MEMORANDUM

No 04/2000

Family Labor Supply when the Husband is Eligible for Early Retirement: Some Empirical Evidences

By

Jia Zhiyang

ISSN: 0801-1117

Department of Economics
University of Oslo

This series is published by the
University of Oslo
Department of Economics

P. O.Box 1095 Blindern
N-0317 OSLO Norway
Telephone: + 47 22855127
Fax: + 47 22855035
Internet: http://www.sv.uio.no/sosoek/
e-mail: econdep@econ.uio.no

In co-operation with
The Frisch Centre for Economic Research

Gaustadalleén 21
N-0371 OSLO Norway
Telephone: +47 22 95 88 20
Fax: +47 22 95 88 25
Internet: http://www.frisch.uio.no/
e-mail: frisch@frisch.uio.no

List of the last 10 Memoranda:

No 03	By Mads Greaker: Strategic Environemental Policy; Eco-dumping or a green strategy? 33 p
No 02	By Ole J. Røgeberg: Married Men and Early Retirement Under the AFP Scheme. 78 p.
No 01	By Steinar Holden: Monetary regime and the co-ordination of wage setting. 35 p.
No 38	By Knut Røed: Relative Unemployment Rates and Skill-Biased Technological Change. 24 p.
No 37	By Finn R. Førsund: Modelling Transboundary air Pollution: The RAINS model approach. 33 p.
No 36	By Christian Brinch: Statistical Discrimination and the Returns to Human Capital and Credentials. 33 p.
No 35	By Halvor Mehlum, Karl Ove Moene and Ragnar Torvik: Crime Induced Poverty Traps. 19 p.
No 34	By Ove Wolfgang: Reflections onAbatement Modelling. 34 p.
No 33	By Jon Strand: Effects of Progressive Taxes under Decentralized Bargaining and Heterogeneous Labor. 29 p.
No 32	By Tor Jakob Klette: Efficiency in the provision of municipal nursing – and home care services: The Norwegian experience. 28 p.

A complete list of this memo-series is available in a PDF® format at:
http://www..sv.uio.no/sosoek/memo/

Familiy Labor Supply when the Husband is Eligible for Early Retirement:
Some Empirical Evidences

By

Jia Zhiyang

Abstract

An econometric model modeling the joint labor market decision of married couples made in the first twelve months following the husband's eligibility of early retirement has been estimated on microdata covering the period 1993-1994. Different specifications of alternative models are presented and compared on the basis of empirical estimation results.

Key words: Early retirement, microeconometrics
JEL classification: D10, J22, J26.

Table of Contents

1 INTRODUCTION 2

1.1 SUMMARY: 2
1.2 BACKGROUND: 3
1.2.1 General 3
1.2.2 Pension system in Norway 3

2 DATA SOURCE AND THE SAMPLE 6

3 THE MODEL 6

3.1 THE PURPOSE. 6
3.2 THE CHOICE SET 7
3.3 DESCRIPTIVE STATISTICS OF THE SAMPLE 9
3.4 THE ECONOMIC ATTRIBUTES IN THE ALTERNATIVES 10
3.4.1 Disposable income and potential income simulation 10
3.4.2 The definition of leisure 12
3.5 MODEL PRESENTATION 13
3.5.1 Model A 14
3.5.2 Model B 15
3.5.3 The identification of parameters when the utility is cardinal 17

4 THE SPECIFICATION OF THE MODEL 19

4.1 UTILITY FUNCTIONAL FORM 20
4.1.1 Linear 20
4.1.2 Cobb-Douglas 20
4.1.3 Box-Cox functional form 20
4.2 LEISURE TERMS 22
4.2.1 Sleeping time 22
4.2.2 Common leisure term 22
4.3 OBS/PRE VERSUS PRE 23

5 EMPRICAL ESTIMATIONS 25

5.1 CRITERIA FOR CHOICE 25
5.2 METHOD OF ESTIMATIONS. 26
5.3 SPLIT THE SAMPLE? 27
5.4 MODEL A OR MODEL B 30
5.5 FUNCTIONAL FORM COMPARISON 33
5.6 TEST OF THE IIA ASSUMPTION 38
5.7 LEISURE TERM COMPARISON 40
5.7.1 Different leisure definition 40
5.7.2 Different common leisure term 41
5.8 PRE VESUS OBS/PRE 42
5.9 WHAT SPECIFICATION SHOULD BE CHOSEN? 43

6 POLICY SIMULATION 44

7 CONCLUSION 49

REFERENCE: 50

1 INTRODUCTION

1.1 Summary:

The focus of this paper is to discuss some empirical evidences of labor supply of elderly married men who are eligible for early retirement under the AFP scheme. Joint decision making of the husband and wife is analyzed in the framework of a discrete choice model.

Retirement behavior is a theme of increasing importance as demographic changes lead to higher proportions of elderly individuals in the population, and as pay-as-you-go public pension systems threaten to become an increasingly heavy fiscal weight for the governments of many social democracies.

Using extensive register data held by Statistics Norway, the project “Pension schemes, work activity and retirement behavior” conducted at the Frisch Centre, we hope to be able to shed some lights on the elderly labor supply in the context of family labor supply. This paper serves as a empirical analysis of the model building underlying one of the papers published as a part of this project.

The basis for the empirical work were various files with register data linked to an individual-specific number, allowing information from the various files to be linked. Information came from the labor market authorities, social security authorities, tax files and official registers containing demographic information. A file containing all married men becoming eligible for early retirement under the AFP scheme at some point in 1993 or 1994, and married to wives not so eligible is created on these basic data files.

An econometric model modeling the joint behavior for the household made in the first twelve months following the husband’s eligibility was specified in two versions. One with an option term capturing the absorbing nature of retirement in the spirit of the option value model of Stock and Wise (1990) and one without such a term.

The models were estimated by Maximum log-likelihood method. Some different specifications of the two models are presented and compared on the basis of empirical estimation results.

The best specification of the model was then used for policy simulations.

1.2 Background:

1.2.1 General

Retirement is an important phenomenon in labor supply. A lot of studies have been carried out to analysis the behavior of retirement. In many countries, the retirement age has declined for a number of years. See Lazear (1986) for detailed discussion.

In Norway, labor force participation of Norwegian males above the age of 60 has gone down over a number of years (Wadensjø, 1996). Part of this decline can be explained by the introduction of early retirement programs. The Norwegian Early Retirement Scheme (AFP) which was the result of a negotiation between employers and unions in 1988. Under this scheme, a large proportion of workers today can decide to retire at the age of 62 instead of the ordinary 67, without (in most cases) receiving lower pensions than they would, had they worked to the age of 67.

In our analysis, we used the introduction of the AFP scheme as an opportunity to study the retirement decision of elderly, married men and the responsiveness of that decision to the level of current earnings and potential. Using extensive register data covering the period 1992-1995, we were able to follow the labor market decisions of a group of couples where the man, and only he, was eligible for AFP retirement, and estimate the importance of various factors on the decision.

In Norway, female labor force participation has increased, over the last two decades by 15 percentage points, although it is still below the male level (Statistics Norway, 1996). Because a majority of men and women are married or cohabitating, it is important to account for the fact that observed behavior may be due to joint decisions by married couples. So, we analyzed the men as parts of a couple, enabling us to study the interdependence of the men's decision with the state and potential states of their wives. Blau (1998), for instance, has found "strong associations between the labor force transition probabilities of one spouse and the labor force status of the other spouse."

1.2.2 Pension system in Norway

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. Because we study the retirement decision given accumulated rights, the description below focuses on the regulations determining the benefits. Regarding the financing of the system, we will just mention that contributions to the system are levied on employers and employees as percentages of total earnings and on self-employed as a percentage of their income, as part of the income tax system. Although there is a central pension fund, it is not required that this should meet future net expected obligations, and the system is based on yearly contributions from the government.

The benefits consist of two main components. One component is a minimum pension, paid to all persons who are permanently residing in the country. With less than 40 years of residence, the pension is reduced proportionally. This reduction mainly applies to immigrants, of which there are very few in the sample, and we will not pay any attention to this feature of the system in the following. The other main component is an earnings based pension.

A crucial parameter in the system, used for defining contributions as well as benefits, is the basic pension. The basic pension in most of 1994 was 38,080 NOK. There were small adjustments during the observation period, and these were accounted for when calculating potential pensions on the basis of the basic pension. The earnings based pension in the private sector depends on the basic pension and the individual earnings history in several ways. 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. The yearly points are then multiplied by 0.45 (points obtained after 1992 are multiplied by 0.42) and the average yearly points over the 20 best years are calculated. These points multiplied by the basic pension give the earnings based component, and adding the basic pension gives the total public pension. If a person has had less than 40 years with earnings above the basic pension, the earnings based pension is reduced proportionally.

The public pension system also has a number of additional regulations, which we will only briefly recount here. Firstly, since we are still in the process of phasing in the public pension system, a special 'overcompensation' program is in operation for persons born before 1928. Secondly, there is a supplementary pension for those without any earnings based pension component, giving a minimum pension level of 1.605 times the basic pension. This means that income below 2.344 times the minimum pension does not influence the public pension. Thirdly, there is co-ordination of the pensions for married couples, mainly reducing their joint pension compared to the sum for two single persons. All of these features have been taken into account when we calculated potential pension.

Keeping 1994 regulations constant, the maximum future pension level will be 4.75 times the basic pension (G), 180,080 NOK (at the end of 1999, 1 USD is approximately 8 NOK) . This pension level requires 20 year with earnings of at least 456,960 NOK and another 20 years with earnings of at least 38,080 NOK. Although there is a re-distributive effect of the tax system also for pre-retirement earnings, this effect is much stronger after retirement. For pre-retirement earnings up to around 100,000 NOK, after-tax pension is actually higher than after-tax earnings. Also, the after-tax public pension curve is fairly flat, implying a strong re-distributive effect. The replacement level implied by the public pension curve falls from one at an income level of 2.344 G (below that level income does not influence the public pension). At earnings just giving the maximum pension, the replacement level is between 0.3 and 0.4.

State and local government employees have alternative pensions, co-ordinated so that benefits will be the maximum of the public and the government pension. The government pension is calculated in much the same way as the public pension, but with some important distinctions. First, it is based on the earnings level immediately prior to retirement and not on the previous earnings history. Secondly, the reduction in accrued pension points starts at 8 times the basic pension, allowing the maximum employer-based public sector pension to be 6.16 times the basic pension in the public system, giving a replacement ratio at that level of 0.51. In addition, there are employer based and private, additional pensions (tax deductible and widespread).

Finally, and central in our analysis, employers and unions in 1988 negotiated an early retirement scheme (AFP). Under this scheme, persons working for employers who are participating (today about 43 % of private employees and all employees of central and local government) and meeting individual requirement could retire at an earlier age than the ordinary 67. 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 was as it would have been from age 67, had the person continued till that age in the job they held at the time of early retirement.

Up to January 1 1997 (that is, in our observation period but not any more) the pension from the public system for those aged 67 to 70 was also conditioned on earnings. Firstly, 50 per cent of labor income above the basic pension – when aged 67 to 70 – was deducted from the pension. Secondly, the sum of pension and earnings were capped to the level of previous earnings.

Finally, there are also special tax rules, which apply to retirement benefits. These will be described in Haugen (2000). In the early retirement program a tax-free lump-sum amount was given to those who retired from a job in the private sector. In the government sector a higher, but taxed lump-sum amount was awarded.

2 DATA SOURCE AND THE SAMPLE

The basis for the analysis is register files held by Statistics Norway. The files are all based on a personal identification number that allows linking of files with different kinds of information and covering different periods in time.

In the present study, we analyze retirement behavior of married men who became eligible for the early retirement scheme (AFP) during 1993 or 1994 and labor supply of their wives, who are required not to qualify. Since the scheme is employer-based, we identify employers where some of the employees took out early retirement and identify all other employees in those companies. With this procedure, we may miss some companies, but are certain that those companies that are identified are participating.

Furthermore, the unit of analysis is the couple rather than the individual, and we therefore consider the male's retirement decision as part of the couple's decision on the optimal labor supply option available to them as a couple. This means that the wife's characteristics may influence the husband's retirement decision, and that the options available to the husband may influence the wife's labor supply decision. However, the couple is not symmetric, since we have restricted the sample to couples in which the wife does not have the option of retiring under the early retirement scheme in operation during the observation period. This reflects the most common situation facing older couples today, since female labor force participation was less common than male among the cohorts here nearing retirement age, and since the wife is generally younger the husband.

So for our analysis, the sample contains all couples such that the husband became eligible for early retirement under the AFP scheme during either 1993 or 1994, while the wife didn't become so eligible.

3 THE MODEL

3.1 The purpose.

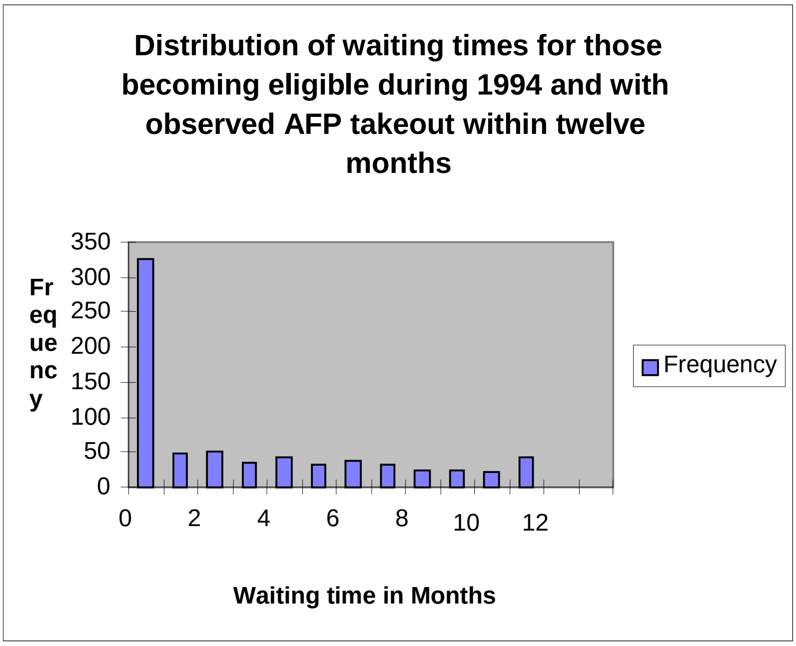
The purpose of our model is to analyze labor choice behavior of couples in which the husband is eligible to the early retirement. By labor choice, we mean the choice of work or not, how many hours to work, when to retire and so on. In fact, the decision problem facing the household can be regarded as a multinomial discrete choice problem.

3.2 The choice set

However, without simplification, the choice sets turn out to be very complicated. In fact the individual (by individual we mean the husband or the wife) is allowed to choose to work from 4 hours a week to 37.5 hours. However, it is not practical or necessary to distinguish the different alternative by each possible work hours. In most data sets hours of work is either observed in broad categories or the observations are contaminated with severe measurement errors. Although we have access to hours observed, these are only in broad categories and also probably unreliable. Therefore we let hours of work be represented by two values only, full-time work equal to 46x37.5 hours a year (1725 hours) and part-time work which is set to half of this annual load.

Meanwhile, The husband may want to take early retirement as soon as he is eligible, or he may want to wait for some months before he really take the retirement. Figure 3.1 shows the distribution of waiting time for those observed with early retirement within 12 months of becoming eligible during 1994.

Figure 3.1 the distribution of waiting time



Based on the distribution in AFP take-up frequency, we can simply split the retirement into immediate and delayed retirement. Delayed retirement means that the male does not take up retirement the very first moth he becomes eligible, but does so during the first twelve months.

For wives, things are relatively easier. They don't have the choice of retirement. But they may be out of labor force.

Following the above discussion, the choices for the household are defined as following:

For husband:

1, full time work
2, part time wok
3,delayed retirement
4, immediate retirement

For wife

1, full time work
2, part time work
5, out of labor force

Full time work is defined to be equal to 46x37.5 hours a year, while part-time work is the half of this working load. Part-time work is accessible both to persons initially working full-time and to persons initially working part-time.

Note that for the husband, state 3 and 4 are absorbing. It is to say that if the husband come into these two states, he can never be back to wage work again!

For each household, there are 4*3=12 different destination states.

Initially, the eligible individuals (husbands) are in one of two states: Full Time Work, or Part Time Work. The non-eligible individuals (wives) are in one of three states: Full Time Work, Part Time Work, or Out of Labour Force.

The following classification-criteria were used for individuals in our sample:

- **Immediate Retirement:** If the waiting time (time from eligibility to take-out) was less than two months (that is, if it was 0 or 1 month). A few individuals were observed with negative waiting times, probably revealing the fact that the implementation of eligibility criteria and the records of the take up dates were both imperfect. These individuals were classified as immediate retirement, but their low number means that their inclusion or exclusion would not affect the estimation much in any case.
- **Delayed Retirement:** If the individual waited two months or more (up to 11).
- **Full Time Work:** If the person, in the eligibility year, is observed with a job classified as 30+ hours a week, the classification is 'full time worker.'

- **Part Time Work:** If the person, in the eligibility year, is observed with a job classified as between four and 29 hours a week.
- **Out of Labor Force:** If the wife is not qualified either full time work or part time work.

Due to some problems of reliably of monthly income data, we use year as time unit. So during our observed period, they only make one choice. So there is no dynamic timing optimization involved in this version of the model.

3.3 Descriptive statistics of the sample.

Based on the choice set defined in section 3.2, Røgeberg assigns the state for all households in our sample. See Røgeberg (1999) for detailed information.

In the sample, for both the husband and wife in each household, the information available is: state occupied, age, observed income, observed income of the year before, average of the best 20 pension points, dummy for education level, dummy for working in private sector, dummy for working in service sector and etc.

The sample size was 6142, which means we had information on 12284 individuals. Table 3.1 displays the distribution of the men over the state chosen, while Table 3.2 is for women. Figure 3.2 below displays the two-way distribution of the couples over the states.

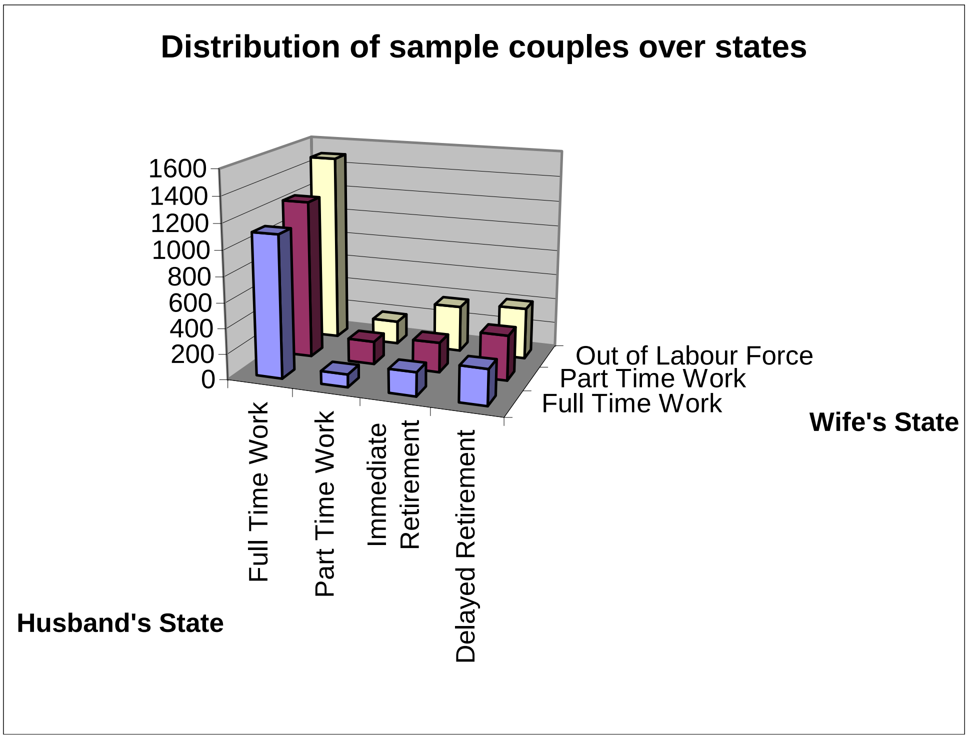
Table 3.1 states chosen by husbands

State	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
Full Time Work	3853	62,7	3853	62,7
Part Time Work	467	7,6	4320	70,3
Immediate Retirement	785	12,8	5105	83,1
Delayed Retirement	1037	16,9	6142	100

Table 3.2 states chosen by wives

State	*Frequency*	*Percent*	*Cumulative Frequency*	*Cumulative Percent*
Full Time Work	1677	27,3	1677	27,3
Part Time Work	2015	32,8	3692	60,1
Out of Labour Force	2450	39,9	6142	100

Figure 3.2 distribution of couples over states



3.4 The economic attributes in the alternatives

We assume that household, faced by the feasible opportunity set, will select the one that yields highest utility. The attractiveness or the utility of an alternative is evaluated in terms of attribute values. In our model, the choices are characterized by different level of disposable income, and different level of leisure for the household.

Blau (1998) analyses the dynamics of joint labor force behavior of older couples in the U.S., and finds strong associations between the labor force transition probabilities of one spouse and the labor force status of the other spouse, which he feels may partly be caused by a strong preference for common leisure. So in our model, we also include the common leisure term as one attribute of the alternatives.

3.4.1 Disposable income and potential income simulation

Disposable income C_{ij} (after tax income when husband is in state i, wife in state j.) is used to denote the consumption of the household. $C_{ij} = r_{Mi} + r_{Fj} - T(r_{Mi}, r_{Fj}) \quad ; i = 1,2,3,4; \; j = 1,2,5.$

where r_{Mi} is the gross income when husband is in state i , and r_{Fj} is the gross income when wife is in state j, and T() is tax function. In calculating taxes we use all details of the tax system. On average pension incomes are taxed at somewhat lower rates than labor income. The tax structure is progressive, but marginal tax rates are not uniformly increasing with income. Thus, the tax rules imply non-convex budget sets. In the estimation of the model all details of the tax structure are accounted for. A detailed description of the tax rules is available in Haugen (2000).

Note that the individual can only be observed in one state, which means that we can only observe the gross income of the individual in that state. In order to compare between the different possible outcomes, we need to impute or simulate the gross income for those states, which the individual is not observed. We need some principles to predict the potential pre-tax (gross) income. In this analysis, we use the following methods:

Full time work and part time:

There are two alternatives for predicting potential earnings in the two states part-time and full-time work.

- Use observed earnings last calendar year, and increase or reduce proportionally to obtain potential full-time earnings for part-timers and vice versa.
- Predict on the basis of an earnings function estimated on observed earnings last year.

In the first alternative we assume that if people continue to work at the same level without taking out any pension, they earn as much as they did last year, and if they move to part-time from full-time or the other way round, they face proportional increases/reductions.

In the second alternative, we remove transitory fluctuations and measurement errors in earnings, but also permanent individual variation apart from what is captured by covariates: Education, gender, industry, weekly hours group. Gross annual labor income, r_s, s=M,F, if working full-time or part-time is predicted from the estimated annual income function given below:

$$\ln r_s = X_s \lambda_s + \tau_s \ \textit{where}\ s = M, F$$

where τ_s is a normal distributed error term. The covariates entering the X-vector are:

Working full time=1, Working part-time=0,

Age,

Education, with 15 years of education or more as a reference category, otherwise three categories: less than 8 years of education, less than 10 years of education, less than 15 years of education,

Working in private sector=1, =0 otherwise, Number of years before the observation period with less than full-time work.

We observe that the income function is allowed to differ across gender and that annual gross income as a full-timer/part-timer is measured as an impact of a covariate on income.

Immediate retirement:

Potential pension following eligibility is calculated according to rules applied to an earnings history. Details are set out in Haugen (2000).

The pension level is calculated in several steps. We start by calculating potential public pension on the basis of accumulated rights, which are registered. Although this is only a part of the total pension rights it is strongly correlated with full pension. Also, since we assume that people may receive private or public pension according to the sector they work in the year they become eligible, we implicitly assume that those working in the public sector have done so for a period of time long enough for them to qualify for public pensions.

Delayed retirement:

Based on the observed take-up profile, we predict 6 more months of work and 6 months of retirement within the year we are modeling.

Out of labor force:

Wife's income when she is out the labor force is either zero or equal to the capital income and/or government transfer allocated to her.

3.4.2 The definition of leisure

We simply use the fraction of free time out of the total hours in a year as definition of leisure. Obviously, for those males in state 4 (immediate retired) and females in state 5(out of labor force), the leisure equals to one.

So we have

LM_1=1-(37.5*46)/8760;	full time work
LM_2=1-(37.5*0.5*46)/8760;	part time work
LM_3=1-(37.5*23)/8760;	delayed retirement
LM_4=1;	retirement

LF_1=1-(37.5*46)/8760;	full time work
LF_2=1-(37.5*23)/8760;	part time work
LF_5=1;	out of labor force

However, in a previous study, Dagsvik and Strøm (1997) argue that one should subtract a lower threshold from leisure, where this threshold may represent the necessary time for sleep and rest. So in this analysis I would like to also compare the estimation results to find out if the empirical evidence support their argument or not.

So we have another leisure term specification like following (we assume that the lower thresh hold for sleep and rest is 8 hours a day.):

LM1=1-(37.5*46)/8760-(8*365)/8760;	full time work
LM2=1-(37.5*23)/8760-(8*365)/8760;	part time work
LM4=1-(37.5*23)/8760-(8*365)/8760;	delayed retirement
LM3=1;	retirement

LF1=1-(37.5*46)/8760-(8*365)/8760;	full time work
LF2=1-(37.5*23)/8760-(8*365)/8760;	part time work
LF5=1;	out of labor force

3.5 Model presentation

Because the choice set, S, consist of more than two alternatives, we need to apply multinomial choice model for random utility of the couples (household). In fact, our model is based on multinomial logit (MNL).

In explaining the choices made by the couple we will employ two different, but related, models. In the first model, Model A, choice probabilities are derived from a static, utility maximizing framework. In the second model, Model B, the utility maximizing couple takes into account that if the male has chosen immediate or delayed retirement, only retirement is a feasible state next year. Thus, retirement is an absorbing state. Therefore, if the male in period t occupies states 1 or 2, then states 1,2 and 4 are feasible for him also the next period, t+1. If the male is in state 3 or 4 in period t, only state 4 is feasible in period t+1.

The reasons why it could be of interest to leave options open for flexible choices in the next period are

- pension may rise for government employee; pension is related to last year income; income in the year preceding retirement may increase due to seniority rules,
- labor income may rise next year,
- tax and pension rules may change.

Of course, the labor attachment of the female the first year puts no limitation on her choice set the following year.

Table 3.3 Feasible states

States	Male, i	Female, j
1	Full-time work	Full-time work
2	Part-time work	Part-time work
3	Delayed retirement	
4	Immediate retirement	
5		Out of the labour force

3.5.1 Model A

Note that because of a variety of reasons, the utility from our point of view should be best viewed as a random variable. See Ben-Akiva, Lerman (1985). Thus assume the following random utility function:

$$U_{ij} = u_{ij} + \varepsilon_{ij} \qquad for\ i = 1,2,3,4 \quad j = 1,2,5; \tag{1}$$

Where u_{ij} is the systematic or deterministic components of the utility and u_{ij} depends on household consumption and leisure of each of the spouses. Note that the u_{ij} may differ from household to household. So be more precisely, we should use u^h_{ij} instead of u_{ij}. But for simplicity of notation, I compress the superscript h from now on. ε_{ij} is disturbances or random components. And we assume that ε_{ij} is an extreme value distributed (Gumbel distributed) random variable. These variables account for unobserved attributes of the states that affect preferences, unobserved taste variation across the households and alternatives. The ε_{ij}'s are assumed to be IID (independent and identical distributed) across states and households with a location parameter 0, and a scale parameter σ.

Under the assumption of utility maximization, the probability of state (i,j) is chosen by the decision maker (household) is:

$$P(i, j) = \Pr(U_{ij} \geq U_{ks}, \forall (k, s) \in (1,2,3,4) \times (1,2,5)) . \tag{2}$$

Then we have

$$P(i, j) = \frac{e^{\sigma u_{ij}}}{\sum_k \sum_s e^{\sigma u_{ks}}}; i = 1,2,3,4; j = 1,2,5. \tag{3}$$

The detailed procedure derivation of the model will be given in section 3.5.3, where we discuss the identification problem of the model when the utility is assumed to be cardinal.

The deterministic part of the utility function can be defined as $v_{ij} = v(C_{ij}, L_{Mi}, L_{Fj}, L_{ij})$ where:

C_{ij} denote the disposable income, henceforth denoted household consumption
L_{Mi} denote leisure of the male in state i
L_{Fj} denote leisure of the female in state j
L_{ij} denote the common leisure of the household in state i,j

We will discuss the specification of the function v later.

3.5.2 Model B

Model B is a revised version of model A. In this case the couple is forward looking. That is to say the household take into consideration that the choices made now (or period t) has consequences for the available choice set the next period. Remember that state 3,4 is absorbing, so the choice set available at period t+1 depends on the decision made at this period. Although terms for both periods are included in Model B, the model does not attempt to explain choices made in the second period. We are only attempting to model the choice made in period t.

We assume that the household – but not the analyst – knows the current stochastic part of the utility function before deciding, but not the future ones. This assumption is the standard assumption in stochastic dynamic optimization models.

To model the behavior of the couple we thus have to specify the utility functions for this 2 periods case: $W_{ij=}U_{ij}(t)+ \gamma E(\max[U_{ks(i,j)}(t+1)])$, where ks(i,j) means the available choices at period (t+1) when the household choose state (i,j) at period t.

where γ is the discount factor. Following O'Donoghue and Rabin (1998) this discount factor can decomposed into a parameter representing "time-consistent" impatience and a parameter that represent a bias for the present. Let 1/(1+r) be the first parameter and let γ* represent the second. Thus

$\gamma= \gamma^*(1/1+r)$.

r can be associated with the rate of interest and will range from 0.02 to 0.05. If γ* equals 1, then there is no bias for the present in the intertemporal preferences and γ ranges from 0.95 to 0.98. However, if γ* is less than 1, then there is an element of time-inconsistency in the intertemporal preferences of the households. O'Donoghue and Rabin (1998) have coined this case procrastination. Although it is not possible to obtain separate estimates of r and γ*, γ can be estimated on household data like the one we employ here. Hence, one can get an indication of time-inconsistency in preferences if the estimate of γ is considerably below 0.95.

If the husband chooses 1,2 (no matter what state the wife chooses) in the first period t, the choice set available at second period is S_1={(1,1),(1,2),…,(2,5),(4,1),…,(4,5)}. Note that the delayed retirement is not an option any more.

Let $Y_1(t+1) = \max[U_{11}(t+1), \cdots, U_{25}(t+1), U_{41}(t+1), \cdots, U_{45}(t+1)]$, note that $U_{ij} = u_{ij} + \varepsilon_{ij}$ which means U_{ij} is extreme value distributed with location parameter u_{ij} and scale parameter σ, then Y_1 is also extreme value distributed. Followed from the property of the extreme value distribution, we have:

$$y_1(t+1) = EY_1(t+1) = \frac{1}{\sigma}\ln(\sum_{i=1,2,4}\sum_{j=1,2,5}\exp(\sigma u_{ij}(t+1))) + \eta/\sigma \qquad (4)$$

where η is Euler constant. ($\eta \approx 0.577$)

Similarly, if the husband choose state 3 or 4, the choice set available for the household S_2={(4,1),(4,2),(4,5)}. Because the state 3 and state 4 is absorbing, the only choice the husband is retirement immediately.

Let $Y_2(t+1) = \max[U_{41}(t+1), U_{42}(t+1), U_{45}(t+1)]$

Then $y_2(t+1) = EY_2(t+1) = \frac{1}{\sigma}\ln(\sum_{j=1,2,5}\exp(\sigma u_{4j}(t+1))) + \eta/\sigma$ (5)

So the utility function is

$$W_{ij} = \begin{cases} u_{ij}(t) + \gamma y_1(t+1) + \varepsilon_{ij} + c & for \quad i=1,2 \;\; j=1,2,5 \\ u_{ij}(t) + \gamma y_2(t+1) + \varepsilon_{ij} + c & for \quad i=3,4 \;\; j=1,2,5 \end{cases}$$

$$\begin{aligned} where \quad & y_1(t+1) = \frac{1}{\sigma}\ln(\sum_{i=1,2,4}\sum_{j=1,2,5}\exp(\sigma u_{ij}(t+1))) \\ & y_2(t+1) = \frac{1}{\sigma}\ln(\sum_{j=1,2,5}\exp(\sigma u_{4j}(t+1))) \\ & c = \eta/\sigma \end{aligned} \qquad (6)$$

we can rewrite it like following:

$$W_{ij} = w_{ij} + \varepsilon_{ij}$$

$$\text{where} \quad w_{ij} = \begin{cases} \frac{1}{\sigma}(v_{ij}(t) + \eta + \gamma\hat{y}_1(t+1)) & for \quad i=1,2 \;\; j=1,2,5 \\ \frac{1}{\sigma}(v_{ij}(t) + \eta + \gamma\hat{y}_2(t+1)) & for \quad i=3,4 \;\; j=1,2,5 \end{cases}$$

$$\hat{y}_1(t+1) = \ln(\sum_{i=1,2,4}\sum_{j=1,2,5}\exp(v_{ij}(t+1))) \tag{7}$$

$$\hat{y}_2(t+1) = \ln(\sum_{j=1,2,5}\exp(v_{4j}(t+1)))$$

$$v_{ij} = \sigma u_{ij}$$

Assuming utility maximization and similar to model A, we get the following choice probabilities

$$P_B(i,j) = \frac{e^{\sigma w_{ij}}}{\sum_k \sum_s e^{\sigma w_{ks}}};\; i=1,2,3,4;\; j=1,2,5. \tag{8}$$

using (7), it can be rewritten as following:

$$P_B(i,j) = \begin{cases} \dfrac{e^{v_{ij}(t)}\tilde{y}_1^{\,\gamma}}{\tilde{y}_1^{\gamma}\sum_{k=1,2}\sum_{s=1,2,5}e^{v_{ks}(t)} + \tilde{y}_2^{\gamma}\sum_{k=3,4}\sum_{s=1,2,5}e^{v_{ks}(t)}}; & for\; i=1,2;\; j=1,2,5. \\ \dfrac{e^{v_{ij}(t)}\tilde{y}_2^{\,\gamma}}{\tilde{y}_1^{\gamma}\sum_{k=1,2}\sum_{s=1,2,5}e^{v_{ks}(t)} + \tilde{y}_2^{\gamma}\sum_{k=3,4}\sum_{s=1,2,5}e^{v_{ks}(t)}}; & for\; i=3,4;\; j=1,2,5 \end{cases} \tag{9}$$

where

$$\tilde{y}_1 = \sum_{i=1,2,4}\sum_{j=1,2,5}\exp(v_{ij}(t+1))$$

$$\tilde{y}_2 = \sum_{j=1,2,5}\exp(v_{4j}(t+1))$$

It is easy to see that when γ=0, model B is exactly model A. So model A is an extreme case of procrastination.

3.5.3 The identification of parameters when the utility is cardinal

Utility here is used in the sense that it is an index of attractiveness of an alternative in terms of its attributes. So the household attempts to maximize it through its choice.

First of all, we need to distinguish the two important different utilities: ordinal utility and cardinal utility. Ordinal utility is only a numerical representation of the order of different choices (or more generally speaking, the consumer's rank of different bundles). It has no significance meaning of the size of the differences in utility. On the other hand, a cardinal utility implies some uniqueness of its numerical assignment. But it doesn't mean that we can measure utility as if it was a physical magnitude in the same sense as weight or height when it is cardinal.

Here we assume that the preferences of household is cardinal. The problem is that when we have a specific functional form for the deterministic part utility of the utility, can we get the 'true' estimation of the parameters through the model above or not?

Let's examine it carefully. As in section 3.5.1, we have the utility $U_{ij} = u_{ij} + \varepsilon_{ij}$. Additional, we assume that ε_{ij}'s are IID, and extreme value distributed (actually this can be defended as an approximation of the multi-normal distribution, while IID is not so easy to justify, we test part of this assumption later on by testing IIA). However, we know nothing about the mean and the variance of the ε_{ij}. Denote the location parameter of the random variable as η and the scale parameter of it as σ, then we can rewrite the utility as $U_{ij} = u_{ij} + \eta + \frac{1}{\sigma}\varepsilon'_{ij}$, where ε'_{ij} is extreme value distributed with parameter (0,1).

Then probability of state (i, j) is chosen by the household:

$$\begin{aligned}P(i,j) &= \Pr(U_{ij} \geq U_{ks}, \forall (k,s) \in (1,2,3,4)\times(1,2,5)) \\ &= \Pr(u_{ij} + \eta + \frac{1}{\sigma}\varepsilon'_{ij} \geq u_{ks} + \eta + \frac{1}{\sigma}\varepsilon'_{ks}, \forall (k,s) \in (1,2,3,4)\times(1,2,5)) \\ &= \Pr(u_{ij} + \frac{1}{\sigma}\varepsilon'_{ij} \geq u_{ks} + \frac{1}{\sigma}\varepsilon'_{ks}, \forall (k,s) \in (1,2,3,4)\times(1,2,5)) \\ &= \Pr(\sigma u_{ij} + \varepsilon'_{ij} \geq \sigma u_{ks} + \varepsilon'_{ks}, \forall (k,s) \in (1,2,3,4)\times(1,2,5))\end{aligned} \quad . \qquad (10)$$

If we define $v_{ij} = \sigma\, u_{ij}$,

We have $P(i,j) = \Pr(v_{ij} + \varepsilon'_{ij} > \max(v_{ks} + \varepsilon'_{ks}), \quad \forall\,(k,s) \neq (i,j)\ (k,s) \in (1,2,3,4)\times(1,2,5)))$

Define $U^*(i,j) = \max(v_{ks} + \varepsilon'_{ks}), \quad \forall\,(k,s) \neq (i,j)\ (k,s) \in (1,2,3,4)\times(1,2,5)))$ Then U^* is also extreme value distributed with parameter ($\ln(\sum_{(k,s)\neq(i,j)} \exp(v_{ks}))$, 1), then we can write

$U^*(i, j) = v^*(i, j) + \varepsilon^*(i,j)$, where $v^*(i, j) = \ln(\sum_{(k,s)\neq(i,j)} \exp(v_{ks}))$, and $\varepsilon^*(i,j)$ is extreme value distributed with parameter (0,1).

So:

$$
\begin{aligned}
P(i, j) &= \Pr(v_{ij} + \varepsilon'_{ij} > v^*(i, j) + \varepsilon^*(i, j) \ ) \\
&= \Pr(\varepsilon^*(i, j) - \varepsilon' < v^*(i, j) - v_{ij}) \\
&= \frac{1}{1 + \exp(v^*(i, j) - v_{ij})} \\
&= \frac{\exp(v_{ij})}{\sum_{k=1,2,3,4} \sum_{s=1,2,5} \exp(v_{ks})}
\end{aligned}
\qquad (11)
$$

Notice that this one is identical to equation (3), because we have defined $v_{ij} = \sigma u_{ij}$. In fact, it is one way that we can use to prove equation (3). It is very clear by now that we are only able to estimate v but not u, and there is no way to identify the scale parameter σ. So those parameters which enter linearly into the utility function can not be identified.

Note that although in model B, we are not able to identify γ* and r separately, we can get the 'true' estimations for γ= γ*(1/1+r). At the first thought, it seems impossible. Because from equation (7) , we may think γ is also a linear parameter of the utility function w_{ij} . However, if we exam the utility function in model B carefully, we will find out it is not the case. It is out of question that we can estimate σw_{ij} , recall the definition of the w_{ij} , equation (7), we have

$$\sigma \cdot w_{ij} = \sigma \cdot \frac{1}{\sigma}(v_{ij}(t) + \eta + \gamma\hat{y}_k(t+1)) = v_{ij}(t) + \eta + \gamma\hat{y}_k(t+1). \quad k = 1, if\ i = 1,2; k = 2, if\ i = 3,4;$$

So we are able to identify v (v=σu) and γ.

4 THE SPECIFICATION OF THE MODEL

We have discussed the main frame of the two models so far. The remaining problem for us is the specification of the model. What kind of functional form we should use in the deterministic part of the utility function? How to define the leisure term for the individual. And the common (joint) leisure term for the household? What set of rules should we imply when we impute the gross income for the household? We will discuss these questions in this and the following chapter.

4.1 Utility functional form

Up to this points in our discussion we have not imposed any functional forms on the deterministic component of the utility function. There are lots of different utility functional forms available. However, there are not enough theoretical principles on which assumptions about functional form can be made. Generally, we will be concerned with two criteria for selecting a functional form. First we would like the function to reflect any theory we have about how the various elements of attributes influence utility; second, we would like to use functions that have convenient computational properties that make it easy to estimate their unknown parameters. But sometimes, these two criteria are conflicting. In what follows we present three different functional forms to be compared with each other.

4.1.1 Linear

The simplest form is the linear functions. It is generally computationally convenient when it comes to estimate the parameters using maximum log likelihood method.

$$v_{ij} = v(C_{ij}, L_{Mi}L_{Fj}, L_{ij}) = \alpha C_{ij} + \beta_1 L_{Mi} + \beta_2 L_{Fj} + \beta_3 L_{ij}; \qquad (12)$$

α, β_1 and β_2 are constants while β_3 is assume to depend on the age difference, A_M-A_F, between the husband and the wife, $\beta_3 = \beta_{30} + \beta_{31}(A_M - A_F) + \beta_{32}(A_M - A_F)^2$. We need to point out that if β_3 is also assumed to be constant, we will have a very standard linear-in-parameters logit model. However, here the household utility function is different across the sample, it also depends on the age difference of the couple.

4.1.2 Cobb-Douglas

We can choose also the very famous functional form for utility — the Cobb-Douglas functional form.

At this case

$$v_{ij} = v(C_{ij}, L_{Mi}L_{Fj}, L_{ij}) = \alpha \ln C_{ij} + \beta_1 \ln L_{Mi} + \beta_2 \ln L_{Fj} + \beta_3 \ln L_{ij}; \qquad (13)$$

α, β_1,β_2 and β_3 have the same meaning as in section 4.1.1. Actually, from a computational point of view, it is the same as the linear function. Both of them are linear in parameters.

4.1.3 Box-Cox functional form

Another useful utility functional form is proposed by Box and Cox (1964). And discussed by many economists later, such as Hensher and Johnson (1981). Dagsvik and Strøm (1997) argued that this functional form (with some parameter restriction) is the proper functional form in theory for deterministic part of the random utility function under some reasonable assumptions and cardinal preference.

In our model, the Box-Cox function is as following:

$$v_{ij} = v(C_{ij}, L_{Mi}L_{Fj}, L_{ij}) = \alpha \frac{C_{ij}^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\beta_{11}} - 1}{\beta_{11}} + \beta_2 \frac{L_{Fj}^{\beta_{22}} - 1}{\beta_{22}} + \beta_3 \frac{L_{ij}^{\beta_{33}} - 1}{\beta_{33}}; \qquad (15)$$

α, β_1,β_2 and β_3 have the same meaning as in section 4.1.1.

α_1, β_{11},β_{22} and β_{33} are so called shape parameters, which also need to be estimated. Because they are highly nonlinear in the utility function, there are some difficulties to estimate them all. A possible solution is that assuming that all the shape parameters are equal. That is $\alpha_1 = \beta_{11} = \beta_{22} = \beta_{33}$.

Note that the first two functional forms are both special case of Box-Cox function.

The linear-in-parameters utility is the special case when all the shape parameters equal to 1:

Let $\alpha_1 = \beta_{11} = \beta_{22} = \beta_{33} = 1$, equation (15) simplify to

$$v_{ij} = \alpha C_{ij} + \beta_1 L_{Mi} + \beta_2 L_{Fj} + \beta_3 L_{ij} - (\alpha + \beta_1 + \beta_2 + \beta_3); \qquad (16)$$

The only difference between equation (16) and equation (13) is the constant term. But note that this constant term will cancel out when we estimate the model, so from an empirical point of view they are identical.

The Cobb-Douglas case can be obtained by let all the shape parameters goes to 0;

Note that we have

$$\lim_{\alpha_1 \to 0} (\alpha \frac{C_{ij}^{\alpha_1} - 1}{\alpha_1}) = \lim_{\alpha_1 \to 0} (\alpha \frac{\ln(C_{ij}) \cdot C_{ij}^{\alpha_1}}{1}) = \alpha \ln C_{ij}$$

So we have

$$\begin{aligned} &\lim_{\alpha_1 \to 0} (\alpha \frac{C_{ij}^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\alpha_1} - 1}{\alpha_1} + \beta_2 \frac{L_{Fj}^{\alpha_1} - 1}{\alpha_1} + \beta_3 \frac{L_{ij}^{\alpha_1} - 1}{\alpha_1}) \\ &\qquad = \alpha \ln C_{ij} + \beta_1 \ln L_{Mi} + \beta_2 \ln L_{Fj} + \beta_3 \ln L_{ij} \end{aligned} \qquad (17)$$

4.2 Leisure terms

4.2.1 Sleeping time

Then it comes to the leisure term, here we have two different suggestions, as I had mentioned in section 3.4.2.

One way is to define the leisure as ratio between the free hours and the total hours of a year (365x24=8760). Actually, the leisure is defined by the free hours of the individual in the coming year. And for easy understanding and computational convenience, we scale down it by divide the total hours of a year (8760 hours).

The other was suggested by Dagsvik and Strøm (1997), which says that we should deduct the sleeping time for those who works. Here working means work full time, work part time, and delayed retirement, Which is state 1,2 for both husband and wife and state 3 for husband.

4.2.2 Common leisure term

We also need to define the common (joint) leisure term for the households. By common leisure we mean the leisure enjoyed by the couple. Hurd (1997) concerns itself with the joint decision of couples. His main finding is that husbands and wives tend to retire at the same time. He himself describe it as following: 'The joint retirement hypothesis implies that as the age difference [wife's age subtracted from husband's age] increases the probability that the husband retires at an early age decreases; that is, the entire distribution of retirement ages shifts towards greater ages.' In our analysis, we like to find out if there is the same effect.

The problem here is that we cannot find a precise variable to denote the common leisure from the information available. We don't have a detailed time schedule for each of the spouses. Then we need an instrumental variable or a proxy for common leisure term. It is straightforward that we need to define it through the leisure term of husband and wife. And it is non-decreasing on both the individual leisure. That is to say if any of individual's leisure is increased, the common leisure term cannot be decreasing. Easily we will be thinking about the following possible candidates:

Arithmetic average

We can use arithmetic average of men's leisure and wife's leisure to define the common leisure $L_{ij} = (L_{Mi} + L_{Fj})/2$.

At the first thought we may think this may be a proper candidate for the common leisure term. However, if we interpret the common leisure term as the leisure enjoyed by the couple together, this one is obviously wrong. Because the leisure time together can never be greater than the minimum leisure of the spouses. And it makes a lot of trouble when it comes to interpret the estimation result. Røgeberg (1999) has given a much more detailed discussion about this issue.

Min[Lm,Lf]

A more intuitive way to measure the common leisure might be the minimum value of the spouses' leisure, simply $L_{ij} = \min(L_{Mi}, L_{Fj})$. So the common leisure is exactly the leisure of the spouse with the less leisure. But the problem for this definition is that if both of the couple are working, the time they can spend together may be less than the minimum of their free time.

Production of leisure terms for the individual.

Another possible candidate is to use the production of the leisure terms of the spouses, $L_{ij} = L_{Mi} \cdot L_{Fj}$. Note that the leisure term is less or equal to 1 for anybody, so the common leisure is always less than $\min(L_{Mi}, L_{Fj})$. And if the husband retired or the wife is out of labor force, it is the same as $\min(L_{Mi}, L_{Fj})$. But if we use the product of leisure one other problem arises, any increase of leisure of one of the spouses will necessarily increase the common leisure term.

It is not easy to decide from the theory point of view which one should be used in the estimation, although we can exclude the case of arithmetic average. The best way is to let the data to decide. We can check the estimation result of the utility function, and use the one that gives us better estimation.

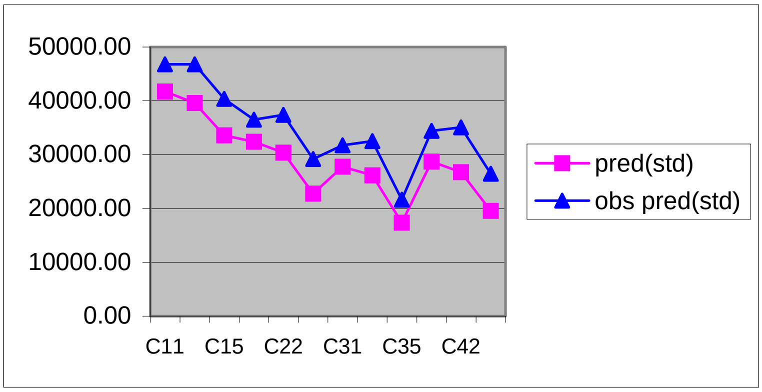
4.3 Obs/pre versus pre

As we have seen in section 3.4.1, we need to impute or simulate the gross incomes for each household for all the states except the one the household is really observed in.

For immediate retirement income of the husband, we have detailed rules to calculate the pension income of his. But for the labor income of the husband and the wife, we have different ways to predict the gross income. This is described in section 3.4.1. And we call the data set generated by the regression prediction as predicted, while the data set generated by the other method as obs\pre dataset. Which one we should use? Can this be solved by using estimation of our model and finding out the better one?

However, comparing the log-likelihood of the estimation can not solve this. Because the date set used for the estimation is different in these two cases, and the two log-likelihood values are not connected in any obvious way. Even if we can determine which set of the data fits the model better, we still cannot be sure that it is the one we should use.

Figure 4.1: Standard deviation of the disposable income



Both methods have their own advantages and disadvantages. If we use observed income the data may be more influenced by measurement error and transitory fluctuations in income. If we use predicted earnings, we will only be able to capture the variation that is linked to the covariates we observe and use in the income regression. Unobserved covariates will mean that there is permanent variation we do not capture. Besides, when constructing a regression model we have to specify how income varies with the observable covariates. Misspecification of the model may lead to biased results. And we lose some variability of the income when we use the predicted data. The Figure 4.1 shows the standard deviation for disposable incomes for the household under these two different methods of income imputations.

Table 4.1: Log labor income regressed against covariates, males

Variable	estimates	std error	t-value
Constant	12.983	0.0323	402.4
Age	0.12	0.0158	7.6
edu < 8 year	-0.487	0.0210	-23.2
edu < 8 year	-0.341	0.0210	-16.2
edu < 8 year	-0.247	0.0201	-12.3
Privat sector	0.031	0.0163	1.9
work history	-0.046	0.0027	-17.2
R^2= 0.11	Observation number = 2654		

The problem mostly concerns me is that the goodness of fit measures of the income regression R^2 is only around 10%. Although it doesn't necessary mean that the income regression model is bad. But it does mean that high variance of the individual prediction and relatively large confidence interval. Table 4.1, it is regression result for the labor income when the male is working full time.

And, Røgeberg (1999) argued that the workers in this analysis will 'to a large extent have jobs covered by union wage settlements (since this is a characteristic of most of the jobs covered by the AFP scheme), and the amount of transitory fluctuations is likely to be quite low'.

Besides, in our model, when the couples makes its choice, they will consider their gross income for the possible states, they most probably will use the last years income as a reference, rather than some regression predicts. So I personally prefer the observed\predicted data set than the predicted data set.

But whether the choice of the different data set significantly affects the estimation results or not? That is a question worthy a little more econometric estimation. So we will still do some estimations based on these two data sets, and compare their estimation results in the next chapter.

A note on the sample used for ML estimation.

After gross income simulation (before tax) we get rid of the households whose gross income of any household for any states is greater than 250,000 or less than 50,000. The reason for this restriction is that we believe that the household must earn more than 50,000 NOK a year to make a living. And those households, which earn more than 250.000 NOK a year, have different incentives when it comes to consideration of retirement decision.

5 EMPRICAL ESTIMATIONS

5.1 Criteria for choice

So far, we have several different specifications of the utility function, common leisure term and also two different models. Which one to choose?

Unfortunately, we do not have a definitive set of rules or theoretical information to tell us which one to use. In fact the model specification requires many intuitive, model-building judgments and it is very demanding. We have to rely on the comparison of the empirical results.

As a rule of thumb, there are some criteria which we can employ to say that one model is better than the other. According to Gajarati (1995), the following aspects are important issues when we compare different models:

Parsimony: A model can never be a completely accurate description of reality, to describe reality one may have to develop such a complex model that it will be of little practical use. A model should be kept as simple as possible or as Milton Friedman would say, "a hypothesis [model] is important if it 'explains' much by little". One should introduce in the model a few key variables that capture the essence of the phenomenon under study and relegate all minor and random influences to the error term.

Identifiability: For a given set of data this means that the estimated parameters must have unique values or, what amounts to the same ting, there is only one estimate for a given parameter.

Goodness of fit: since the basic thrust of modeling is to explain as much of reality, a model is judged good if the goodness of fit is high. (in regression we use R square)

Theoretical consistency: a model may not be good, despite high goodness of fit, if one or more of the estimated coefficients have the wrong signs.

Predictive power: To quote Friedman again, "the only relevant test of the validity of a hypothesis [model] is comparison of its predictions with experience."

In general, there will be a learning process that consists of a sequence of model estimation and some formal and informal tests designed to narrow down the range of alternative specifications. We need to keep these criteria in mind during the procedure.

5.2 Method of estimations.

For multinomial logit model, the most widely used method is Maximum likelihood (ML) method. Although there are still other methods that can be applied to logit model. Such as least squares, it may be computational difficult and has no significant theoretical advantage over maximum likelihood.

Simply, a maximum likelihood estimator is the value of the parameters for which the observed sample is most likely to have occurred. Although maximum likelihood estimators are not in general unbiased, they are consistent and asymptotically normal. So we can apply asymptotic t test to test whether a particular parameter in the model differs from some known constant, and the likelihood Ratio (LR) test to test some linear constraints of the parameters.

In all the estimation results reported in this paper, we also include two informal goodness-of-fit measures ρ^2 and $\bar{\rho}^2$.

$$\rho^2 = 1 - \frac{\ell(\hat{\beta})}{\ell(0)}$$

$$\bar{\rho}^2 = 1 - \frac{\ell(\hat{\beta}) - K}{\ell(0)}$$

where $\ell(\hat{\beta})$ is the value of the log likelihood at its maximum

$\ell(0)$ is the value of the log likelihood when all the parameters $= 0$;

K is the number of parameters

All the ML estimations are carried out by using econometric package TSP 4.4 on a DELL Pentium II pc. The algorithm used is Berndt-Hall-Hall-Hausman (BHHH), see Berndt *etal* (1974) for detailed information for the algorithm.

The consumption is scaled down by 100,000 at all estimations. We should be aware of this when we interpret the estimation result.

5.3 Split the sample?

There is suggestion that we should choose to divide the sample in two parts according to whether the husband works in the service sector or not. In Table 5.1 we can see that proportion of taking retirement (delayed or immediate) for those who doesn't work in service sector is much higher than those who work in services sector.

Table 5.1: Proportion of males in different states by sectors

	in service	not in service
full time work	65.98%	56.02%
part time work	8.76%	4.98%
delayed retirement	16.53%	20.16%
immediate retirement	8.72%	18.85%

In general, government employees work in the service sector and they have incentives to postpone retirement since pensions are related to the earnings the very last year of working. We thus expect the bias for the present (procrastination, Model B) to be less in the service sector than in the non-service sector.

Persons who have been working in the non-service sector may have had a harder working history and they will thus be more inclined to immediate retirement than those working in the service sector. We thus expect the leisure term for the male to be of greater importance for the retirement decision if working in the non-service sector than if working in the service sector.

Two estimations are designed here to test if data set we have support the above arguments or not. The first estimation use the full data set, model B, common leisure term $L_{ij} = \min(L_{Mi}, L_{Fj})$ and linear utility functional form. $v_{ij} = \alpha C_{ij} + \beta_1 L_{Mi} + \beta_2 L_{Fj} + \beta_3 L_{ij}$.

In the second estimation, we keep all the other specification. The only difference is that for each parameter in the utility function, we redefine it as a constant plus a parameter times the dummy related to the husband works in service sector or not. For example $\alpha = \alpha \ + D_1 s$ where s is 1 if the husband works in service, 0 otherwise.

The estimation results are set out in Table 5.2 and Table 5.3. We can see from the results, that except those dummy parameters, which are relevant to the common leisure term, the others are all significant from 0. On the other hand, we can apply LR test to compare these two estimations, The test statistic is –2(Likelihood (restricted model)-likelihood (unrestricted model)), which is χ^2 distributed with (K_u-K_r) where K_u and K_r are the numbers of estimated coefficients in the unrestricted and restricted models, respectively.

So for the null hypothesis that all the dummy parameters are 0, the LR statistic will be χ^2 distributed 6 d.f. but $LR = -2(-10528.1 + 10431.4) = 181.4$, p value<0.001. It means that we can reject the null hypothesis at very low level of significance.

Table 5.2: Estimates of linear model

Coef	variable	Estimate	t-value
β_{30}	Common leisure: constant	2.4830	5.5338
β_{31}	Common leisure: linear	-0.5182	-6.1986
β_{32}	Common leisure: quadratic	0.0134	2.4408
α	Consumption	3.3233	55.1686
β_1	Male leisure	0.2034	0.5253
β_2	Female leisure	22.5928	51.4090
γ	Discount and procrastination	0.4892	11.1205
	observations	5529	
	log-likelihood	-10528.1	
	ρ^2	0.2337	
	$\bar{\rho}^2$	0.2332	

Table 5.3: Estimates of linear model with sector dummies

Coef	variable	Estimate	t-value	Coef	variable	Estimate	t-value
β_{30}	Common leisure:const	1.6028	2.6131	D_{30}	Dummy Common leisure:const	1.6580	1.8118
β_{31}	Common leisure:linear	-0.7100	-5.3772	D_{31}	Dummy Common leisure:linear	0.2691	1.5746
β_{32}	Common leisure:quadratic	0.0229	2.4728	D_{32}	dummy Common leisure: quadratic	-0.0112	-0.9833
α	Consumption	3.6928	40.7869	D_1	D Consumption	-0.7029	-5.7447
β_1	Male leisure	2.9229	5.8763	D_2	D Male leisure	-4.8030	-5.9567
β_2	Female leisure	25.6214	39.2091	D_3	D Female leisure	-5.7858	-6.4994
γ	Discount and procrastination	0.3967	7.1284	D_4		0.2917	2.9959
	observations	5529					
	log-likelihood	-10431.4					
	ρ^2	0.2407					
	$\bar{\rho}^2$	0.2397					

However, it doesn't necessary mean that we should split the sample by the dummy that husband works in service sector or not. We can include these dummies in our estimation without splitting the sample. But at some cases, it causes computational difficulties. So we choose to split the sample when we do all the following estimations.

We noticed that the dummy parameter for the parameter for discount and procrastination is greater than 0, which mean that the estimates of the parameter for discount and procrastination is greater when the husband work in service sector than when he doesn't work in service sector. This is consistent with our expectation.

The other conjecture we have in mind is that the value of husband leisure is greater when he is in service sector than when he is not. We observe a very strong negative dummy parameter value for husband's leisure. So it is justified also.

5.4 Model A or model B

For model A we may use ML (maximum likelihood) to estimate the coefficients $\vec{\beta} = (\alpha, \beta_1, \beta_2, \beta_3, \alpha_{31}, \alpha_{32})$ (in the case of using the linear-in-parameters deterministic utility functions) or $\vec{\beta} = (\alpha, \alpha_1, \beta_1, \beta_2, \beta_3, \alpha_{31}, \alpha_{32}, \beta_{11}, \beta_{22}, \beta_{33})$ (in the case of using the Box-Cox deterministic utility functional form) by maximize the log likelihood

$$\varphi_A = \sum_{households} \log(P^h(i,j)) = \sum_{households} \log\left(\frac{e^{v_{ij}}}{\sum_{k=1,2,3,4} \sum_{s=1,2,5} e^{v_{ks}}}\right) = f(\vec{\beta});$$

The only difference model B is that we have one more coefficient γ= γ*(1/1+r), and more complicated two-periods utility function. Model A is just a particular case of model B when γ=0.

In this section, we use the obs/pre data set, and the leisure term definition 1 (no deduction of sleeping time), common leisure term 2 ($L_{ij} = \min(L_{Mi}, L_{Fj})$).

As mentioned in the last section, we split the sample into two. To the end of this chapter the estimation result reported are all based on the sub-sample in which the husband doesn't work in service sector unless it is otherwise stated.

Cobb-Douglas functional form: (estimation result is in Table 5.4)

$$v_{ij} = v(C_{ij}, L_{Mi} L_{Fj}, L_{ij}) = \alpha \ln C_{ij} + \beta_1 \ln L_{Mi} + \beta_2 \ln L_{Fj} + \beta_3 \ln L_{ij}$$

From this estimation, we can easily say that model B is much better than model A.

The null hypothesis γ=0 was rejected at very small level of significance. (P-value <0.001). Furthermore it is significant different from 0.95, $\text{t value} = \frac{0.7615 - 0.95}{0.0890} = -2.1180$, at 5% level of significance. This means that we have an indication of time-inconsistency in preferences of the household.

The goodness-of-fit measure, both ρ^2 and $\bar{\rho}^2$, increase and suggest a better fit of data in model B than model A.

Table 5.4: Estimates of a log linear model

		model A		model B	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: constant	2.9665	5.0089	3.0869	5.8150
β_{31}	Common leisure: linear	-0.8307	-6.2009	-0.5131	-4.9296
β_{32}	Common leisure: quadratic	0.0280	2.9084	0.0162	2.2356
α	Consumption	5.2834	35.1945	5.4356	39.0856
β_1	Male leisure	-1.7970	-4.4401	1.2853	2.7515
β_2	Female leisure	16.6224	34.4208	16.7046	36.4897
γ	Discount and procrastination	NA	NA	0.7615	8.5578
	observations	2892		2892	
	log-likelihood	-5891.32		-5811.39	
	ρ^2	0.1802		0.1913	
	$\bar{\rho}^2$	0.1794		0.1904	

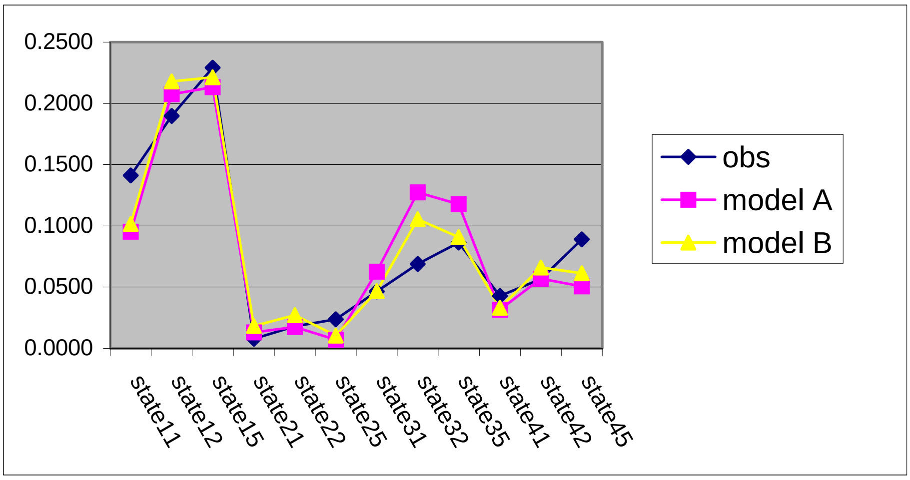
The most important aspect here, is the theory consistency. The most basic test of the model estimation output is the examination of the values of the coefficient estimates. Usually we have a priori expectation with respect to the signs and relative values of coefficients. In the estimations of model B given in the Table 5.4, all the estimates have the expected signs and expected relative values. For example, the common leisure term is decreasing when the age difference increase, and it is greater than 0, given that the age difference of the couples less than 8.5. And the leisure is much more valuable to the wife than to the husband. This fact can be explained by that in most households, wife takes the more responsibility of housework than her husband does.

However, there is one serious problem with the estimation of model A. First of all, we can reject the null hypothesis of $\beta_1 > 0$, by less than 2.5% significance. Which means that the marginal utility of the husband is less than 0, which is not so obvious to justify from the standard theoretical point of view. Even we may argue that if the husband's leisure is less or equal than the wife's leisure, the common leisure term is also the leisure of the husband, so the marginal utility of leisure for the men may be positive. But through very simple algebra, we can justify that as long as the age difference is greater than 1.5, (I restrict my discussion on the decreasing part of the quadric function) the marginal utility of the leisure for the husband is less than 0. In model B, we don't have this problem for the subset of the sample when the husband doesn't work in service sector. This can be also regarded a reason to say that model B is better than model A.

Moreover, Figure 5.1 shows the average of probability predicted by these two models across the household as well as the observed fractions in our sample. We can see that both models' predictions are quite well. But there are several relatively large deviations. Such as for state45, both models under-predict the proportion of household which choose state45. It may suggest that our model underestimate the leisure value of the males.

In general model B predicts the choice probabilities better than model A does. In particular, and as to be expected, the choices of delayed retirements far better predicted in model B.

Figure 5.1: The average of predicted probabilities across households and observed fractions



So from all these arguments above, we can have the conclusion that model B is far better than model A.

What if we change our utility functional form? Whether the conclusion is sensitive to the choice of functional form or not? Following we use the linear functional form to do the same estimation again. The results are given in Table 5.5.

This time both model A and model B give reasonable sign and relative value of the estimates. The parameter for discount and procrastination γ is still significant from 0 at very low level of significance. Most estimates are quite similar this time, with only one exception the leisure term of the husband.

Through these two comparisons we can see that both of them suggest that the parameter for discount and procrastination is significant from 0. And the estimation of model B is generally more theory-consistent than model A. So as long as it is computational practical, I suggest we should use model B rather than model A.

Table 5.5: Estimates of the linear model

		model A		model B	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: constant	1.5680	2.3521	1.6028	2.6131
β_{31}	Common leisure: linear	-0.9305	-6.0079	-0.7100	-5.3766
β_{32}	Common leisure: quadratic	0.0311	2.7802	0.0229	2.4723
α	Consumption	3.9217	41.3757	3.6929	40.7874
β_1	Male leisure	1.4055	2.9643	2.9229	5.8763
β_2	Female leisure	27.2523	40.3916	25.6216	39.2095
γ	Discount and procrastination	NA	NA	0.3968	7.1290
	observations	2892		2892	
	log-likelihood	-5571.65		-5534.61	
	ρ^2	0.2247		0.2298	
	$\bar{\rho}^2$	0.2239		0.2289	

Another thing we should be aware of is that although model A can be thought of as a special case of model B, it is constructed on different assumptions. The econometric findings above don’t necessary mean that model A is wrong! If there is a strong priori reasons to believe that when people make decisions they don’t pay any attention to the option value of the choice, we should use model A despite of the above empirical estimations.

We can also give the estimation results in other cases, such as when different leisure definition or different data set are used. But in general, these alternatives give similar results.

5.5 Functional form comparison

From now on, we use model B to compare the other specifications as long as we can estimate the coefficient when we compare the other specification effects.

Here, we still using the obs/pre data set, and the leisure term definition 1 (no deduction of sleeping time), common leisure term 2 ($L_{ij} = \min(L_{Mi}, L_{Fj})$).

We already have the estimates on the first two functional forms. To be able to compare it more easily we reproduce the results in Table 5.6.

Table 5.6: Linear (exponent=1) and log-linear utility (exponent=0)

		Exponent=1		Exponent=0	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: constant	1.6028	2.6131	3.0869	5.8150
β_{31}	Common leisure: linear	-0.7100	-5.3766	-0.5131	-4.9296
β_{32}	Common leisure: quadratic	0.0229	2.4723	0.0162	2.2356
α	Consumption	3.6929	40.7874	5.4356	39.0856
β_1	Male leisure	2.9229	5.8763	1.2853	2.7515
β_2	Female leisure	25.6216	39.2095	16.7046	36.4897
γ	Discount and procrastination	0.3968	7.1290	0.7615	8.5578
	observations	2892		2892	
	log-likelihood	-5534.61		-5811.39	
	ρ^2	0.2298		0.1913	
	$\bar{\rho}^2$	0.2289		0.1904	

The problem here is how can wen compare these 2 results? Both of them are theory consistent but give rather different estimates. For example, the parameter for discount and procrastination of the two estimations are quite different. If we only consider goodness of fit, we should choose the first functional form — $v_{ij} = \alpha C_{ij} + \beta_1 L_{Mi} + \beta_2 L_{Fj} + \beta_3 L_{ij}$. Is there any statistical test, which we can use to help us to make a decision?

The classical statistical hypothesis tests are always expressed as a comparison between restricted and unrestricted models like we have used in section 5.4 to test if the parameter for discount and procrastination is equal to 0. This kind of hypothesis is called nested hypothesis. But now we wish to compare two models, which we are not able to nest

One way to do it is to use the Cox test suggested by Cox (1962). We need to estimate a composite specification, which combine these two functional forms. And perform two likelihood ratio tests for each of the two specific functional form against the composite model. Although we managed to estimate the composite model, both LR tests rejected the null hypothesis at very low level of significance.

Then we turn to the method suggested by Ben-Akiva and Swait (1984), under the null hypothesis that the Cobb-Douglas functional form is right, the following holds asymptotically:

$$\Pr(\bar{\rho}^2_{linear} - \bar{\rho}^2_{cd} > z) \le \Phi\left\{-\left[2Nz \ln J + (K_{linear} - K_{cd})\right]^{1/2}\right\}$$

where: $\bar{\rho}^2$ is the adjusted likelihood index,

K is the number of the estimated parameters

Φ is the standard normal cumulative distribution function.

Using the estimates in Table 5.6, we find the difference between the two adjusted likelihood ratios is approximately 0.0385, and the probability that such a difference would be exceeded for a sample of 2892 observations and 12 alternatives is less than 0.001. So we should say that the linear functional form is considerably better than the Cobb-Douglas functional form!

So far we have just compared the first two functional form. We have already noticed that those two are just the particular case of the Box-Cox functional form. If we can estimate the model with the more general Box-Cox functional form, it will be able to give us some useful insights. Unfortunately, due to the computational difficulty, we are not able to estimate all the 11 variables (10 in the utility function and one parameter for discount and procrastination γ) together when we apply the Box-Cox functional form:

$$v_{ij} = \alpha \frac{C_{ij}^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\beta_{11}} - 1}{\beta_{11}} + \beta_2 \frac{L_{Fj}^{\beta_{22}} - 1}{\beta_{22}} + \beta_3 \frac{L_{ij}^{\beta_{33}} - 1}{\beta_{33}};$$

Actually the estimation difficulties arise from the nonlinearly of α_1 and $\beta_{ii}, i = 1,2,3$ and the number of exponent parameters. By restricting all exponents to be the same value , then the Box-Cox functional form will be :

$$v_{ij} = \alpha \frac{C_{ij}^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\alpha_1} - 1}{\alpha_1} + \beta_2 \frac{L_{Fj}^{\alpha_1} - 1}{\alpha_1} + \beta_3 \frac{L_{ij}^{\alpha_1} - 1}{\alpha_1}$$

at this case, we are able to estimate all 8 coefficients of the model at the same time. The result given in the following Table 5.7.

Table 5.7: Estimates of the Box-Cox utility

coef	variable	Estimate	t-value
β_{30}	Common leisure: constant	0.8691	1.3242
β_{31}	Common leisure: linear	-0.7518	-5.3998
β_{32}	Common leisure: quadratic	0.0243	2.4718
α	Consumption	3.2222	23.5617
β_1	Male leisure	3.2617	6.3783
β_2	Female leisure	26.8937	38.4305
γ	Discount and procrastination	0.3267	6.1204
α_1	Exponent	1.2049	27.0717
	Observations	2892	
	log-likelihood	-5522.88	
	ρ^2	0.2315	
	$\bar{\rho}^2$	0.2304	

We can test two null hypothesizes separately

a) $\alpha_1 = 0$, From the table we can see directly that a) will be rejected at very low significance(its t value is 27).

b) $\alpha_1 = 1$, We can simply compute the t value $= \frac{1.2049 - 1}{0.0445} = 5.4134$, and again hypothesis b is rejected at less than 0.01 level of significance.

But there is a very serious problem. Although the other variables like leisure term for husband, the parameter for discount and procrastination have right sign and relative value, a very important parameter, α_1 doesn't have the right value at all!

Note that if we have $v_{ij} = \alpha \frac{C_{ij}^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\alpha_1} - 1}{\alpha_1} + \beta_2 \frac{L_{Fj}^{\alpha_1} - 1}{\alpha_1} + \beta_3 \frac{L_{ij}^{\alpha_1} - 1}{\alpha_1}$, Then $\frac{\partial v}{\partial C} = \alpha C^{(\alpha_1 - 1)}$ and $\frac{\partial^2 v}{\partial C^2} = \alpha \cdot (\alpha_1 - 1) C^{(\alpha_1 - 2)}$, similarly we can calculate the second derivative of the leisure term. It is easy to find out the necessary condition for the utility is concave is that $\frac{\partial^2 v}{\partial C^2} \leq 0 \Leftrightarrow \alpha_1 \leq 1$!

Based on our estimation, we can even reject the hypothesis that $\alpha_1 \leq 1,$ at very low level of significance. It is very problematic for our inference of estimation and it is not consistent with the classical economical theory. How can this be accounted for?

Our restriction that all the four exponents will be same may be a reason. But unfortunately we are not able to estimate the complete model to test our hypothesis that all the parameter will be same.

Of course we can try to carry on grid searches on the exponent parameters. At least for the reason of getting better start values. Ben-Akiva and Lerman (1985) had suggested that a simple procedure is to estimate models for fixed values of the exponents and to search for the value of exponents that maximize the likelihood ratio index. But the problem is that I only manage to estimate two exponent parameters in one ML estimation, so it demands a lot of work, it need N^2 ML estimations to complete a grid search, where N is the number of different values for one parameter. And most important, even if we can use this method, we can never be sure the point we get is the maximum point. So it is not very helpful for our analysis.

Stevens (1975) have found that a modified power function fits data well. The modification consists in replacing the consumption C by income minus a lower threshold level, which can be interpreted as a subsistence level. So it may be also the case in our analysis. The estimation of the subsistence income for a household per year in Norway at 1992 to 1993 was estimated around 40,000 NOK (5700 USD at 1993). Then we rewrite the utility function as following:

$$v_{ij} = \alpha \frac{(C_{ij} - 0.4)^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\alpha_1} - 1}{\alpha_1} + \beta_2 \frac{L_{Fj}^{\alpha_1} - 1}{\alpha_1} + \beta_3 \frac{L_{ij}^{\ \alpha_1} - 1}{\alpha_1}$$

and estimate the model again. The result is given in Table 5.8,

When we compare Table 5.7 with Table 5.8, we will find out that these two estimations are almost identical including the log-likelihood and $\rho^2, \bar{\rho}^2$. With only one exception that the exponent α_1 is different. Although when we deduct the substance level from disposable income, the estimation for exponent is less than before, it is still significant from 1, not to mention that the null hypothesis that $\alpha_1 \leq 1$ is still rejected at very low level of significance.

Table 5.8: Estimates of the Box-Cox utility function with a consumption threshold

Coef	variable	Estimate	t-value
β_{30}	Common leisure: constant	0.8474	1.2933
β_{31}	Common leisure: linear	-0.7512	-5.4143
β_{32}	Common leisure: quadratic	0.0242	2.4790
α	Consumption	3.4538	31.8532
β_1	Male leisure	3.2627	6.4000
β_2	Female leisure	26.8839	38.5376
γ	discount and procrastination	0.3227	6.0568
α_1	Exponent	1.1661	32.4725
	observations	2892	
	log-likelihood	-5522.69	
	ρ^2	0.2315	
	$\bar{\rho}^2$	0.2304	

5.6 Test of the IIA assumption

At this point, we suspect that this non-concave utility function estimation may arise because of some kind of violations of the independent of irrelevant alternatives (IIA) assumptions. And if so, we need to find a way to remedy the problem to obtain an acceptable model.

There are a wide range of computationally feasible tests to detect violations of IIA assumption. See McFadden, Tye, and Train (1977) for detailed information. The test involves comparison of logit models estimated with subsets of alternatives from the full choice set.

We construct the restricted choice set $\bar{S}$ by eliminating the part time work choices for the husband. It is to say that $\bar{S} = \{(1,1),(1,2),(1,3),(3,1),(3,2),(3,3),(4,1),(4,2),(4,3)\}$. The estimation data used for the model with the $\bar{S}$ are a subset of the full data set, omitting observations with chosen alternatives not in the restricted choice set. So all those household in which the husband is working part time are omitted from the sample when we estimated the model on $\bar{S}$.

Table 5.9: Estimates on restricted choices sets and full choices set

		Restricted choice set		Full choice set	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: constant	0.2429	0.3514	0.8691	1.3242
β_{31}	Common leisure: linear	-0.7965	-5.2963	-0.7518	-5.3998
β_{32}	Common leisure: quadratic	0.0241	2.2201	0.0243	2.4718
α	Consumption	3.4283	17.6672	3.2222	23.5617
β_1	Male leisure	3.3628	6.4021	3.2617	6.3783
β_2	Female leisure	29.1302	31.0465	26.8937	38.4305
γ	Discount and procrastination	0.1857	3.5304	0.3267	6.1204
α_1	Exponent	1.2277	25.0926	1.2049	27.0717
	observations	2748		2892	
	log-likelihood	-4940.75		-5522.88	
	ρ^2	0.1817		0.2315	
	$\bar{\rho}^2$	0.1804		0.2304	

We should obtain consistent coefficient estimates under these two different choice sets. Let the estimated values for the restricted choices set $\bar{S}$ be $\hat{\beta}_{\bar{S}}$ and the estimated value for the universal choices set S be $\hat{\beta}_S$. Denote analogously the covariance matrices as $\Sigma_{\bar{S}}$ and Σ_S. Hausman and McFadden (1984) developed a test for the null hypothesis that $\hat{\beta}_S = \hat{\beta}_{\bar{S}}$. The test statistic $(\hat{\beta}_{\bar{S}} - \beta_S)'(\Sigma_{\bar{S}} - \Sigma_S)^{-1}(\hat{\beta}_{\bar{S}} - \beta_S)$ is asymptotically χ^2 distributed with K degrees of freedom, where K is the number of elements in $\hat{\beta}_S$.

The estimation results are given in Table 5.9

A comparison of the values in the table shows that almost all the coefficients for both models are quite stable. All the coefficients of the restricted choice set are within one standard error of those of the unrestricted choice set model with only one exception of the parameter for discount and procrastination.

The covariance matrixes of these two estimations are also estimated. And the statistic for Hausman and McFadden test is 14.98143. This value is compared to the critical chi-squared value with 8 degree of freedom and 95% confidence level, which is 15.5. Hence we cannot reject the null hypothesis of IIA at this level of significance.

So we have to consider the probability of using different leisure definition and/or different common leisure terms.

5.7 Leisure term comparison

5.7.1 Different leisure definition

We have mentioned that Dagsvik and Strøm (1997) suggests that we should deduct the sleeping time for those who work in section 3.4.2 and 4.2.1. What will be the result if we use their definition of leisure instead? Does the data fit well to their suggestions in our analysis?

Can this solve the problem of last section that the utility function is not concave?

We use the utility functional form

$$v_{ij} = \alpha \frac{(C_{ij} - 0.4)^{\alpha_1} - 1}{\alpha_1} + \beta_1 \frac{L_{Mi}^{\alpha_1} - 1}{\alpha_1} + \beta_2 \frac{L_{Fj}^{\alpha_1} - 1}{\alpha_1} + \beta_3 \frac{L_{ij}{}^{\alpha_1} - 1}{\alpha_1}$$

and the leisure term after subtraction for the sleeping time for those who are working. The parameter estimations are given in Table 5.10. To make it easy to compare, I reproduce the estimation result of Table 5.8(without deduction of sleeping time and subsistence income) here too.

This time all the parameter estimations have expected sign and relative value. Most important the estimation of α_1 is less than 1, which means that the utility function is concave. So far, all the parameter estimation is theory consistent!

Note that it makes no any sense to compare those two models' likelihood index. Because these two estimations are based on different data set! The value of leisure for each household is not a parameter in the utility function, it is part of the data set! So changing the definition of leisure means change of data set. So the argument that the first one gives a better fit of data or the attempt to use Ben-Akiva and Swait test we used in section 5.5 is not right thing to do.

Table 5.10: Estimates when sleeping hours are deducted

		Without deduction		With deduction	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: Constant	0.8691	1.3242	1.3766	7.2488
β_{31}	Common leisure: Linear	-0.7518	-5.3998	-0.1806	-3.8661
β_{32}	Common leisure: Quadratic	0.0243	2.4718	0.0047	1.3014
α	Consumption	3.2222	23.5617	3.1871	33.1272
β_1	Male leisure	3.2617	6.3783	0.3838	2.7365
β_2	Female leisure	26.8937	38.4305	5.5341	32.3257
γ	Discount and procrastination	0.3267	6.1204	0.6715	9.2511
α_1	Exponent	1.2049	27.0717	0.6497	16.6049
	Observations	2892		2892	
	log-likelihood	-5522.88		-5816.26	
	ρ^2	0.2315		0.1907	
	$\bar{\rho}^2$	0.2304		0.1895	

5.7.2 Different common leisure term

So far all the estimation is carried out by using the common leisure term $L_{ij} = \min(L_{Mi}, L_{Fj})$. We still have another option here, which is using the production of the leisure of the husband and the wife as the common leisure of the household. That is to say to use $L_{ij} = L_{Mi} \cdot L_{Fj}$ What will be the estimation result? Which one is better?

The estimation result is given in Table 5.11,

From the table we can see that these two estimation results are quite similar. For example, the parameter for the consumption, the parameter for discount and procrastination, the exponents are almost the same in these two estimations. Although the log-likelihood of the production case is slightly greater than that of the minimum case, it pays the price that the parameter of male leisure is not significant from 0. We can not reject that it is equal to 0 even at 0.5 level of significance.

Table 5.11: Estimates with different common leisure terms

		$L_{ij} = L_{Mi} \cdot L_{Fj}$			$L_{ij} = \min(L_{Mi}, L_{Fj})$		
Coef	variable	Estimate	std error	t-value	Estimate	std error	t-value
β_{30}	Common leisure: constant	1.5449	0.4779	3.2326	1.3766	0.1899	7.2488
β_{31}	Common leisure: linear	-0.2039	0.0330	-6.1777	-0.1806	0.0467	-3.8661
β_{32}	Common leisure: quadratic	0.0061	0.0022	2.7340	0.0047	0.0036	1.3014
α	Consumption	3.1481	0.0966	32.5836	3.1871	0.0962	33.1272
β_1	Male leisure	0.0282	0.4220	0.0668	0.3838	0.1402	2.7365
β_2	Female leisure	5.1215	0.4159	12.3151	5.5341	0.1712	32.3257
γ	discount and procrastination	0.6218	0.0692	8.9921	0.6715	0.0726	9.2511
α_1	Exponent	0.6645	0.0405	16.4188	0.6497	0.0391	16.6049
	observations	2892			2892		
	log-likelihood	-5522.88			-5816.26		
	ρ^2	0.1930			0.1907		
	$\bar{\rho}^2$	0.1919			0.1895		

5.8 Pre vesus obs/pre

So far all the estimations are based on the obs/pre data set. As we mentioned in section 4.3, whether the choice of the different data set significantly affects the estimation results or not is a still interesting issue.

The estimation results of same model, different data set are listed in Table 5.12. (Note that here I used the split sample with the husband working in service sector.)

From this table we can see that the parameters are precisely estimated. And at both cases only the parameter for common leisure quadratic term is not significant from 0. For all the parameter estimations, the signs are the same in both cases. Actually, the value of most estimates is not far from each other either.

Table 5.12: Estimates using obs/pre and pre

		Obs/pre		pre	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: constant	1.8956	7.7312	1.6548	6.8779
β_{31}	Common leisure: linear	-0.1742	-3.9058	-0.1867	-4.0799
β_{32}	Common leisure: quadratic	0.0049	1.7769	0.0046	1.4786
α	Consumption	2.7809	29.0137	1.9231	24.4756
β_1	Male leisure	-1.2772	-6.2207	-2.0149	-9.8119
β_2	Female leisure	5.1661	30.6234	3.4575	21.8276
γ	Discount and procrastination	0.7177	9.5039	0.6015	9.5658
α_1	Exponent	0.7771	18.7429	0.6919	14.5329
	observations	2637		2887	
	log-likelihood	-5060.64		-6006.76	
	ρ^2	0.2277		0.1628	
	$\bar{\rho}^2$	0.2266		0.1618	

5.9 What specification should be chosen?

Form the empirical estimations we have got above, I suggest that we should use model B, based on the obs/pre data set, and the leisure term definition 2 (with deduction of sleeping time), common leisure term 2 ($L_{ij} = \min(L_{Mi}, L_{Fj})$) in the policy simulation. The results are given in Table 5.13 for both sub sets.

From the table, we can see that both the exponent estimates are less than one, which suggests a concave utility function. And the value of the common leisure is positive and decreasing when the age difference of the couple increases.

There is one thing, which need further discussion. We can see from the table that the coefficient for male leisure is significantly less than 0, which is totally different from when we use the sample in which the husband doesn't work in the service sector. This result confirms that leisure plays a more important role in preferences for males working in the non-service sector (harder physical work) than for those working in the services sector. The tendency of a negative marginal utility of leisure for males in the service sector may indicate that there are constraints on how many hours of work one is allowed to do. Hours of work are in part regulated by the government and in part set in negotiations between employers and employees associations. To control for this one could explicitly represent these constraints in the budget constraints and hence in the choice probabilities as done in Aaberge, Colombino and Strøm (1999). Moreover, since these constraints may be of minor importance when the husband has retired, one could let the leisure term in preferences differ across states.

Table 5.13: Recommend model estimates

		Not in service		In service	
Coef	variable	Estimate	t-value	Estimate	t-value
β_{30}	Common leisure: constant	1.3766	7.2488	1.8956	7.7312
β_{31}	Common leisure: linear	-0.1806	-3.8661	-0.1742	-3.9058
β_{32}	Common leisure: quadratic	0.0047	1.3014	0.0049	1.7769
α	Consumption	3.1871	33.1272	2.7809	29.0137
β_1	Male leisure	0.3838	2.7365	-1.2772	-6.2207
β_2	Female leisure	5.5341	32.3257	5.1661	30.6234
γ	Discount and procrastination	0.6715	9.2511	0.7177	9.5039
α_1	Exponent	0.6497	16.6049	0.7771	18.7429
	observations	2892		2637	
	log-likelihood	-5816.26		-5060.64	
	ρ^2	0.1907		0.2277	
	$\bar{\rho}^2$	0.1895		0.2268	

6 POLICY SIMULATION

In order to illustrate the magnitude of the estimated relationship and the corresponding impact of potential policy changes, we have performed two simulations with the model. In the first simulation, to this end called Policy 1, pensions are taxed the same way as labor earnings. In the second simulation, Policy 2, pension is taxed as labor earnings, and pre-tax pension is reduced by 10 per cent.

Table 6.1: Choice probability in policy simulation

The state specification			Husband works in service sector			Husband does not work in service sector		
States	Husband	Wife	Model B	Policy1	Policy2	Model B	Policy1	Policy2
11	Full-time	Full-time	0.2721	0.2925	0.3040	0.1926	0.2298	0.2452
12	Full-time	Part-time	0.1440	0.1538	0.1597	0.1068	0.1250	0.1331
15	Full-time	Out of labor force	0.2445	0.3114	0.3239	0.2445	0.3357	0.3570
21	Part-time	Full-time	0.0357	0.0405	0.0422	0.0253	0.0317	0.0341
22	Part-time	Part-time	0.0200	0.0224	0.0234	0.0122	0.0149	0.0160
25	Part-time	Out of labor force	0.0249	0.0318	0.0332	0.0161	0.0222	0.0238
31	Delayed retirem.	Full-time	0.0649	0.0377	0.0305	0.0779	0.0476	0.0409
32	Delayed retirem.	Part-time	0.0391	0.0219	0.0176	0.0450	0.0260	0.0221
35	Delayed retirem.	Out of labor force	0.0695	0.0445	0.0350	0.0989	0.0639	0.0526
41	Immed. retirem.	Full-time	0.0233	0.0116	0.0084	0.0540	0.0312	0.0241
42	Immed. retirem.	Part-time	0.0137	0.0065	0.0046	0.0296	0.0159	0.0120
45	Immed. retirem.	Out of labor force	0.0482	0.0256	0.0175	0.0971	0.0560	0.0393

Table 6.2: Marginal choice probabilities for husband and wife, percent

	Husband works in service sector			Husband is not in service sector		
Husband	Model B	Policy 1	Policy 2	Model B	Policy 1	Policy 2
Full-time	66.06	75.77	78.76	54.39	69.05	73.53
Part-time	8.06	9.47	9.88	5.36	6.88	7.39
Delayed retirement	17.35	10.41	8.31	22.18	13.75	11.56
Immed. retirement	8.52	4.37	3.05	18.07	10.31	7.54
Sum	100.0	100.0	100.0	100.0	100.0	100.0
Wife	Model B	Policy 1	Policy 2	Model B	Policy 1	Policy 2
Full.time	39.6	38.23	38.51	34.98	34.03	34.43
Part-time	21.68	20.46	20.53	19.36	18.18	18.32
Out of labor force	38.71	41.33	40.96	45.66	47.78	47.27
Sum	100.0	100.0	100.0	100.0	100.0	100.0

How the average choice probabilities (the approximation of the fraction) across the sample are affected by the policy changes are set out in Table 6.1. In Table 6.2, we show how the marginal probabilities across gender are affected.

The tax system strongly favors retirement (Hernæs et. al., 1999). The columns labeled Policy 1 in Table 6.1 and Table 6.2 analyze the effect of taxing pension as earnings. As expected, this change of the budget constraint for the retired person reduces immediate retirement substantially. The proportion of households in which the husband retires immediately goes down from 8.52 per cent (Table 6.2) to 4.37 per cent when the husband worked in the service sector, and from 18.07 per cent to 10.31 per cent when the husband worked in some other sector than services. A further reduction, but not so marked, is obtained by reducing the pre-tax pension by 10 per cent and taxing pensions the same way as earnings. Also delayed retirement is drastically reduced when pensions are taxed as earnings. As expected, it is the full-time work that increase for the husband when pensions are taxed as earnings, and by as much as 10 percentage points when he works in the service sector and by 15 percentage points when he works in other sectors. These results clearly indicate that the current tax system favors retirement to a great extent and that a change in the tax rules may have a considerable impact on male labor supply among those males who are eligible for early retirement.

It is not easy to say what will happen to the wives' choices. Because wives are not eligible to the retirement, so these two policies have no direct effect on the wives' choices. Although in our model the household is analyzed as a unit, the effect to each household may be different. So the aggregate effect on the female's labor supply is ambiguous from theory point of view. In our simulation, the female labor supply is nearly not affected by the policy.

Summing up, labor income taxes on pensions give more full-time work among the husbands. Retirement, delayed and immediate, is clearly postponed. This change in the taxation of pension incomes clearly increase family labor supply among the elderly and is thus a good policy candidate if one wants to counteract the negative effects on labor supply implied by the early retirement programs.

Reducing pensions by 10 per cent also has a positive effect on family labor supply when the husband is eligible for early retirement, but the effect is a magnitude smaller compared to the effect of changing tax rules. When pension income is taxed as earnings, the tax on lower pension income is increased much more than the tax on higher income, We would thus expect that the labor supply increases are much stronger in households with lower income than in households with higher incomes. To check on this we have run regressions of the responses against the income of the household in the period prior to estimation, that is period t-1.

Let $\varphi_{i.}$(Pol r) denote the marginal choice probability for the male under policy regime r; r=1,2 and i=1,2,3,4, and let $\varphi_{i.}$(B) denote these marginal choice probabilities before the policy change and when Model B is employed. Furthermore let R denote the income of the household. In Table 6.3 below we give the result of regressing log [$\varphi_{i.}$(Pol r)/ $\varphi_{i.}$(B)] against log R. Similar calculations can be done for females. We only show the estimates and t-values of the slope coefficient and for males.

Table 6.3: The relationship between log [$\varphi_{i.}$(Pol r)/ $\varphi_{i.}$(B)] and log R

	Husband works in the service sector				Husband is not in the service sector			
	Pol 1/ModB		Pol 2/Mod B		Pol 1/Mod B		Pol 2/ Mod B	
Hus-band	Estimate	t-value	estimate	t.-value	Estimate	t-value	estimate	t-value
Full-time	-0.035	-10.9	-0.036	-9.91	-0.167	-22.3	-0.191	-21.7
Part-time	-0.020	-8.00	-0.022	-7.12	-0.133	-20.4	-0.157	-19.7
Delay. retire.	0.067	10.2	0.053	8.93	0.0854	7.92	0.0397	3.55
Immed .retire.	0.038	4.60	0.019	2.40	-0.046	-3.78	-0.106	-7.54

The pattern is as expected: the lower the household income is, the more likely it is that the husband will respond to the policy changes by increasing his labor supply.

From the model it follows that the expected consumer surplus for an household, denoted S, is given by

$$S = \ln \sum_{k=1,2,3,4} \sum_{s=1,2,5} y_k^{\gamma} \exp(v_{ks}(t))$$

where
$$y_1 = y_2 = \sum_{i=1,2,4} \sum_{j=1,2,5} \exp(v_{ij}(t+1))$$
$$y_3 = y_4 = \sum_{j=1,2,5} \exp(v_{4j}(t+1))$$

Let S(Pol r) denote the expected consumer surplus under policy regime r, r=1,2 and let S(B) denote the surplus before the policy change and when model B is employed. To check how the loss in consumer surplus from introducing the considered worsening of the budget constraints is distributed across households, we have regressed

S(pol r)-S(B) against household income for the period prior to estimation. The results of these four regressions are given in Table 6.4. Because the change in taxation of pension income hits the low income groups harder than the high income groups we will expect that the loss in expected consumer surplus is higher for the lower income households than for the higher income households.

Note that since there is a loss for all households S(pol r)-S(B) will be negative. If households with low prior income lose more than households with high prior income, we would expect that the coefficient in front of log household income is positive.

Table 6.4: The relationship between the change in expected consumer surplus, S(pol r)-S(B), and the log of household income prior to estimation. T-values in parentheses.

	Husband in service sector		Husband not in service sector	
Policy regime	Intercept	Slope	Intercept	Slope
Policy 1	-0.6468	0.0152	-3.12	0.199
	(-10.7)	(3.19)	(-20.1)	(16.16)
Policy 2	-0.7368	.016528	-3.67	0.233
	(-10.6)	(3.01)	(-19.9)	(15.99)

These results confirm the conjecture that the poor households will suffer more from the consider change in tax rules and pensions than the rich. Thus, the policy experiments considered here imply that a higher labor force participation can be obtained by changing tax and pension rules but at the expense of a less even distribution of household welfare. Therefore one has to make the familiar trade off between efficiency and equity. It should be noted that equity here is related to the distribution of household income to estimation and hence also prior to the policy experiments. The r- square coefficients related to the regressions in Table 6.4 are rather low, which indicate that prior household income is only variable among a possible large number of variables that can explain the heterogeneity in the population.

7 CONCLUSION

Among men who became eligible for early retirement during 1993 and 1994 and who were married to wives who did not become eligible, the take-up rate of early retirement was around one quarter during the first year after eligibility.

Models are constructed to explain the observed choices the very first years of eligibility for the husband under the AFP scheme. We found that joint leisure had a positive value, which means that there is a strong correlation between the labor supply decision between husband and wife. We also find indication of time-inconsistency in preferences of the household in our analysis. Different leisure definitions are discussed too, and the empirical results support the conjectures suggested by Dagsvik and Strøm (1997).

Based on the model, policy simulations are carried out. Taxing pensions as labor earnings induced a substantial decline in immediate and delayed retirement and a substantial shift towards full-time work among males Female labor supply is affected to a much smaller extent, but also the females respond by increasing full-time work when pension income are taxed as earnings. An additional 10 per cent cut in the pre-tax pension income has a positive impact on full-time work among both spouses, but the effect is a magnitude smaller than the effect obtained by changing taxation. Husbands in poor households tend to increase their labor supply more than husbands in rich households. Poor households are also more negatively hit in terms of loss in expected household welfare than the rich households are.

REFERENCE:

Aaberge, Rolf, Colombino,Ugo and Strøm, Steinar (1999): "Labour supply in Italy: An empirical analysis of joint household decisions, with taxes and quantity constraints", *Journal of Applied Econometrics*, 14

Ben-Akiva, M. and S.R. Lerman (1985): *Discrete Choice Analysis, Theory and Application to Travel Demand*, The MIT Press, Cambridge, Massachusetts, London, England.

Ben-Akiva, M and Swait, J (1984): "the Akaike Likelihood Ratio Index" Working paper, Department of civil engineering, MIT, Cambridge, Mass, USA.

Berndt.E.K., B.H.Hall, R.E.Hall, and J.A.Hausman (1974): "Estimation and Inference in Nonlinear Structural Models", *Annals of Economic and Social Measurement*, October 1974, pp. 653-665.

Blau, David M. (1998): "Labour Force Dynamics of Older Married Couples", *Journal of Labour Economics*, 1998, vol. 16, no. 3: 595-629

Cox, D.R (1962): "Further results on tests of separate families of hypotheses" *Journals of the Royal Statistical Society*, B 24:406-424.

Dagsvik,J.K. and Strøm,S (1997): "A frame work for labor supply analysis in the presence of complicated budget restrictions and qualitative opportunity aspects". Memorandum from Department of Economics, University of Oslo, No 22/97.

Haugen,F (2000): "Insentiverkninger av skatte og pensjons vegler" Master thesis, University of Oslo, (Norwegian only)

Gajarati, Damodar N (1995): *Basic Econometrics*, McGraw-Hill, Inc. Singapore.

Hausman, J.,and D.McFadden. (1984): "specification tests for the multinomial Logit Model". *Econometrica* 46.

Hensher, D.A: and L.W. Johnson. (1981): *Applied Discrete Choice Modeling*, Croom Helm, London, England.

Hurd, M. D. (1997): “The Joint Retirement Decisions of Husbands and Wives”, In *Issues in the Economics of Aging*, edited by David A. Wise, pp. 231-254. Chicago: University of Chicago Press

Hernæs, Erik, Sollie, Marte and Strøm, Steinar (1999): “Early Retirement and Economic Incentives” Memorandum from Department of Economics, University of Oslo, No 17/99

Lazear, E.P (1986): “Retirement from the labor force” In *Handbook of Labor Economics* ,edited by O.Ashenfelter and R.Layard, volume I, pp. 305-355. Elsevier Science Publishers B.V., Amsterdam, the Netherlands.

O’Donoghue, Ted and Rabin, Matthew (1998): “Procrastination in Preparing for Retirement” In Henry Aaron, editor: *Behavioral Dimensions of Retirement Economics*, pp 125-156 Brookings Institution Press & Russel sage Foundation.

Røgeberg, O. J. (1999): “Married Men and Early Retirement Under the AFP Scheme” Master Thesis, University of Oslo.

Statistics Norway (1996):"Labor Market Statistics 1995", Official Statistics of Norway, C 325, Oslo-Kongsvinger: (Dual Norwegian and English text)

Stevens,S.S (1975): *Psychophysics: Introduction to its perceptual neural, and social prospects*, Wiley, New York.

Stock, James and Wise, David E.(1990): “Pensions, the Option Value of Work, and Retirement” *Econometrica* 58, 5: 1151-1180,

Wadensjø, Eskil (ed.)(1996): *Leaving the Labor Market Early in the Welfare State* North-Holland Elsevier,. (Part 2 of Wadensjø, Eskil (ed.) *The Nordic Labor Markets in the 1990s* North-Holland Elsevier,)