Master Thesis 2, Spring 2010
School of Economics and Management
Lund University

LUND UNIVERSITY
School of Economics and Management

Modelling Retirement Determinants in Sweden on Qualitative Microdata

Author:
Malin Ivarsson

Supervisors:
Martin Nordin
Inga Persson

Abstract

The aim of this thesis is to investigate the determinants of the retirement decisions in Sweden by using qualitative micro data from The Survey on Health and Retirement in probit models. The models investigate the change in job status between the first survey wave in year 2004 and the second survey wave in year 2006/2007 by adding background variables measured in years 2004. The study show that despite there being a very distinct age pattern in retirement decisions in Sweden, factors like health, expected pension and gender do affect the retirement decisions as well. The main findings of the study are that women are found to retire earlier than men and expected pension has a negative effect on the probability of having retire in the time-span and health has a large marginal effect of the probability retirement. The empirical results also show that a large sample is not needed to clearly see the effects of these qualitative variables.

Keywords: Pension, retirement determinants, Sweden, expected pension and health

Table of Contents

List of Tables

1. Introduction

This chapter introduces the topic, and the purpose of this study. It also describes in short the methodology used to fulfil the purpose and gives a description of the disposition of the thesis.

1.1 Purpose

All over Europe the demographics has experienced substantial changes the last decades, leading to changes in the prerequisites for the retirement systems. The average age of the population has increased, the retirement age and the fertility rate has decreased in Sweden as in the rest of Europe. All these factors create new important challenges and questions for the retirement systems and labour markets in the European countries. The demographic changes make it important to examine the determinants of retirement in order for governments to make adequate adjustments. Retirement decisions have in the past often been modelled as depending on income, education, pension wealth and other types of wealth. But despite large similarities between the European countries, the institutional differences are large which makes for differences in the role and effects of such general variables. By using qualitative micro data on self-perceived health and subjective expected pension the retirement decisions of individuals in Sweden might be better explained. Including such qualitative micro data variables have in the past been difficult due to the lack of available data. But thanks to a new European database including cross-sectional and longitudinal micro data, the availability has improved greatly. The Survey on Health and Retirement (SHARE) contains qualitative micro data on health, retirement and expected pension incomes for individuals ages 50 and above. The surveys have been conducted three times, but only data from the first two survey waves have been released yet (Share-project.org, 2010).

The aim of the thesis is to, using a PROBIT, first test the ability of a model only including age, gender and then to proceed by gradually adding educational, health and total expected pension variables to explain the retirement decisions for Sweden. The probit models will be performed in Stata. The empirical study will test whether models that also include qualitative micro data on self-perceived health-status and expected pension income variables helps explain the retirement decisions better. The question to be answered is: Do pension expectations and self-perceived health help explain the retirement decisions in Sweden?

1.2 Delimitations

The study only concerns Swedish residents due to the large institutional differences between the European countries. The data was collected at two separate time points, in the years 2004 and in 2006/2007. As the retirement determinants will be investigated, only individuals in the relevant age span, age 59 to 69 in 2004, and who are not retired when the first wave was conducted in 2004 will be included in the study. The explanatory variables used are total gender, age health and total expected pension. The health variable will only be measured by self-perceived health and the education variable will only be measured by years of schooling. The age variable is measured by the age of the respondents in the first survey wave in 2004. The total expected pension variable includes the public and the occupational pension.

1.3 Disposition

The first chapter in this study introduces retirement decisions and describes the aim and method. Following this is the second chapter that presents the background and describes the Swedish retirement system, both the old and the new. Chapter two also includes retirement decision theory. The third chapter presents some of the previous research in retirement decision determinants. Chapter four describes the data used in the study and also presents some descriptive statistics. In this chapter the limitations of the study are also discussed. In the fifth chapter the econometric model used is described. Following that, chapter six presents the model specifications used in the study. The seventh chapter presents the empirical analysis and discusses the results of the empirical tests. The results reported in chapter seven are discussed and analysed in chapter eight where and in this chapter further research is also presented and discussed. Last in chapter eight the conclusions of the thesis also are presented

2. Background and Theory

This chapter presents the essential background to understand how to model retirement decisions in Sweden, how the Swedish retirement system works and how retirement decisions have previously been modelled.

2.1 Background

The demographics in Sweden and all over Europe is changing towards an older population due to longer life spans and decreased fertility. Despite the increases in the expected lifespan the retirement age in the OECD countries decreases (Blöndal & Scarpetta, 1998). The decrease in retirement age has large effects on the society and the economy in a country. For example, it results in a lower productive capacity and thereby a large unused capacity (Brugaivaini, 2005). These demographic changes pose important questions for the governments about how to increase the individuals' working lives and how to design a sustainable retirement system given the new prerequisites. A prolonged working life depends on the individuals' ability to work, i.e. their health-status, but also on institutions that allow and create incentives for a longer working life, such as social security and pension systems.

In cross-sectional studies on retirement and labour markets in the European countries, large differences are found between the countries in both age and gender distribution due to institutional differences. The northern countries have both higher average retirement ages and a larger part of the healthy population aged 50 and above in the work force (Brugaivaini, 2005). But despite the high retirement ages and the high share of the population classifying themselves as employed, there is also a large share of individuals who become disabled before becoming retired (Brugaivaini, 2008). The institutional importance also depends on to which extent the population is dependent on the public retirement system as an income source when retired, i.e. how widespread the public pension system is in the country (Brunner et al., Share 2005). In comparison with other European countries the Swedish pension system covers a larger share of the population. Both the so called second pillar and the so called third pillar have a high coverage share in Sweden in comparison, due to the strong tradition of collective bargaining by the unions (Brugaivaini, 2005).

As most European countries have a pay-as-you-go pension system with defined benefit rate, the decrease in retirement age has been beneficial for retired individuals but at a large and growing cost to the society. In a pay-as-you go pension system the contemporary working population pays for the currently retired which makes it important for the sustainability that the number of retired individuals does not grow faster than the working population. Thus the demographic changes lead to new economic and social challenges for the pension systems in terms of fiscal stability (Börsch-Supan et al., 2009).

The demographic changes are to some extent known and adjusted to but the true effect will not be possible to establish until both the indirect and the direct behavioural reactions of individuals are known (Börsch-Supan et al., 2009). According to Börsch-Supan et al. (2009) understanding the individual and societal behaviour reactions is a key feature for the social policy in its adaptation to the new demography.

In Sweden the average effective retirement age is high, 62 years, relative to other European countries. But despite this the sustainability of the Swedish pension system was in danger in the beginning of the 1990's (Brugaivaini, 2005). Because the old system's indexation was based on the consumer price index and the connection to the economic development thereby was weak, the old system was sensitive to recessions, like the one in the 1990's in Sweden. A combination of the economic state and the demographic changes, especially the increase in life length, made the system unsustainable. If the system had not been reformed, the system would have gone bankrupt in the beginning or middle of the 2100 century (Pensionsmynidgheten, 2010).

2.2 The Swedish Retirement System

In the timing decisions of retirement, the rules determining the available social security pension benefit provides strong incentives. This makes the specific social security pension system in a country a very important determinant of the retirement decision (Zamarro et al, 2007). In 1957 the pay-as-you-go system ATP (Allmänna Tillägspensionen), was introduced in Sweden. Due to demographic changes and worsened economic conditions this pension system was reformed and a new system, a Notional Defined Contribution system (NDC), was introduced in 1999 with benefits paid out from 2001 and onwards. The NDC system has a defined contribution rate and is financed like a pay-as–you-go system except for the funded individual account component. The new system was introduced to assure a long run financially and politically stable pension system. By changing from a defined benefit system to a defined contribution system, the responsibility for the financing of the system relies on the currently living instead of on future generations as in the old system (Palmer, 2002).

The NDC system was introduced using transition rules meaning that for individuals born between the years 1938 to 1953 the pension consists of a mix between the two different systems. Individuals born 1938 receive one-fifth of their benefits from the new system and four-fifth from the old system. From 1938 each cohort will receive 1/20 more from the new system (Sundén, 2004). The individuals included in the sample in our study are born between 1935 to 1945 and will thus receive more than half of their pension from the old system, except for the two youngest cohorts who will receive half from each (individuals born 1944) and 55 percent from the new and 45 percent from old system (individuals born 1945). Thus the individuals in this study belong to the cohorts that receive a mix of the two pension systems. How much of each system depends on which cohort. The survivor pension (efterlevnadspension) as well as the previous pre-retirement pension were reformed in 2003 (Pensionsmyndigheten, 2010).

2.2.1 The Old Swedish Retirement System

The old pension system in Sweden was a pay-as-you-go pension system. The system consisted of three parts, a basic pension (Grundpension), the so called public pension (Folkpension), and a supplementary pension, the so called occupational pension (Allmänna tilläggspensionen, ATP) (Sundén, 2004). The basic and the public pension were income independent and available to all Swedish residents. The normal retirement age in Sweden was 65 years, but the basic pension was available from age 60 but then with a reduction by 0.5 percent for each month below the normal retirement age (Karlstrom et al, 2004). The occupational pension was dependent on the individuals' earnings, pension points for the fifteen years with the highest income determined the size of the pension. In addition, to get a full pension 30 years with pension points was required (Pensionsmynsigheten, 2010).

The occupational pension was indexed by the consumer price index and based on the individuals' earnings (Karlstrom et al, 2004). As a consequence of the price indexation, the benefits were not linked to the wage growth, which in turn resulted in that the benefit could rise faster and more than the wage growth and the contributions. This was also possible when the economic productivity growth was negative, which made the skewed relationship between contribution rate and benefit rate even more skewed (Sundén, 2004). The skewed relationship between the contributions and the benefits was also due to that the contributions were paid on all the individual's years as active in the labour force (or from 16 years of age as earliest) whereas the benefit was based on the 15 years with highest earnings and 30 years of earnings. These rules on the contribution and benefit base were profitable for highly educated individuals with a steep income curve but less profitable for individuals with a long working life who typically are low income earners (Sundén, 2004).

2.2.2 The New Swedish Retirement System

The new pension system is, as mentioned above, a Notional Defined Contribution system, meaning that each worker has an individual account where his/her contributions and thereby future pensions benefits are recorded but the contemporary incomes are used for the currently retired, like in a pay-as-you go system. The new system has no official retirement age, but the pension is available from age 61. In the new system the pension is based on the income over the entire life-cycle in contrast to only on the fifteen highest income years (Sundén, 2004). The indexation of the new pension system also differs from that of the old, where the new system uses the national per capita wage growth. This change was made to make the new system more sustainable to changes in the economy (Pensionsmyndigheten, 2010). One important feature of the new system was to strengthen the connection between the payments and the benefits and to remove the defined benefit promise. The new system also contains an automatic balance system to make sure that the debts of the system never exceed the assets. If the balance would be negative the pensions are recalculated with a so called balance index (Sundén, 2004).

The new system consists of three different parts, a public pension (Allmän pension), occupational pension (Tjänstepension) and possible voluntary private pension (privat pension) (Pensionsmyndigheten, 2010). The public pension in turn consists of three different parts, an income pension (Inkomstpension), a premium pension (Premiepension) and for some a guarantee pension (Garantipension). The occupational pension covers more than 60 per cent of the population and more than one fifth has a private old-age pension in Sweden (Brunner et al. 2005). To the public pension system 18.5 percent of the individual's income is contributed. From this the income pension, the notional account, receives 16 percent and the premium pension the remaining 2.5 percent (Pensionsmyndigheten, 2010). The 2.5 percent is the individual defined contribution part of the NDC system which the individual can choose to invest in up to five different funds (Sundén, 2004). There is an income ceiling meaning that despite that the pension contribution is paid on the entire income of an individual, only incomes up to the income ceiling give pension benefits. The contributions paid on income above the incomes ceiling are paid to the national budget (Pensionsmyndigheten, 2010).

The income pension is based on the individual's income over the entire life-cycle and the development of the benefit depends on the development of the income index and thereby differs from year to year. For those individuals who are born before 1954, and thereby receive a mix of the new and old system, the occupational pension (ATP) is also indexed by the income development after the reform. As the development of the incomes is in a direct relation to the development of the economy as a whole, the new indexation mirrors the state of the economy better than the old system. The development of the premium pension, the notional account, depends on how the securities markets develop and is thereby not affected by the automatic balance system (Pensionssystemet, 2010).

Individuals whose earnings have not been large enough to ensure a substantial pension level receive a guarantee pension which is available at the age of 65 and represents around 35 percent of a blue-collar worker's average income. The guarantee pension is indexed by the consumer price index but also affected depending on if you receive any other pensions such as income pension or ATP. As the guarantee pension is indexed by prices and its development thus has a weak connection to wages, the importance of this pension part will be reduced as wages grow. The guarantee pension is collected by around 30 percent of the presently retired (Sundén, 2004). The development of the occupational pension depends on the type of occupational pension and is managed by the different pension firms (Pensionssystemet, 2010). How much a retired individual will receive per year from the income (and ATP) system is calculated based on the expected life length (a unisex mortality table) for the individual's cohort at age 65 and an imputed real rate of return of 1.6 percent (Sundén, 2004).

The new system will imply that individuals have to work longer in order to reach a specific replacement rate, due to the increase in life expectancy, which may be difficult for some job types that are physically demanding. In the new system there are incentives for continuing to work in high ages, as the extra benefits from working an extra year are actuarially fair and there is no upper age limit for earnings that give pension rights. In conclusion, a greater responsibility is put on the specific individual regarding his/her pension in the new system compared to in the old (Sundén, 2004).

The new pension system was designed to create incentives for working longer. But despite a prolonged working life being a desirable outcome for the Swedish economy there are other practices than the pension system that give opposite incentives. When the new pension system was introduced the Swedish legislation was also changed to enable the legal possibility to work past the age of 65 until the age of 67. But workers older than 65 years old are not entitled to sickness or unemployment insurance. This precaution was taken to ensure that individuals do not use sickness or employment insurance as a early retirement. Besides this, most employers are not interested in employing individuals older than, or in the proximity of, 65 years (Sundén, 2004).

2.3 Retirement Decision Theory

In the most basic model the retirement decision is regarded as a function of earnings from work, private pension and social security benefits (Fields & Mitchell, 1984). The received pension per year for an individual is an increasing function of the numbers of years of working life but postponing retirement also infers fewer years to collect the pension. An individual will try to maximize his/her intertemporal utility function subject to his/her specific intertemporal budget set. The intertemporal budget set is the present discounted value of expected lifetime income (PDVY), which in turn is the sum of the present discounted value of earnings (PDVE) and the present discounted value of pension (PVDP)

$$PDVY = PDVE + PDVP = \int_0^R E_t \delta_t + \int_R^T (P\,P_t + SS_t)\delta_t dt \qquad (2.1)$$

where E_t represents the after-tax earnings of the individual, δ_t represents the discount factor between time preference and mortality, PP_t represents private pension in time t and SS_t represents social security benefits streams. The intertemporal utility function is represented by the number of leisure years (RET), $RET = T - R$ where T is number of total years and R is retirement age, and the utility is increasing with increasing value of RET. The utility is thus a function of present discounted value of expected lifetime and number of leisure years.

$$U = U(PDVY, RET) \qquad (2.2)$$

The utility is maximized when the marginal utility of one more year as retired is equal to the marginal utility of one more year working (Fields & Mitchell, 1984).

The present discounted value of expected lifetime earnings can in turn be divided into two different subgroups. The subgroups are base wealth (YBASE), which is the present value of the income that is to disposal at the earliest retirement age, and the increase in present value of income due to deferring retirement (YSLOPE). Fields & Mitchell (1983) show that due to the income effects, an increase in the base wealth decreases the retirement age whereas an increase in the increase in present value due to working longer theoretically has two different effects on retirement age. If the increase in the present value of income due to working longer increases, the leisure time becomes more expensive as the opportunity cost would be higher and an individual would thereby be willing to substitute leisure time for working time thus higher retirement age, according to the substitution effect.

The income effect on the other hand would have a decreasing effect on the retirement age as a higher present value of income would enable the individual to “afford” to be retired earlier. Empirical tests have been performed on U.S. labour data to determine if the income effect or the substitution effect dominates (Fields & Mitchell, 1984). The results show that as expected there is a negative relationship between an increase in the base wealth and retirement age. Furthermore the increase in present value du to working longer has a positive effect on retirement decisions, i.e. increases the retirement age. The empirical testing also shows that the YBASE and the YSLOPE have a large and significant effect on the retirement decision (Fields & Mitchell, 1984).

As shown by Fields & Mitchell (1984) the economic incentives have large effects on the retirement decisions. The change in Sweden from the old to the new pensions system has resulted in large changes for the individual’s incentives to retire. The new system changes the YSLOPE compared to the old system and is thereby likely to affect the behavioural decisions taken by Swedish individuals approaching retirement. How economical changes in benefit systems can change the behaviour of individuals is further studied in the next chapter where other explanatory variables than economical also are investigated.

3. Previous Research

In this chapter some of the previous research on retirement decisions is presented. This also gives a perception of which variables are found to be important to include in retirement decision models.

3.1 Previous Research

Retirement decisions have frequently and extensively been investigated and modelled in economics. Despite the extensive amount of studies the direct determinant variables have proven to be difficult to establish. That economic factors play an important role in the retirement decision has been concluded, but the magnitude of its importance and whether it is the most important variable differs between different studies. An often agreed conclusion is that the institutional factors, i.e. social security pension rules etc., play a significant part in the decision process and that finding a global model for the retirement decision will thereby be troublesome. Fields & Mitchell (1984) investigated the economics of retirement in the U.S and argued for the economic factors being an important explanation as it gives income and leisure opportunities a large role. They conclude that income is an important variable in individuals' retirement decisions since different income opportunities have a significant impact on when individuals choose to retire and also that there is heterogeneity in the preferences for income and leisure both within firms and across firms. To test how the retirement decision changes with changes in the budget constraint two models were used. The first model is an ordinary least square (OLS) model and the second is a discrete choice model. The discrete choice model used is a conditional ordered logit (OL) motivated by that in the (OL) the dependent variable is permitted to be a function of how appealing the next alternative is (Fields & Mitchell, 1984).

In the ordered logit model they use the following utility function

$$U_{ij} = (\alpha \log PDVY_{ij} + \ \beta \ log \ RET_{ij}) + \varepsilon_{ij}$$

where they test the average taste parameters, α and β, for the ith worker at the retirement age j. The ratio between α / β then determines the preference of working longer, and thus earning more, compared to a longer retirement period with lower income. The empirical tests of the ratios showed that individuals do consider income and leisure opportunities when choosing their retirement age. The sensibility of the different economic parameters was tested by exposing them to shocks. Earned income is found to have an increasing effect on the retirement age while retirement benefits such as social security or private pension have a decreasing effect on the retirement age. A ten percent increase in the retirement benefit is found to lead to around one month decrease in retirement age. About the same increase in retirement age is found from a ten percent increase in earned income. Fields & Mitchell (1984) concludes that despite the impact of changes in the economic parameters on retirement decisions there would have to be large changes in the policy variables to obtain a marked change in the retirement age (Fields & Mitchell, 1984).

The decision of whether to retire or to continue to work does however depend not only on economical factors. Johnson (2009) argues that many studies have found that continuing to work in older ages improves not only the economical factors but also improves the well-being and the physical-health of an individual. The sense of identity, social integration and social support are found to be improved by working longer as one then remains active and socially engaged (Johnson, 2009). Dwyer & Mitchell (1999) also finds that health problems have a significant impact on men's retirement decisions, even more than the economical factors, when regressing health conditions on expected retirement age.

Health is found to have many various effects on labour market participation, affecting it in different directions. Besides that bad health can have limiting effects on the possibility to work, health can also affect the productivity and thus also earnings. A worsened health can alter an individual's relative preferences between consumption (thus work-time) and leisure time. Both the change in productivity and in preferences leads to an increased relative utility from leisure which results in a decrease of the retirement age. There may also be an opposite effect on the retirement as poor health also might have the effect that the marginal utility of consumption increases relative to the one of leisure. If bad health leads to a shorter life, the life horizon also changes making the individual relatively richer as the wealth per year increases, which should in theory lead to an income affect. Despite these contrasting effects in theory, empirically bad health is found to have a decreasing effect on the retirement age due to a domination of the preference and productivity effects (Dwyer & Mitchell, 1999).

Using subjective measures on health is often criticised due to that the health effects may be misestimated since individuals could claim to retire due to non existing health reasons as health sometimes is not directly observable, the so called Justification hypothesis. This may be true especially for low income individuals in countries with generous social security pension, such as Sweden. Some empirical investigations find support for the hypothesis but

Dwyer & Mitchell (1999) find, by using some different measures for health as well, that there is no support for that hypothesis. The reason for the lack of support of the Justification hypothesis may be due to that the data sample used contains relatively young and not yet retired individuals and different measures for the explanatory variables. Other more objective ways of measuring health has however also some problems but in the opposite way, meaning that some individuals who have strong preferences for their current work were found to have a tendency to overestimate their health and vice versa (Dwyer & Mitchell, 1999).

To avoid any possible problems associated with subjective data Dwyer & Mitchell (1999) use both self-perceived health measures, so called unobserved health indicators, and observed health indicators from the Health and Retirement Study in 1992. By using a more extensive database including surveys on economical and personal questions with two types of health measures, they are able to estimate the magnitude of how different health problems affect retirement decisions. As expected, bad health is found to have a negative impact on retirement age. Of the economical variables, having a health insurance or not (health insurance controls) is found to be the one with most effect on retirement age. Large differences are found between different healths problems when it comes to health's impact on the retirement age, where conditions with greater physical limitations have a larger impact (Dwyer & Mitchell, 1999).

Education has also been found to be an important explanatory variable in retirement decisions. The educational variable can be used as a measure for low or high-skilled occupation and can thus have both an income effect and substitution effect, like the income variable discussed above. The family has as well been suggested to have an impact on an individual's retirement decision, for example by coordination of the timing of retirement points between husbands and wives. Coordination between wife and husbands in their retirement decisions results in a lower retirement age for the wife as the age of the wife often is lower than the one of the husbands (Johnson, 2009). If this is true for Sweden will be tested in the empirical models. To conclude, previous research suggests that both economical and subjective factors such as health affect the retirement decisions. How these variables affect the retirement decisions in Sweden will be tested in chapter 7 where the results from the empirical tests will be reported.

4. Data and Descriptive Statistics

Chapter four will describe and discuss the data used and as well present the variables used in the models in this study. Some descriptive statistics of the data sample will also be presented.

4.1 Data and limitations

The data used has been extracted from the Survey of Health, Ageing and Retirement database (SHARE) which contains cross-sectional and longitudinal micro data from 11 European countries. The SHARE database contains in all over 45 000 interviews of individuals aged 50 or older on topics such as economic situation, expected pension and health. The data used in this thesis is only the interviews conducted in Sweden. The survey was conducted at two different times, the first wave in 2004 (and one interview in January 2005) and the second wave from October 2006 to August 2007 in Sweden. All the interviews in Sweden were conducted in Swedish (Share.org, 2010). When modelling retirement decisions, in this study, the relevant age of the respondents will be taken to be between 59 and 69 years old as it is in this age span most individuals retire. All the modelling will be done in stata by using a dprobit that estimates the marginal effects for the variables.

The aim of the models is to investigate the change in current job status between the first survey wave in 2004 and the second survey wave in 2006/2007, for the respondents who were still working and were between 59 and 69 years old in 2004 when the first survey wave was conducted. By adding explanatory variables from the first survey wave in 2004, the model investigates whether these variables can explain the change in job statuses between the two survey waves. So the models will try to explain the change in the respondents' job statuses between 2004 and 2006/2007. In the model the dependant variable is the current job status of the respondents in the second wave. In model one to five a more restrictive definition of the classification of individuals as retired is used compared to model six to ten. To be classified as being retired in the first definition the respondent must receive the majority of his/her income from public and occupational pension and not from labour market income. In the second definition of retired used, the individual will be classified as retired if he/she does not report any labour market income. The first definition of retired respondents in the sample generated a smaller share of retired in the sample than seemed likely. The small share of retired individuals was probably due to problems with the data, so to get around this data problem the second definition of retired respondents in the sample was used in model six to ten.

All the explanatory variables are collected from the first survey wave in 2004, so they are all related to the first survey wave. That implies that the age variables are the respondents age in 2004 and their health status is the respondents health status in 2004 and likewise for the other variables. There is a dummy for age corresponding to each age between 59 and 69. Education is measured by years of schooling and then divided into three different dummies of educational attainment and a reference group. The reference group corresponds to nine years of schooling or less. The first educational attainment dummy, secondary schooling, corresponds to ten to twelve years of schooling and the second dummy, the low university degree, corresponds to thirteen to fifteen years of schooling. The last dummy, the higher university degree includes individual who have studied for sixteen years or more. The total expected pension is measured by percent of the current wage.

Health is difficult to measure since it is not directly observable. Here health is measured as self-perceived health, where the respondents answered how they perceive their health in general. The answer alternatives for the health variables where divided into three categories: excellent to very good, good and pretty bad to bad, where good was used as the reference group. As self-perceived health status is a subjective measure there may be some problems with individuals trying to use bad health as a reason for retiring early, the justification hypothesis, and the health variable should thus be regarded as a proxy with some limitations. The problem arises if the self-estimated health is endogenous to the labour supply choices of the individual, since if there is a correlation between the economic variables and omitted taste variables the behavioural response estimates will suffer from bias. If the health effects on retirement are extremely large and substantially larger than the economical effects, the justification bias is likely to exist in the specific dataset (Dwyer & Mitchell, 1999).

The variables used in this study are shown in the table below together with explanations and definitions. All explanatory variables are collected in the first wave in 2004.

Table 4.1 Variables and Definitions

Variables	Description
Retired definition 1	Job status of the respondents in wave 2. Retired if pension income is larger than wage income. Used in model 1-5. 1= Retired 0 = Working
Retired definition 2	Job status of the respondents in wave 2. All respondent who have not reported a wage income are defined as retired. Used in model 6-10. 1= Retired 0= Working

Gender	Dummy variable for gender where 1= male and 0= female
Reference group for Age	Dummy for age 59 in survey wave 1 (2004). Used as reference group.
Age dummy 2- Age dummy 9	Dummy for age 60- age 69 in wave 1 (2004)
Reference group for Education Basic level	Dummy for educational attainment. Years of schooling 1 to 9. Used as reference group.
Secondary school	Dummy (1) for 10-12 years of schooling.
Low university degree	Dummy (2) for 13-15 years of schooling.
High university degree	Dummy (3) for 16 and more years of schooling.
Very good health	Dummy for self-perceived health status, excellent to very good in wave 1 (2004).
Reference group for health Good health	Dummy for self-perceived health status, good in wave 1 (2004). Used as reference group.
Bad health	Dummy for self-perceived health status, fair to poor in wave 1 (2004).
Dummy no answer, health	Dummy for the missing observations of the respondents on the health variable.
Total expected pension	Total expected public and occupational pension as percent of current wage in wave 1 (2004).
Dummy don't know expected pension	Dummy for the missing observations of the respondents on the expected pension variable in wave 1 (2004).

4.2 Descriptive statistics

The Survey of Health, Ageing and Retirement (SHARE) was conducted at two times in Sweden, year 2004 and in 2006-2007. In the first wave in 2004, 4011 respondents participated where 770 of them were aged between 59 and 69 years when the first survey wave was conducted in 2004. The respondents were somewhat fewer in the second wave in year 2006-2007, and then 3660 respondents participated. Totally 2010 respondents participated in both the first and the second wave. Table 4.1 shows the numbers of respondents in each wave together with interview year, for the respondents who participated in both survey waves.

Table 4.2 Number of Respondents who participated in both Wave 1 and 2.

	Interview year W2		
Interview year W1	2006	2007	total
2004	966	1043	2009
2005	0	1	1
total	966	1044	2010

In our study of retirement decisions the chosen age interval is 59 to 69 years old, as the retirement decision was found to be most central in these ages as the public pension in Sweden is available from age 61. Table 4.3 shows the age frequency of the respondents, who participated in both the survey waves and how many who were classified as workers in the first wave and as retired in the second wave according to the two different definitions.

Table 4.3 Ages and Frequencies for Working in Wave 1 and Retired in Wave 2*

Respondents	**Wave 1**		**Wave 2**	Definition 1	Definition 2
Age	Frequency	Working	Frequency	Retired	Retired
59	82	76			
60	88	79			
61	84	73	39	2	30
62	70	63	66	2	37
63	66	50	88	7	42
64	64	43	64	9	30
65	88	43	61	8	21
66	69	4	43	18	6
67	65	6	43	3	10
68	45	3	26	5	5
69	49	4	4	1	1
70			4	2	0
71			5	0	0
72			1	0	0
Total	770	444	444 (238 & 433)	57	182

*according to the different definitions.

The table shows the large difference in number of respondents who are classified as retired between the two different definitions. The number of observations is 444 in wave two but when the other variables are added in the models the number of observations decrease substantially due to a lot of missing values. In model one to five the sample contains 238 observations and in model six to ten, where the less restrictive definition of job status in wave two is used, 433 observations. When using the second definition of retired respondent in wave two in the sample the share of respondents classified as retired rose from 23.85 (57 of 238) percent to 42.03 (182 of 433) percent. The age distribution of the retired respondents in the second wave clusters around the age 61-65 years old.

5. Econometric Model

Chapter five presents and explains the econometric model used in the study, why it is used and how to interpret the results.

Often in micro-economics the data used in models are on individual level and the dependent variables are limited. To model these types of micro-economic data a different type of method is needed than the ordinary linear regression models. To model choices between two discrete variables Binary Choice Models are often used instead of linear regression models (Verbeek, 2009). If linear models are used on a model where the dependent variable, y_i, only can take two different discrete outcomes, $y_i = 1$ and $y_i = 0$ for individual i where = 1 is being retired and = 0 is not being retired and where there are observations on the explanatory variable x_{i2} such as for example income. In the model there is also an intercept term β_1 which is defined as $x_{i2} = 1$. The model would then have the following appearance

$$y_i = \beta_1 + \beta_2 x_{i2} + \varepsilon_i \tag{5.1}$$

which could also be written as $y = x_i'\beta + \varepsilon_i$ where x_i is a vector of the x's. If we assume that the x variables are exogenous, $E\{\varepsilon_i|x_i\} = 0$, and that the regression line describes the conditional values of the dependent variable given the values of explanatory variable, $E\{y_i|x_i\} = x_i'\beta$, the model could have the following appearance

$$E\{y_i|x_i\} = x_i'\beta \tag{5.2}$$

since,

$$E\{y_i|x_i\} = 1.P\{y_i = 1|x_i\} + 0.P\{y_i = 0|x_i\} = P\{y_i = 1|x_i\} \tag{5.3}$$

(Verbeek, 2009). According to this model $x_i'\beta$ is a probability that the individual will be retired and the value therefore lies between 0 and 1.

There are however some large problems with applying this linear model to the data where the dependant variable is bounded, such as that the distribution of the error term is non-normal and it is heteroskedastic. Some of the restrictions needed for the model to work, but that rarely is fulfilled, are that the explanatory variables have to be bounded and some restrictions on β have to be fulfilled. The limited outcomes of the dependant variable impose the same restrictions on the outcomes of the error term, leading to that the variance of the error term depends on the explanatory variables and the parameters in the model, β, and therefore is not consistant (Verbeek, 2009).

5.1 Probit

A more appropriate method to use on micro-data with limited dependent variable is Binary Choice Models and one type is the probit model which is often used when the dependant variable is a discrete variable (Verbeek, 2009). The Probit model will be used to estimate how the chosen variables, such as income, pension and education and so on directly affect the probability of being retired, i.e. that $y_i = 1$. In the retirement decision model presented in this thesis the dependant variable, $y_i = 1$ if the individual is retired and $y_i = 0$ if the individual is not retired. The model will have the following appearance

$$P\{y_i = 1|x_i\} = G(x_i, \beta) \tag{5.4}$$

and states that the probability of being retired is a function, $G(x_i, \beta)$, of individual specific characteristics such as for example income and pension. The values of the $G(x_i, \beta)$ function can only be in the interval between 0 and 1 and the function should have the shape of $F(x_i'\beta)$ which implies that the F function as well only can take on values in the same interval. Based on this, F can therefore be seen as a distribution function where the choices between the distributions lie between the standard normal distribution function and the standard logistic distribution function. If the normal distribution is chosen the linear probability model will be a probit model and if the logistic distribution is chosen the model will be a so called Logit model. The probit and the logit are similar in their appearance, both models have the expected value of the error term as zero, but the logit model has fatter tails than the Probit (Verbeek, 2009).

As described above in the theory chapter 2.3 the decision to retire or not depends on if the utility as retired is larger than the utility of continuing to work. In the retirement decision there may be some unobserved variable influencing the decision y_i^*, a latent variable, that is a function of the observed variables.

$$\boldsymbol{y_i^* = x_i'\beta + \varepsilon_i} \tag{5.5}$$

If the two outcomes are 0 and 1 as before, the function with a latent variable can be modelled as binary model.

$$P\{y_i = 1\} = P\{y_i^* > 0\} = P\{x_i'\beta + \varepsilon_i > 0\} = P\{-\varepsilon_i \leq x_i'\beta\} = F(x_i'\beta) \tag{5.6}$$

As normal in the binary choice models, the form depends on the assumed distribution of the error term (Verbeek, 2009).

A problem with the Binary Choice Models such as Probit and Logit is that the coefficients are difficult to interpret. The normal interpretation of the signs, that a minus indicates a negative result, can of course be made. In order to be able to compare results from different models, the marginal effects of the change in the independent variables could be calculated through the partial derivative of the probability of being retired (Verbeek, 2009). For the Probit model the marginal effect would be the following

$$\frac{\partial \Phi(x_i'\beta)}{\partial x_{ik}} = \phi(x_i'\beta)\beta_k \qquad (5.7)$$

where $\phi(x_i'\beta)\beta_k$ is the standard normal density function (Verbeek, 2009). The change in the continuous variable, x_{ik}, is dependent on the values x_i, of the explanatory variable. If the explanatory variables as well are discrete, i.e. dummies, which they often are when modelling retirement decisions, the probabilities of the two different outcomes should be computed center paribus in order to get the effect of a change in the variable (Verbeek, 2009).

The probit model is estimated by a likelihood function of the following shape for the sample

$$L(\beta) = \prod_{i=1}^{N} P\{y_i = 1|x_i: \beta\}^{yi} \, P\{y_i = 0|x_i; \beta\}^{1-yi} \qquad (5.8)$$

Where, as shown in equation 5.8 the likelihood is a function of the β. By transforming the function to a log likelihood function, the expression can be easily handled as the product gets transformed to a log of the sum instead and by exchanging the $P\{y_i = 1|x_i: \beta\}$ to$F(x_i'\beta)$. After this manipulation, the function can now be maximized with respect to β. By differentiating 5.8 with respect to β the first order condition is obtained.

$$\frac{\partial \log L\,(\beta)}{\partial \beta} = \sum_{i=i}^{N} \left[\frac{y_i - F(x_i'\beta)}{F(x_i'\beta)(1-F(x_i'\beta))} f(x_i'\beta)\right] x_i = 0 \qquad (5.9)$$

From the first order condition, the probability of being retired conditional on certain explanatory variables can be estimated. In the probit model the predicted frequency will be approximately the same as the actual frequency (Verbeek, 2009). The probit models in this study will be modelled in Stata using a so called dprobit that directly estimates the marginal effects of the explanatory variables.

6. Specifications in Modelling Retirement Decisions

In the sixth chapter the specifications of the retirement decision models used in this study, will be presented and discussed.

As stated in previous chapters, according to the theory of retirement decisions the individual will try to maximize his/her utility based on the information available. In previous research the generosity of the social security systems and pension systems were found to have an important role (Gruber & Wise, 1999). As an individual never knows how long he/she will live there is always some uncertainty even for the individual himself/herself when trying to estimate future income as retired. The classical model used when investigating retirement decisions contains earnings, private pensions and social security benefits (Mitchell & Fields, 1984). These variables were partly chosen due to limitations in the available data. These economical and institutional variables can only however explain the retirement decision up to a certain point as some individuals have other variables that are more important for the decision, such as inability to work due to for example illness (Brugaivaini, 2005). Studies on qualitative data suggest that the decision of retirement is based on a number of different variables and some of theses will be investigated below in our models.

The models aims at explaining the change in the job status for the respondents in the sample between 2004 (the first survey wave) and 2006/2007 (the second survey wave). The approach taken here will be to gradually add explanatory variables on the job status of the respondents in wave two, for the respondents who were working in wave one. The models will test whether the added explanatory variables improve the models ability to explain retirement decisions. Thus the models are step by step extended to include more variables to see if the inclusion of the self-perceived variables and the expectations of future income affect the outcome of the dependant variable i.e. the job status for the respondents in 2006/2007.

All respondents used in all models are classified as working in the first wave in 2004. By using different explanatory variables the effects on the respondents' job status in the second wave are estimated. The first model regresses retirement on age and gender to test if and how age and gender influence the retirement decision. Age is expected to have a large influence on the retirement decision. When retirement determinants are empirically investigated only data on men are often used. This is due to, among other things, that women's labour market participation is more complicated and also currently retiring women belong to a generation where the labour force participation was not as extensive as today and there are thus considerable cohort effects (Zamarro et al., 2007). Women are often as well assumed to retire at the same point in time as their husbands who often are older, thereby leading to that women often retire at a younger age than men.

Model 1 $$Job\ status\ in\ wave\ 2 = \alpha + \beta_1\ Age + \beta_2\ Gender + \varepsilon$$

The second model is extended to also include education. Education can be used as an indirect measure for low-skilled and high-skilled occupations, as a high-skilled job (normally) requires an employee that has a high education and thus also a high number of years of schooling. Education also indicates preferences between work and leisure, where a highly educated individual is assumed to have stronger preferences for work than an individual with lower education.

Model 2 $$Job\ status\ in\ wave\ 2 = \alpha + \beta_1\ Age + \beta_2\ Gender + \beta_3\ Education + \varepsilon$$

In the third model a health variable that controls for the individual's self-perceived health status, is added. The health status is divided into three different categories very good, the good and bad where good is used as the reference group. Very good health is expected to increase the retirement age whereas bad health is expected to have the opposite effect.

Model 3

$$Job\ status\ in\ wave\ 2 = \alpha + \beta_1\ Age + \beta_2\ Gender + \beta_3\ Education + \beta_4\ Health + \varepsilon$$

In the fourth model the health is removed but expected pension is added as the expected income from pension determines a large part of the future budget constraint for the individual. Depending on which cohort the individuals belong to the mix between the new and the old pension system is different. That the expected received pension benefit is a mix between the two systems makes it more difficult to predict the future pension. For the youngest individuals in the sample the prediction is further obstructed due to that the new system makes it more difficult to predict expected pension in advance (Sundén, 2004).

Model 4

$$Job\ status\ in\ wave\ 2 = \alpha + \beta_1\ Age + \beta_2\ Gender + \beta_3\ Education + \beta_4\ Expected\ Pension\ Income + \varepsilon$$

In the fifth model both the expected pension and the health variables are included. This tests whether the variables still maintain their separate effect or if some variable takes more than one factor into account so that the separate effects decrease.

$Model\ 5$

$$Job\ status\ in\ wave\ 2 = \alpha + \beta_1\ Age + \beta_2\ Gender + \beta_3\ Education + \beta_4\ Health + \beta_5 Expected\ Pension\ Income + \varepsilon$$

Defining the number of retired respondents as the respondents who receives the majority of their income from pensions resulted in a lot of missing values. When using the majority income definition, the more restrictive definition, only 23.85 percent of the respondents in the sampled were categorised as retired. This is a lower share of retired respondent than what seems plausible. Especially regarding the age-span the respondents are in the second wave, the low share of retired respondents probably stem from problems with the data. One way to evade this problem is to use an alternative definition of the classification of respondents as retired. The alternative definition, the second definition, of job status categorizes all individuals who do not receive labour market income as retired. When using this definition, the number of retired respondents in the sample rose to 42.03 percent. Models six to ten uses the same variables as model one to five respectively with exception of the dependant variable. Thus in models six to ten a different definition of the job statuses of the respondents in wave two is used. The definition of the job statuses used in model six to ten is that all respondents who did not report a labour market income is classified as retired.

7. Empirical Analysis

In this chapter the results of the different retirement decision models will be presented and compared. The models using the first definition of retired respondents, model one to five, are first presented and then followed by the models using the second definition of retirement status, models six to ten.

7.1 Empirical Results Using Retirement Definition 1

In table 7.1 the results of the empirical tests on the determinants of the retirement decisions are shown. Age dummies are used as the age variable, where respondents aged 59 in wave one in 2004 is used as the reference group. The age of the respondents will thus be two or three years older than in the first wave. Dummy variables are also used for the education variable where one to nine years of schooling is used as the reference group. The reference group, one to nine years of schooling, will be classified as a low educational attainment, so called basic level. The second education category includes ten to twelve years of schooling which corresponds to high school (gymnasium) as the highest attained degree and will thus be called secondary schooling. The third education category number three contains thirteen to fifteen years of schooling, thus corresponding to some years of university studies alternatively some years of vocational school after high school and will be called low university degree. The last education category, the so called high university degree, includes sixteen years of schooling or more and indicates a high educational attainment, for example a masters degree. The health variable is also measured by dummies, where a good health is used as the reference group. The second health category corresponds to a very good health and the third health category represents a bad health. The total expected pension is measured as the expected percent of the current salary that the public pension and the occupational pension will represent in the future. For both the health and the expected pension variable a dummy is used for the respondents who have not reported any answer on the specific question. By including a dummy for those respondents, their observations on other variables are kept and it thus helps to keep up the number of observations.

In table 7.1 the marginal coefficients (dF/dx) are reported and the standard errors are in the brackets. The coefficients show the change in marginal effect of an individual having retired during the time span between the two survey waves, due to the specific variable. The stars report the significance level, where three stars indicate that the coefficients are significant at one percent, two stars at five percent and one star at ten percent. If all observations in a variable have the same value the probit can not estimate the probability of the different outcomes and therefore drops the variable. This happened in models one to five for the age dummies 8 (age 66 in 2004) and 10 (age 68 in 2004), where the variables were dropped due to collinearity. The numbers of observations in models one to five are 238. In table 7.1 the Pseudo R^2 values of each model is also presented which describes the variability, how well the model explains the outcome.

Table 7.1Results Using the Restrictive Retirement Definition

Job status W2	**Model 1**	**Model 2**	**Model 3**	**Model 4**	**Model 5**
Variables					
Gender	-0.1003* (0.0556)	-0.1125** (0.0561)	-0.1196** (0.0568)	-0.1118 ** (0.0564)	-0.1069* (0.0574)
Age dummy 2 (60)	0.0800 (0.1159)	0.0717 (0.1148)	0.1377 (0.1290)	0.1022 (0.1194)	0.1546 (0.1311)
Age dummy 3 (61)	0.2668** (0.1270)	0.2600** (0.1274)	0.3366*** (0.1382)	0.2264 ** (0.1264)	0.3029** (0.1381)
Age dummy 4 (62)	0.2664** (0.1333)	0.2323* (0.1327)	0.3068** (0.1454)	0.2770** (0.1379)	0.3436** (0.1490)
Age dummy 5 (63)	0.6178*** (0.1146)	0.6392*** (0.1108)	0.6602*** (0.1153)	0.5961*** (0.1239)	0.6302*** (0.1257)
Age dummy 6 (64)	0.5526*** (0.1296)	0.5316*** (0.1339)	0.5876*** (0.1350)	0.5351*** (0.1356)	0.5974*** (0.1347)
Age dummy 7 (65)	0.4577*** (0.1568)	0.4192*** (0.1656)	0.5432*** (0.1615)	0.3263** (0 .1737)	0.4767*** (0.1788)
Age dummy 9 (67)	0.6923** (0.1634)	0.7285*** (0.1327)	0.7870*** (0.0875)	0.6741** (0 .1938)	0.7615*** (0.1291)
Education dummy 1 Secondary Schooling		0.1113 (0.0792)	0.1277* (0.0820)	0.1160 (0.0798)	0.1234 (0.0818)
Education dummy 2 Low university degree		0.1235 (0.1033)	0.1299 (0.1060)	0.1514 (0.1076)	0.1466 (0.1093)
Education dummy 3 High university degree		0.0431 (0.0875)	0.0902 (0.0945)	0.0693 (0.0910)	0.1010 (0.0962)
Very good health			-0.1748*** (0.0572)		-0.1711** (0.0601)
Bad health			0.2126 (0.1734)		0.1650 (0.1687)
Dummy no answer health			-0.1321* (0.0688)		-0.1380* (0.0698)
Total expected pension				-0.0011* (0.0006)	-0.0005 (0.0007)
Dummy don't know expected pension				-0.1864*** (0.0520)	-0.1595 ** (0.0550)
Pseudo R^2	0.1724	0.1834	0.2382	0.2492	0.2836

***1%significance level
**5%
* 10%

7.1.1 Model 1

In model one gender has a negative effect, on the ten percent significance level, on the probability given age of having retired during the time span meaning that women compared to men of the same age have a higher probability of having retired, i.e. retire earlier. This result is in line with what was expected, as women are often assumed to leave the labour market earlier for several reasons. The age dummies all have a positive effect on the retirement decision, indicating that you are more likely to have retired during the time span if you are older than the 59 years old in 2004, which was the reference group. The effect is not so large for the first dummy, corresponding to an age of 60 in 2004, and as well not significant, but increases with the age dummy groups. The first age dummy is never significant in any of the models, which may be due to that the difference by a year may not be large and the respondents aged 60 in the first wave are only 63 or 64 years old in the second wave and thus they are still younger than the "normal" retirement age of 65 in Sweden. From age 62 (dummy 3) to age 63 (dummy 4) in 2004 there is a substantial increase in the marginal effect from 27 percent to 62 percent, so more than twice the marginal effect, relative to the reference group of 59 years old. This effect is in line with theory since the individuals aged around 63 are around 66-67 years old in the second survey and have thus passed the "normal" retirement age in Sweden of 65 years old. Interesting is, however, that the probability of having retired is lower for the respondents aged 65 in the first wave (year 2004) than for the respondents age 64 or 63 years old in the first wave. All age dummy variables, except that for the 60-years old, are significant on at least the 5 percent level. The age and gender variables have shown to be robust in models one to five, independently of what other variables that are added. The Pseudo R^2 value is relatively low compared to the other models indicating that the model only explains parts of the retirement decisions for the respondents, which are not surprisingly as only two explanatory variables are used.

7.1.2 Model 2

In model 2 three dummies for education are added to control for educational differences and the effect of education on retirement decisions. The educational variables do not only measure differences in educational attainment but can also be seen as proxies for different occupations within the labour market. A high educational attainment indicates as well that the respondent is most likely to work in a high-skilled sector, and thus to have a less physically demanding and high paid job. An individual working in a high skilled sector may thereby have a higher income and thus also a higher expected pension which makes it possible for him/her to retire earlier. But a higher income also corresponds to a higher opportunity cost of being retired, thereby indicating likely postponing of retirement. An individual with lower education may have a less skilled work and often a more physically demanding job than the high skilled. Due to that the physically demanding aspect of the job puts constraints on health in the long run, a low skilled individual will probably retire earlier.

The educational variables show that more education has a positive effect on retirement meaning that a higher educational attainment increases the respondents' probability of having retired. This is however only up to a certain educational level, as the high university degree level of education shows a less positive effect than the lower university degree and secondary schooling variables. These results imply that having a medium high education increases the probability of having retired during the time span, probably due to that they have a somewhat higher income than the low educated and thereby can afford to retire earlier. Having a very high education also decreases the retirement age compared to that of the low educated but not by as much as for the medium educated respondents. The explanation of the result for the very highly educated respondents may be that they can afford to retire earlier but have a preference for working as they have invested a lot in attaining a high skill. None of the educational variables are however significant which means that any conclusions regarding their effects should be drawn with large caution. The educational variable is difficult to interpret in this study due to the significance problems caused by the small data sample. The model shows similar results regarding gender and age, indicating that these variables maintain their own separate effects. Gender has an even more significant and more negative effect in the second model than in the first. The pseudo R^2 value has increased a bit when education was added to the model indicating that education may be a determinant in retirement decisions.

7.1.3 Model 3

The third model also includes variables that measure self-perceived health status. The very good health variable is expected to have a negative effect on the probability of having retired which the negative coefficient also shows. The variable shows that a very good health decreases the probability (at the one percent significance level) of the respondents to have retired, in the age span between the first and second wave, by around 17 percent compared to those (of the same age, gender and educational level) who had only good health in the first survey wave in 2004. The variable for bad health shows as expected that a bad health has a positive effect on the probability of having retired during the period between the first and second survey waves. This variable is however not significant, thus no conclusions can be drawn on the effect of the bad health variable. The dummy controlling for the respondents who not have answered the health question in the survey wave 1 in 2004 is negative and significant at the ten percent level. The respondents who have not answered the health variable might mentally be farther away from retirement as individuals who are closer to retirement may consider their health status more. The gender and educational effects are about the same in this model as in the previous. The only substantial change in the effects is for the secondary school variable, education dummy 1, which has a slightly larger coefficient but again is not significant. The age variables coefficients show larger values in this model with about the same level of significance. This might be due to that when health also is included in the model the age effects become stronger. When the health variables are added the model is improved, which can be seen on the higher Pseudo R^2 value.

7.1.4 Model 4

In the fourth model no health variables are included, but instead total expected pension, measured as total expected pension as percent of the current wage, is added. In the total expected pension variable both the expected public pension and the occupational pension are included. The expected marginal effect of the total expected pension is difficult to determine as it can work in opposite directions and also be a result of preferences in economical behaviour. The total expected pension can have a positive effect on the probability of having retired since a larger expected pension means that the differences in income between working and being retired will be less. A large expected pension can implies also that the respondents can afford to retire. But a large expected pension can also be a result of economical preferences meaning that a respondent who has a preference for a large pension may choose to work longer in order to achieve that. The empirical results show that in this sample the total expected pension has a negative effect on the probability of having retired during the time-span between the two survey waves. A ten percent increase in the expected pension income in relation to the wage, is found to decrease the probability of being retired in the second wave about one percent. The dummy controlling for the respondents who have not answered the questions in the survey, either due to that they don’t know or that the refuse to answer, is found to affect the probability of being retired negatively as well, at the one percent significance level. The negative coefficient of the dummy variable for the respondents who answered that they do not know or refused to answer may be due to if retirement is mentally far away then one might not have a clear expectation of the size of the future pension. In general, the gender, age and the educational variables are unchanged regarding both the size of their coefficients and their significance level. Including the expected pension variable improves the model more than when including the health variable.

7.1.5 Model 5

In the fifth model both health and total expected pension are included. Very good health has about the same negative marginal effect, -17 percent, as in model three but is slightly less significant. Bad health has a smaller positive effect on the probability of having retired during the time span than in model three, but the effect is not significant. Thus, when both expected pension and bad health is included in the model, the increase in probability of having retired between the first and second wave is not as large as when only bad health is included. Total expected pension maintains its negative effect on the retirement probability but is in this model a bit smaller and not significant. So when including both health and expected pension, the effect of the expected pension decreases a bit even if the effect is still negative. The dummy for the respondents who have not reported an expected pension income is negative still but smaller than in model four and less significant. The gender effect is about the same as in previous models. The age and education coefficients are a bit larger in this model compare to the age effects in model one, two and four but education is still not significant. According to the Pseudo R^2 value this model explains the retirement decision better than of all the previous models, indicating that all these variables may be determinants in the retirement decision.

7.2 Empirical Results Using Retirement Decision 2

In the previous models the reported pension incomes of the respondents were few due to a lot of missing data. This resulted in a low sample of respondents who reported top have retired between the first and second wave. The reason for the low number of respondents on pension income was however not that the respondents were working, since no labour market income was reported. It is not unlikely that all respondents who have not reported a labour market income are retired. In the following models, the specific assumption that all those without reported labour market income are retired is therefore made, which leads to an increase in the number of observations to 433 compared to 238 in the previous models. Number of observations in the sample decreased from 444 to 433 due to missing values on the gender variable. In the models below there are consequently strong assumptions made regarding the categorization of the respondent's labour market status in wave two in 2006/2007. Despite this being a strong assumption it is however not an unlikely one. If the job status is equal to one the respondent is retired in the second wave and if the job status is equal to 0 the respondent is still working. Table 7.2 reports the results obtained for model six to model ten, where the second definition of retired individual in the data sample is used as the dependent variable.

Table 7.2 Results Using Alternative Retirement Definition

Job status W2	**Model 6**	**Model 7**	**Model 8**	**Model 9**	**Model 10**
Variables					
Gender	-0.0484 (0.0498)	-0.0497 (0.0502)	-0.0673 (0.0513)	-0.0776 (0.0520)	-0.0863 (0.0528)
Age dummy 2 (60)	0.1206 (0.0744)	0.1241 (0.0748)	0.1447* (0.0744)	0.1462* (0.0753)	0.1588** (0.0748)
Age dummy 3 (61)	0.2141*** (0.0680)	0.2191*** (0.0681)	0.2424 *** (0.0667)	0.2128*** (0.0698)	0.2289*** 0.0687
Age dummy 4 (62)	0.2169*** (0.0694)	0.2201*** (0.0699)	0.2370*** (0.0686)	0.2572*** (0.0673)	0.2658*** (0.0664)
Age dummy 5 (63)	0.3968*** (0.0459)	0.3940*** (0.0467)	0.3907*** (0.0469)	0.3809*** (0.0493)	0.3779*** (0.0494)
Age dummy 6 (64)	0.3535*** (0.0532)	0.3554*** (0.0529)	0.3587*** (0.0523)	0.3503*** (0.0542)	0.3547*** (0.0531)
Age dummy 7 (65)	0.3673*** (0.0507)	0.3715*** (0.0503)	0.3882*** (0.0467)	0.3204*** (0.0621)	0.3426*** (0.0576)
Age dummy 9 (67)	0.3667 ** (0.0707)	0.3613** (0.0758)	0.3887* (0.0490)	(0.3042) (0.1208)	0.3502** (0.0839)
Education dummy 2 Secondary schooling		-0.0402 (0.0628)	-0.0126 (0.0638)	0.0103 (0.0645)	0.0227 (0.0649)
Education dummy 3 Lower university degree		0.0267 (0.0744)	0.0399 (0.0754)	0.0520 (0.0760)	0.0600 (0.0766)
Education dummy 4 Higher university degree		-0.1011 (0.0728)	-0.0415 (0.0742)	-0.0373 (0.0745)	0.0030 (0.0748)
Very good health			-0.2322*** (0.0775)		-0.1814 (0.0821)
Bad health			0.2012* (0.1004)		0.1740 (0.1093)
Dummy no answer health			-0.0198 (0.0678)		-0.0061 (0.0697)
Total expected pension				-0.0030***	-0.0025***

				(0.0006)	(0.0007)
Dummy don't know expected pension				-0.3372*** (0.0644)	-0.3188*** (0.0661)
Pseudo R2	0.0943	0.0990	0.1310	0.1819	0.1988

***1%significance level
**5% significance level
* 10% significance level

7.2.1 Model 6

In the sixth model gender and age in year 2004 are regressed on the job status of the respondents in wave 2 in 2006/2007. The gender coefficient is relatively small indicating that there is a lower probability by 4.84 percent of having retired, all things the same, if you are male than if you are female in the period between the first and the second survey wave. This effect is however not significant so no conclusions can be drawn with certainty. The model also shows a positive and significant age effect, except for the age 60 in year 2004, meaning that the older you are the greater probability are there of having retired between the first and the second wave. There is a substantial difference in the magnitude of the age effect from age 63 to age 64 in 2004. According to the model there is a 22 percent larger probability of having retired if the respondent was 63 years old in the first survey wave in year 2004 than if 59 years old in 2004 and for a respondent who was 64 in 2004 there is a 40 percent larger marginal effect, thus almost twice the marginal effect. The results for this model are roughly the same as for model one with the exception that the gender variable was significant in model one. The Pseudo R^2 value was, however, much smaller in this model than in model one indicating that model one better explains the outcome of the job status in wave two. This difference in the Pseudo R^2 values could be due to the limited data sample.

7.2.2 Model 7

The seventh model is extended to also include variables controlling for education. The educational variable for the secondary schooling (10-12 years of schooling) indicate a decrease in the probability of having retired between the first and second wave. This result is different compare to the result obtained in model two, where the secondary schooling increase the probability of having retired by 11 percent. The effect of the lower university degree increases the marginal effect by 3 percent, which is less than in model two where the marginal effect was 12 percent. The higher university degree decreases the marginal effect by 10 percent whilst in the model two the marginal effect of high university degree is positive by

4 percent. So in this model high university degree and the secondary schooling decreases the probability of having retired between 2004 and 2006/2007 whilst in model two the same variables have the opposite marginal effects. Neither of these educational variables are however significant so the results should be concluded with caution. The results for the gender and age effects in this model are about the same as in the sixth model. The Pseudo R^2 value is higher for this model than for the previous (model six) but it is still lower than the pseudo R^2 for the second model. Model two thus better explain the change in job statuses for the respondents between the first and second wave than this seventh model despite that the seventh model has a larger sample.

7.2.3 Model 8

The eighth model is extended to also include health. The variable for very good health showed a negative and significant marginal effect by 17 percent meaning that a very good health decreases the probability of having retired between the first and second wave by 17 percent, compared to a good health. The model also showed that a very bad health increase the marginal effect of having retired in the time interval between the first and the second survey wave by 20 percent, at a ten percent significance level. These health effects are about the same as in model three, but more significant. The age variables have increased slightly, which also happened when the health variables were added in model three. The gender and educational variables showed roughly the same as in the previous models with the exception of the secondary schooling variable that showed a smaller effect, in contrast to model two to five where the educational variables increased when health was added to the model. Adding health improved the model substantially, when regarding the Pseudo R^2 value, but this model perform worse than model three where the more restrictive definition of retired respondents were used.

7.2.4 Model 9

The total expected pension is now added to the model and the health variables are taken away for now. The total expected pension is found to have a negative and highly significant marginal effect by -0.3 percent on the probability of having retired between wave one and two, meaning that if the total expected pension in relation to the current wage increase by ten percent the probability of having retired in the time-span decrease by 3 percent. The increase in significance between this model and model four, where the more restrictive definition of retired respondents in the sample was used, probably depends on the few observations in the sample in model one to five. The dummy that controls for the respondents who have not answered the expected pension question is highly significant and negative. A possible explanation for this can be that for respondents that mentally are far from retirement does not have a clear estimation of their expected pension. The gender and age effects are about the same as in previous models where the increase in probability of having retired between the first and second wave lies around 21-25 percent for the age 61-62 in the first wave (2004) and around 35 percent for ages 63-65 in the first wave. The education coefficient for secondary school is suddenly positive by 1.3 percent whereas the other educational variables coefficients are roughly the same. The differences in the educational marginal effects are probably due to the small data sample. Adding total expected pension to the model improves it significantly, according to the Pseudo R^2 value, even more than when health was added (in model eight). The same model but with the more restrictive definition of job status in wave two (model four) performed better despite that this model has a larger sample.

7.2.5 Model 10

In the tenth model all variables are included, both health and total expected pension. The gender coefficient is larger than before by -8.63 percent but still not significant. The age variables marginal effects have increased some but are still around 25 percent for the respondents aged 61-62 in 2004 and around 35 percent for the respondents aged 63-67 in 2004. The same result could be seen in model five where the age effects also increased when all variables where included. The educational variables are all positive but none are significant. The health variables marginal effects are -18 percent for very god health and 17 percent for a bad health, which is less than in previous model including health but they are not significant. As the health variables are not significant no specific result could be drawn but it could possibly be that the health effects decrease when the total expected pension is included in the model as well. The health effects in this model are, however, very similar to the health effects in model five. The coefficient for the total expected pension is still significant but slightly smaller, in contrast to model five where the coefficients were insignificant but small. So inclusion of both health and expected pension decreases the marginal effect of the expected pension variable. Adding all variables improves the model and makes it the best performing of all models using the definition of job status where all respondents are retired if they do not report a labour market income. But this model still performs worse than the model with all variables from the first definition of job status in the second wave, model five.

To sum up the results, the gender variable is negative in all models but only significant in model one to five. This implies that if you are male (with certain characteristics) you have a lower probability of having retired during the time-span between wave 1 and wave 2 than if you are female. All age variables show that, other things the same, the older you are the greater the probability that you will have retired. The effect of education on retirement is difficult to determine as it shows contrasting results in the different models where different definitions of being retired are used. A very good health, everything the same, makes the respondents postpone retirement as compared to those with good health whilst bad health results in earlier retirement than if the health was good. The total expected pension decreases the probability of having retired in the time-span between 2004 and 2006/2007. The variable found to be most determinant for retirement is a high age.

8. Discussion and Conclusion

In this chapter the results will be discussed and analysed. Here possible explanations for the results will also be suggested. The chapter ends with suggestions for further research.

This study aimed at modelling the retirement determinant for Swedish residents by using a probit model on qualitative micro data. The models investigated the probability of change in job status between the first survey wave in 2004 and the second survey wave in 2006/2007 by looking at some background variables measured in the first survey wave. When using a more restrictive definition of the respondents' job status in the second wave (in model one to five) the sample became very small so an alternative definition, that increased the sample, was used for model six to ten. The results from the models using the different definitions of the retirement classification show mostly similar results. The more restrictive definition of the retired classifications limited the sample substantially and results that only show significance in the model one to five, where the restrictive definition was used, should therefore be interpreted with caution.

The results of the empirical tests showed that women in the sample have a larger probability of having retired between the first and second wave than men, all things the same. This implies that women retire earlier than men, which also is consistent with existing theory that couples coordinate their retirement decisions but due to the wife often being younger than the husband this results in a lower retirement age for women. The difference in probability is around ten percent but is only significant when using the restrictive definition of retirement, the definition that generates a smaller sample.

All models show, with high significance level, that the older you are the higher is the probability of having retired within the time-span. The empirical results also show that in all models there is a large difference in the marginal probability of having retired between wave one and wave two between the age 62 in the first wave and the age 63 in the first wave. As the age variables is the age of the respondents in the first survey wave in 2004, the large increase in the probability of having retired during the time-span between the two waves are around the ages 66-68 for the respondents. This result indicates that it is around these ages that most Swedes retire. These results are in line with the "normal" retirement age in Sweden which is approximately around 65 years old. In the new pension system in Sweden the public pension can be collected from age 61 but the system is designed to give incentives for continuing to work. As the cohorts in this sample receive a mix of the old and the new pension system the incentives for continuing to work might be stronger in the future for the cohorts who will only receive pension from the new system, thus (hopefully) leading to an increase in the effective retirement age.

The effects of different educational attainments have showed contradicting results in some of the models. In model two to model five all educational attainment levels above the basic level of nine years of schooling showed a positive marginal effect on the probability of having retired between wave one in 2004 and wave 2 in 2006/2007, compared to the reference level basic education, thereby indicating that education decrease the retirement age. Whilst model seven to nine showed that having a secondary schooling level or having a high university degree have an increasing effect on retirement age. All educational variables suffer from low significance, which makes them difficult to interpret. In model three the secondary educational level is significant, indicating that there is a positive effect on the probability of having retired in the time-span retired. This might be a result of the income effect, i.e. that the respondents with ten to twelve years of schooling can afford to retire earlier than those respondent with only nine or less years of schooling.

The retirement probability is as expected negative for the respondents who report a very good health and positive for the one who report a bad health. A good health decreases the probability by 17 to 23 percent in the models, indicating that health is an important variable in determining retirement. This also implies that only using economical measures when trying to model retirement decisions might be misleading. It should be noted that the health variable used here is self-perceived, which could lead to that the respondents underestimate their health conditions to justify their retirement.

The effect of the total expected pension is in theory ambiguous as it can work in opposite directions and as well be a result of economic preferences. The total expected pension can have a positive effect on the probability of having retired since a larger expected pension means that the differences in income between working and being retired will be less. But a large expected pension can also be a result of economical preferences meaning that a respondent who has a preference for a large pension may choose to work longer in order to achieve that. In the empirical test in this study the total expected pension has a negative impact on the probability of having retired in the time-span meaning that a higher future expected income decreases the probability of being retired, thus increasing the retirement age. By estimating the total expected pensions' marginal effects on retirement decision, it can not be concluded whether the effect is due to that respondents who is planning to postpone their retirement have a high pension or if respondents with a high pension chooses to work longer. What can be concluded from the empirical test of the expected pension variable is that expectations affect the retirement decision and that there are different types of economical behaviour. To conclude, there is a very distinct age pattern in retirement but despite that the retirement decisions are still substantially affected and modified by both economical (in this study represented by total expected pension) and non-economical factors (here represented by health status). The study also shows that gender and education also play an important role.

An interesting feature of this study is that even in model one to five, where the more restrictive definition of retired respondents are used and thus a smaller sample, a lot of significant result are generated that remains significant when a larger sample is used. The study also shows the importance of including both economical and subjective variables when modelling retirement decision determinants. Further research on these variables on a larger sample would be very interesting and also use interaction variables to control for different effects for men and women.

References

Published references

Brugiavini, A., (2005) “Work and Retirement“, Chapter 5, Health, Ageing and Retirement in Europe - First Results from SHARE, Mannheim: MEA

Brugiavini, A., (2008) “Work and Retirement“, Chapter 6
Health, Ageing and Retirement in Europe (2004-2007), Mannheim: MEA

Börsch-Supa, A., Hank, H., Jürges, H. and Schröder, M. (2009), ”Introduction: empirical research on health, ageing and retirement in Europe”, *Journal of European Social Policy,* Vol. 19(4):293-300

Blöndal & Scarpetta (1999) “The Retirement Decisions in OECD Countries” OECD Economics Department Working Papers, No. 202, OECD Publishing.

Dwyer, D.S, Mitchell, O.S., (1999) “Health Problems as Determinants of Retirement: Are Self-rated Measured Endogenous?”, *Journal of Health Economics*, Vol. 18

Fields, G.S., Mitchell, O.S., (1984) “The Economic Behaviour of Retirement”, *Journal of Labour Economics,* Vol 2, no 1

Fields, G.S., Mitchell, O.S., (1983) “Economic Determinants of Optimal Retirement Age: An Empirical Investigation” NBER Working Paper no 876, April 1982; revisited June 1983

Gruber & Wise (1999), ”Social Security and Retirement around the World”, Chicago: University of Chicago Press

Johnson, R.W. (2009), “Family, Public Policy, and Retirement Decisions”, *Research on Ageing*, Vol. 31, No 2

Karlstrom, A., Palme, M., Svensson, I. (2004) ”A Dynamic Programming Approach to Model the Retirement Behaviour of Blue-Collar Workers in Sweden”, *Journal of Applied Econometrics*, Vol 19, No 6

Palmer, E., (2002) “Swedish Pension Reform: Its Past and Its future” In Feldstein, M., & Siebert, H., Eds. *Social Security Pension Reform in Europe*, Chicago: The University of Chicago Press: 171-210

Sundén, A., (2004) ”The Future of Retirement in Sweden” Pension Research Council Working Paper, Warton School of the University of Pennsylvania

Verbeek, M., (2009) “A Guide to Modern Econometrics” John Wiley & Sons, Ltd

Zamarro, G., Meijer, E., Fernandes, M., (2007)”Labour Force Participation” Chapter 2.4
Health, Ageing and Retirement in Europe (2004-2007), Mannheim: MEA

Electronic references

Pensionssystemet http://www.pensionssystemet.se/ (2010-04-16)
http://www.pensionssystemet.se/allmanna-pensionssystemet

Pensionsmyndigheten
(http://www.pensionsmyndigheten.se/PensionFranFleraHall.html) (2010-04-21)
http://www.pensionsmyndigheten.se/OmPensionsonssystemet.html (2010-05-12)

Blöndal &Scarpetta (1998)
http://www.oecdilibrary.org/oecd/content/workingpaper/565174210530