our sample) by around 0.02. Similar results have been found in Hernæs, Jia, and Strøm (2001). However, the magnitude of the reductions we find here is smaller than that of Hernæs, Jia, and Strøm (2001) (typically 0.05-0.07). Of course, this might be simply due to cohorts effect since these two studies are based on different data. We also notice that the relative reduction of the retirement hazard \((H_s - H_p)/H_p\), where \(H_s\) is the hazard rate after simulation and \(H_p\) is the predicted hazard rate, generally decreases with age. This is consistent with the fact that older individuals normally have less to lose due to the restriction of compulsory retirement age.
In short, this hypothetical policy reform increases the labor participation rate for both husband and wife by around 4 percentage points at the age of 65. Although the effects are not as strong as reported in previous literature, this policy is still a good candidate for the purpose of counteracting the negative effects on labor supply implied by the early retirement programs.

7 Conclusion

In this paper, we develop a dynamic programming model for joint household retirement behavior that acknowledges the institutional features of the Norwegian Social Security and Tax system. The model is then estimated on a sample of Norwegian working couples in which the husband qualified for a special early retirement scheme in 1997 or 1998. The model provides fairly well within sample predictions of labor force participation rates.

Interestingly, for both husband and wife, the estimated function of the preference for leisure is a constant for all relevant age interval. It suggests that the changes in preferences by age are rather little.

We find that the parameters corresponding to the wife have smaller numerical magnitude when compared with their counterparts for the husband in the joint utility function. This can be a sign of unbalanced bargaining power within the household. In contrast to Christensen and Gupta (1994), we find that wives’ leisure is valued more than their husbands’ leisure, which is consistent with the view that wives are more efficient in household production than husbands.

A hypothetical policy simulation is performed using the estimates of the model. In the simulation, pension benefits are taxed the same way as labor earnings. It increases the labor supply for both husband and wife by around 4 percentage points at the age of 65.

As a first step to develop a fully dynamic model for analyzing the joint retirement behavior of Norwegian couples, the present study inevitably has some limitations. The most important limitation is that we do not model the savings behavior of the household. A more realistic dynamic programming model should include also wealth accumulation and allow for the consumption/savings decision as the single agent model in French (2001). Another limitation is that we are not able to model disability pension scheme in our analysis. So in the current analysis, an important path way out of the labor
force is excluded. Further research is required to resolve these problems.

References


Recent publications in the series Discussion Papers

313  G.H. Bjertnær (2001): Optimal Combinations of Income Tax and Subsidies for Education


316  A. Bruvoll and K. Nyborg (2002): On the value of households' recycling efforts

317  E. Bjørn and T. Skjerpen (2002): Aggregation and Aggregation Biases in Production Functions: A Panel Data Analysis of Translog Models


320  T. J. Klette and A. Raknerud (2002): How and why do Firms differ?


323  E. Røed Larsen (2002): Searching for Basic Consumption Patterns: Is the Engel Elasticity of Housing Unity?


325  E. Røed Larsen (2002): Consumption Inequality in Norway in the 80s and 90s.


327  M. Søberg (2002): A laboratory stress-test of bid, double and offer auctions


332  M. Greker (2002): Eco-labels, Production Related Externalities and Trade

333  J. T. Lind (2002): Small continuous surveys and the Kalman filter

334  B. Halvorsen and T. Willumsen (2002): Willingness to Pay for Dental Fear Treatment. Is Supplying Fear Treatment Social Beneficial?

335  T. O. Thoresen (2002): Reduced Tax Progressivity in Norway in the Nineties. The Effect from Tax Changes


340  H. C. Bjørnland and H. Hungnes (2003): The importance of interest rates for forecasting the exchange rate


343  B. Bye, B. Strøm and T. Åvitsland (2003): Welfare effects of VAT reforms: A general equilibrium analysis


346  B.M. Larsen and R.Neshakken (2003): How to quantify household electricity end-use consumption

347  B. Halvorsen, B. M. Larsen and R. Neshakken (2003): Possibility for hedging from price increases in residential energy demand


349  B. Holtsmark (2003): The Kyoto Protocol without USA and Australia - with the Russian Federation as a strategic permit seller

350  J. Larsson (2003): Testing the Multiproduct Hypothesis on Norwegian Aluminium Industry Plants


352  E. Holmøy (2003): Aggregate Industry Behaviour in a Monopolistic Competition Model with Heterogeneous Firms


354  E. Holmøy (2003): Aggregate Industry Behaviour in a Monopolistic Competition Model with Heterogeneous Firms


