SUBJECTIVE SURVIVAL PROBABILITIES AND THEIR ROLE IN LABOUR SUPPLY DECISIONS

Tesis doctoral presentada en la FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES de la UNIVERSIDAD DE NAVARRA para la obtención del grado de Doctor en Economía Aplicada por JORGE ESTEBAN LEÓN ARELLANO, bajo la dirección de la profesora DRA. DÑA. JUNCAL CUÑADO EIZAGUIRRE.

PAMPLONA, A 4 DE DICIEMBRE DE 2012

AGRADECIMIENTOS

Agradezco a mi directora de Tesis, la profesora Dra. Dña. Juncal Cuñado Eizaguirre, todo su apoyo a lo largo de estos años. La consecución de este proyecto no habría sido posible sin su constante orientación y supervisión desinteresada.

Hago extensivo mi agradecimiento a los profesores de la Facultad de Ciencias Económicas y Empresariales de la Universidad de Navarra, y a los del Centro de Estudios Monetarios y Financieros del Banco de España (CEMFI), que han contribuido a mi formación académica.

Agradezco igualmente a la consultora económica Frontier Economics por su apoyo para compaginar mis actividades profesionales con la realización de esta Tesis.

Esta Tesis está dedicada a mis padres, por inculcarme el valor del esfuerzo y trabajo diario; a mis hermanos, por enseñarme el valor de la amistad; y a Rosana, por mostrarme la belleza de la vida.

INDEX

ABBREVIATIONS

ADL	Activities of Daily Living
AHEAD	Asset and Health Dynamics among the Oldest-Old survey
ALSA	Australian Longitudinal Study of Ageing
BNHI	Bureau of National Health Insurance of Taiwan
ELSA	English Longitudinal Study on Ageing
GALI	Global Activity Limitation Indicator
HMD	Human Mortality Database
HRS	Health and Retirement Study
IADL	Instrumental Activities of Daily Living
ISCED	International Standard Classification of Education
IV	Instrumental Variables
LCM	Life Cycle Model
MEA	Mannheim Research Institute for the Economics of Aging
OLS	Ordinary Least Squares
SHARE	Survey on Health, Ageing and Retirement in Europe
SSP	Subjective Survival Probabilities
UK	United Kingdom
US	United States

GENERAL INTRODUCTION

As economic agents we take economic decisions every day. Our decisions can go from trivial choices, such as deciding whether to spend money on a product or saving it, to more vital choices, such as buying an annuity or deciding when to retire.

Basic economic theory often assumes that decision-makers have all the information needed to take optimal decisions. However, this might not always be true in real life. Most of our choices as economic agents are made under uncertainty. This does not mean that economic agents only consider those aspects that are perfectly observable when taking their decisions. We actually incorporate the uncertainty element into our decision making process through expectations. For example, even though we are not certain about the future behaviour of the interest rate, we could form our own expectations about this and decide whether a mortgage based on variable interest rate is preferred to a fixed interest one.

The role of uncertainty, and therefore of expectations, is especially important in late life decisions. This is true because the scope for revision in case we take wrong economic decisions is limited. Older individuals are faced with one of the most important decisions an economic agent has to take: when to retire. Deciding when to retire is deciding when to stop working, and therefore, stop receiving labour income and start collecting pension income, which is usually lower. In order to overcome this decrease in income, economic agents normally accumulate wealth during working years and spend it during retirement.

Hence, an economic agent is faced with a difficult decision. Let us assume that someone is deciding whether to delay retirement by one year or not. If she delays retirement she could accumulate more wealth to finance retirement years. However, if she dies shortly after retirement, she will have accumulated too much wealth and it would have been best not to delay retirement. On the other hand, if the economic agent does not delay retirement, she will not accumulate more wealth. Yet, if she happens to live many years after retirement, the accumulated wealth may not be enough to finance consumption in the late years.

How can an economic agent take a rational retirement decision given this uncertainty? As mentioned above, we incorporate the uncertainty element into our decision making process through expectations. In this case, economic agents are expected to shape longevity expectations to take rational decisions. A rational decision in this context would mean that if an individual expects to live exceptionally long, she will have to retire at a later age than individuals who expect to die early. This is because she will need more wealth to finance more years of retirement.

This idea is a direct consequence of the Life Cycle Model (LCM) developed by Modigliani and Brumberg (1954), and Friedman (1957). According to this theory, a rational economic agent would take a retirement decision so that she maximizes *lifetime* utility. However, because there is uncertainty about when someone will die, longevity expectations should play a significant role on deciding when to stop working.

The fact that longevity expectations are taken into account when taking retirement decisions seems reasonable in theory. But is this truly the case in real life? Do people think about mortality when deciding when to stop working and do they take coherent retirement choices? In order to answer to these questions one should take one step backwards and firstly understand and analyse how economic agents form and update their expectations about longevity.

The analysis performed in this Thesis is based on the data drawn from the "Survey of Health, Ageing and Retirement in Europe" (SHARE). This is one of the most important novelties of the Thesis. SHARE is a multidisciplinary and cross-national panel database of micro data on health, socio-economic status and social and family networks of more than 55,000 individuals from 20 European countries aged 50 or over. SHARE is coordinated centrally at the Mannheim Research Institute for the Economics of Aging (MEA) and it is harmonized with the HRS and the ELSA surveys.

Currently, the database comprises two waves and an additional survey which concentrates on people's life histories. The next wave is being collected and is expected to be released at the beginning of 2013. The first wave was released in 2004 and it included information from 31,115 individuals from eleven European countries (Austria, Belgium, Denmark, France, Germany, Greece, Italy, Netherlands, Spain, Sweden, and Switzerland). Further data was collected in 2005-06 in Israel. For the second wave, which was released in 2008 and includes 33,281 respondents, two "new" EU member states - the Czech Republic and Poland - as well as Ireland joined SHARE. The latest wave to be released in 2013 will also include data from Estonia, Hungary, Portugal and Slovenia.

Since its release in 2004 SHARE has become increasingly popular among social scientists and researchers and particularly among labour economists. SHARE has been the data source for many books, book chapters, journal articles and working papers.

For the analysis performed we make use of the individuals' subjective assessment of their mortality risk drawn from SHARE. In particular, respondents are asked about their chances to live to age T or more, where the target age, T, is chosen conditional on the respondent's age. We concentrate on individuals aged between 45 and 65, so that they are asked about a common target age: 75 years old. Additionally, SHARE provides substantial information on socio-demographic characteristics, health variables, and economic and labour conditions at an individual level along waves. This information is used in each of the three Chapters that comprise this Thesis.

This Thesis analyses people's longevity expectations and their role in labour supply decisions in three different Chapters. It seeks two concrete objectives. Firstly, in the first and second Chapter we explore individuals' longevity expectations – more commonly known in economic literature as Subjective Survival Probabilities (SSPs) – in order to determine if we could confidently use them as main input in a retirement model. Secondly, once SSPs are validated, in the third Chapter we to test if they play a role in labour supply decisions in the same fashion the LCM suggests.

For the first objective we analyse whether SSPs satisfy three specific properties that expectations should satisfy according to Hurd and McGarry (1997): (i) individuals understand questions about probabilities and accurately report their beliefs about the likelihood of future events; (ii) individuals adjust their reported probabilities in response to new information; and (iii) the reported probabilities predict outcome.

Chapter 1 examines whether individuals understand the question about longevity expectation and that they accurately report their beliefs (first condition). In particular, we analyse if SSPs are consistent across waves, average close to objective probabilities; are consistent with mortality and epidemiological data and with economic literature. For this matter, we first introduce the concept of SSPs. We present SHARE's question on SSPs and we explore the respondents' answers for both waves. We also conduct a comparison of Subjective versus Objective Survival Probabilities using national life tables in order to determine if people's expectations are coherent. Finally, we focus on understanding what the determinants of SSPs are and if they are consistent with mortality and epidemiological data. Based on the approach followed by Hurd and McGarry (1995, 1997), Hurd *et al.* (2001) and Liu *et al.* (2007) among others, a cross sectional econometric analysis for each of the two waves of SHARE (2004 and 2008) is performed. We define a relationship where SSP is a function of socio-demographic characteristics, health condition variables, health behaviour variables, self-reported health, and parental mortality experience.

Chapter 2 studies how SSPs evolve over time as new information, whether this is medical, socio-economical, or personal, is revealed to the respondent (second condition). We examine if SSPs drawn from the SHARE database update in a systematic and reasonable manner with the arrival of new information. For this purpose we perform an econometric analysis of the determinants of the changes in SSPs from the first to the second wave of SHARE at an individual level. Our results are compared with those of previous studies that use different national databases such as Hurd and McGarry (1997), Hurd *et al.* (2001), Smith *et al.* (2001), Benítez-Silva *et al.* (2008), Liu *et al.* (2007), and Steffen (2009).

Additionally, the second chapter examines whether SSPs predict actual mortality (third condition). Following Hurd and McGarry (1997), Van Doorn and Kasl (1998), Smith *et al.* (2001), Hurd *et al.* (2001) and Siegel *et al.* (2003) we use a probit specification to model observed mortality between wave 1 and wave 2 as a function of SSP in wave 1 and test if SSP adds any predicting power when forecasting decease. By controlling for observable risk factors we can find if the reported SSP incorporates reliable information on survival chances that would otherwise not be observed.

The second objective of the Thesis is explored in Chapter 3. We test whether SSPs play a significant role in the labour supply decisions of economic agents. Based on the LCM it is expected to observe that lower SSP is associated with a higher probability of retirement. We propose a probit specification in order to model the probability of an individual being retired in wave 2 (who was in the labour force in wave 1) as a function of SSP and a set of independent variables. This approach has been used by several authors. They include Hurd *et al.* (2004), Delavande *et al.* (2006), Bloom *et al.* (2007), Bloom *et al.* (2006), and O`Donnell *et al.* (2008).

A rather important aspect that is dealt in the Thesis is the treatment of SSP when modelling entering into retirement. Authors such as Stern (1989), Delavande *et al.* (2006), Bloom *et al.* (2006), Bloom *et al.* (2007), and O`Donnell *et al.* (2008), among others, argue that SSP could suffer from endogeneity. This endogeneity could arise due to two factors: SSP might suffer from classical measurement error, and due to, what in the economic literature is known as "justification bias". Based on economic literature we use parental mortality and smoking behaviour as instruments to correct the possible endogeneity of SSP.

Furthermore, several sensitivity scenarios are constructed in order to test the consistency of the results. In particular, we include different definition of labour supply that other authors have used in the past such as Hurd *et al.* (2002), Bloom *et al.* (2006), Fischer and Sousa-Poza (2006), Bloom *et al.* (2007), O'Donnell *et al.* (2008), and Hospido and Zamarro (2012). Additionally, because it might be the case that Europeans choose alternative retirement pathways such as decreasing the number of hours worked gradually, instead of taking a one-off retirement decision, we analyse the effect of SSPs on the change in the number of hours worked between wave 1 and wave 2.

Even though SSPs have been explored using other surveys, mainly the Health and Retirement Study (HRS) in the US and the English Longitudinal Study on Ageing (ELSA) in the UK, they have not been analysed in detail using SHARE data. There are not too many studies that rely on SHARE's expected longevity data. We can highlight the work of Balia (2007), Delavande and Rohwedeer (2008), Hurd *et al.* (2008) and Post and Hanewald (2011). Similarly, to our knowledge, the role of SSPs on labour supply decisions for SHARE respondents has not been explored in detail. This Thesis intends to fill this gap and to provide a comparison of the European results with those obtained by other authors using other national databases.

Additionally, this Thesis also intends to contribute to the public debate on the retirement behaviour of Europeans and the sustainability of the pension system. We believe that understanding how expected longevity affects labour supply decisions is a crucial topic for policymakers. We are currently living in an environment of financial crisis, high unemployment rates, low fertility rates and increasing life expectancy. This has increased the financial pressure on Social Security systems across Europe and has increased the doubts about the sustainability of the whole system. Policymakers should be interested in understanding how to incentivise people to work longer in a situation characterized by increasing longevity.

Overall, we have found that SSPs do satisfy the three specific properties that expectations should satisfy according to Hurd and McGarry (1997) and therefore they can be used confidently to estimate models of decision-making under uncertainty. Specifically, using cross-sectional analysis it is found that SSPs in SHARE behave reasonably well. They average close to actuarial probabilities and they covary with other variables in the same way actual outcomes vary with the variables. Moreover, our results are in line with economic literature based on alternative data. We have found that the evolution of SSPs is coherent with epidemiological evidence and with previous studies done based on other surveys. We conclude that individuals adjust their reported expected longevity in response to new information and that this adjustment is consistent. Finally, we have found that SSPs predict mortality.

We have also found that expectations of longevity do play a significant role when taking retirement decisions only in the case of females. In particular, it is found that females who expect to live longer have a lower probability of retiring. This finding is consistent with the LCM. This suggests that females tend to act as rational economic agents when taking late life decisions such as continue working or not. We found that this might not be true for males as our results suggest that males do not take into account longevity expectations when deciding when to retire. This interesting result could be related to the fact that males and females have different retirement length expectations. In particular, wives typically anticipate a longer retirement period than their husbands, and part of that retirement period is likely to be as widow. For this reason, females tend to be more risk averse to fall short of savings and wealth accumulation.

Even though this Thesis concentrates on the role of SSP on labour supply decisions, we believe that these types of studies, where SSPs are incorporated as main input, are quite relevant in other fields of Economics. In particular, areas such as family economics, health economics, behavioural economics, and macroeconomics, among others, could benefit from the use of SSPs.

CHAPTER 1.
SUBJECTIVE SURVIVAL PROBABILITIES AND THEIR DETERMINANTS

1. INTRODUCTION

Even though expectations play a very important role in models of decision-making under uncertainty, economists have been deeply sceptical about using data on expectations when modelling. Normally, due to the lack of information, it has been necessary to make assumptions about agents' expectations. For example, macroeconomic models often assume rational expectations achieved by a trial-and-error mechanism where agents adjust their initial beliefs if they are incorrect. On the other hand microeconomic estimations often use observable population probabilities. It is true that economists are often much more interested in what people do rather than what people think.

Nevertheless, examining the history of economic thinking about expectations' data we find that the general scepticism is only based on a narrow foundation. From 1940s the Federal Reserve Board stated to fund annual surveys on consumer finance that included expectation questions such as:

"How about a year from now - do you think you people will be making more money or less money than you are now, or what do you expect?"[1]

Economist at that time criticized the usefulness of the responses to such questions and argued that their predictive power was very limited. By mid-1960s the general view was firmly negative.

However, since mid-1990s this view started to change and economists increasingly started to use data on probabilistic expectations of significant personal events. For example, economists' attention turned to expectations on job loss, crime, earnings, bequests, social security, personal choices, and most importantly for this Thesis, expectations on longevity.

[1] Dominitz and Manski (1997a).

Why did economists suddenly found that expectations and subjective probabilities were now of economic interest? Hamermesh (2004) justified the sudden use of expectations partly because economists are seeking things to do with the large amounts of data that they have suddenly discovered—the Mont Everest phenomenon. If it's there, we must climb it. Since mid-1990s large scale surveys where popularized in the US, together with the development of more powerful statistical packages facilitated the analysis of expectations.

But Hamermesh (2004) also points out to a second reason why economist suddenly found expectations to be interested. The author argues that results are important inputs to economists' basic research and are not likely to be considered by other social scientists.

A clear example of this is the role of longevity expectations in the Life Cycle Model (LCM). According to the LCM there is a well-defined link between the consumption plans of an individual and her income and income expectations as she passes from childhood, through the work participating years, into retirement and eventual decease. The hypothesis was developed by Modigliani and Brumberg (1954) in line with the Permanent Income Hypothesis developed by Friedman (1957) which states in its simplest form, that the choices made by consumers regarding their consumption patterns are determined not by current income but by their longer-term income expectations.

The LCM is based on the assumption that a rational economic agent maximizes lifetime utility. This lifetime utility is given by consumption and leisure. However, there is, obviously uncertainty on when someone will die, and here is where longevity expectations could play a significant role.

As noted by Hamermesh (2004) expectations are inherently subjective as they deal with events on which objective measures cannot yet be provided. Subjective expectations are central to testing basic theories of economics and the LCM provides an excellent example.

For instance, when we write down the lifetime utility maximization of the economic agent suggested above, we derive paths of consumption, labour supply, wealth and other outcomes from the dynamic optimization problem defined over a horizon, T. But how long is T? This is a very important element of the solution.

One direct consequence of the LCM is that individuals who expect to live exceptionally longer will retire at a later age than individuals who expect to die early. This is because they will need more wealth to finance more years of retirement, as income from a pension is substantially lower than labour income. Wolfe (1983) demonstrated this hypothesis using an actuarial, an economic and a statistic model. In this line of argument, Bloom *et al.* (2003) found that increases in life expectancy play a large role in savings behaviour and can account for the observed saving boom in East Asia during 1950 and 1990.

But before assessing whether the LCM holds in real life and therefore individuals do consider expected longevity when taking decisions like consumption, saving and wealth accumulation, and late life decisions like retirement as rational economic agents suggested by this hypothesis, it is essential taking one step backwards and analysing the main ingredient: expected longevity or how they are commonly known in economic literature, Subjective Survival Probabilities (hereafter, we refer to them as SSPs).

SSPs are individuals' subjective assessment of their mortality risk. They are typically elicited from surveys that ask respondents about the numerical probability that he or she will survive to a given age that is a number of years in the future.

In order to examine whether we can confidently use individual-level probability distributions to estimate models of decision-making under uncertainty, Hurd and McGarry (1997) stated that we need to understand and test their properties. The authors mention three properties that should be satisfied:

1. Individuals understand questions about probabilities and accurately report their beliefs about the likelihood of future events;
2. Individuals adjust their reported probabilities in response to new information; and
3. The reported probabilities predict outcome.

We consider that taking this step backwards is crucial for the analysis of the economic agents and their economic behaviour. When testing whether people take into account their expected longevity when taking late life decisions, it is necessary firstly to test the validity of those expectations. Otherwise, we may find ourselves, for example, in a situation where our analysis suggests SSPs are irrelevant when deciding when to retire. But we cannot establish whether this is because individuals do not act as rational economic agents, as LCM suggests, or because expectations on longevity are not consistent or are not in line with mortality data. As Hamermesh (2004) clearly states, the justification for an economist undertaking this kind of analysis is that the results are important inputs to our basic research and are not likely to be considered (and were not discussed) by other social scientists.

The goal of this Chapter is to test whether SSPs fulfil the first of the three properties that Hurd and McGarry (1997) stated. We concentrate on SSPs drawn from the "Survey on Health, Ageing and Retirement in Europe" (SHARE), which is presented in Section 3.

In particular, testing whether SSPs fulfil the first of the three properties means assessing if SSPs are:

- Consistent across waves;
- Average close to objective probabilities;
- Consistent with mortality data;
- Consistent with epidemiological data; and
- Consistent with economic literature.

For this matter, we first introduce the concept of SSPs. We present SHARE's question on SSPs and we explore the respondents' answers for both waves. We also conduct a comparison of Subjective versus Objective Survival Probabilities using national life tables in order to determine if people's expectations are coherent. We then focus on understanding what are the determinants of SSPs and if they are consistent with mortality and epidemiological data. We, therefore, conduct a cross sectional econometric analysis for each of the two waves of SHARE (2004 and 2008) and contrast results obtained with previous economic and medical studies.

We want to respond to questions like "how are SSPs related to personal characteristics of older Europeans?", "are they consistent with epidemiological data and risk factors?", "do they average close to actual average probabilities in the population using life tables?", "are they consistent throughout time between years 2004 and 2008?", "do they co-vary coherently with economic literature in relation to income, education and other socioeconomic variables?".

In this sense we would expect, for example, that on average a smoker respondent reports lower survival probability while, on average, females report a higher survival probability and poorer respondents express lower expected longevity. This is because medical evidence suggests that people who smoke, on average, do not live as long as non-smokers, mortality data suggests that on average females live longer than males do, and economic literature has proven that, on average, richer people live longer than individuals in the lower economic scale.

The rest of the Chapter is structured as follows: Section 2 provides a literature review on expectations and on SSPs in particular. Section 3 describes the data we have used in our analysis. Section 4 focuses on SSPs, where we include a descriptive analysis of them, modelling strategy followed, and results obtained in the analysis. Finally, in Section 5 we present the conclusions obtained in this Chapter and we introduce Chapter 2.

2. LITERATURE REVIEW

Next, we provide a literature review on the field of subjective expectations and more concretely, on SSPs. As mentioned above, given that only since mid-1990s economists have stated to get real interest in expectations, the literature in this field is recent and not very extensive. Nevertheless, we present the most relevant literature that has shaped this new field of Economics.

Firstly, we present on overview of the literature around the broad area of expectations. We then provide a more detailed review of the work previously done on SSP where we highlight the contribution of Hurd (Hurd and McGarry, 1995; Hurd and McGarry, 1997; and Hurd *et al.*, 2001). We concentrate on Hurd's contribution with respect to initial inspection of SSPs and their determinants. Further contribution by this author is considered in the literature review of Chapter 2 and 3. We also present how SSPs have been analysed outside the US. Finally, we present a brief overview of other surveys which incorporate questions on expectations, apart from SSPs.

2.1. EXPECTATIONS

Among economists, the idea that probabilistic measurement of expectations might improve on attitudinal research appears to have originated with Juster (1966). The author presented a set of experiments designed to test the hypothesis that buying intentions were essentially probability statements in disguise. The experiments were conducted between late 1963 and 1964 and are based on surveys of consumer anticipations. The idea behind was that consumers purchase durable goods as houses, cars, and appliances based on observable variables, such as income, income change, assets, etc. but also based on consumer optimism or pessimism.

Specifically, he considered how responses to traditional yes/no buying intentions questions should properly be interpreted. In particular, the author considered that consumers reporting that they ‘intend to buy A within X months’ can be thought of as saying that the probability of their purchasing A within X months is high enough so that some form of ‘yes’ answer is more accurate than a ‘no’ answer.

He found that survey consumer intentions to buy are not efficient predictors of actual purchase rate, because the adjectival scale reduces the accuracy of the probability judgment. A quantitative scale is, instead, more precise and the mean value of the distribution of probabilities is a good predictor of future purchasing behaviour. The author obtained that the purchase probabilities explain about twice as much of the cross-section variance in automobile purchase rates as buying intentions. Similar but not quite so conclusive differences are obtained from analysis of selected household durables.

He concluded that it would be more informative to ask consumers for their purchase probabilities than for their buying intentions[2]. Regardless of this very important finding, a

[2] Juster (1966) used the following Purchase Probability Questions:

Taking everything into account, what are the prospects that some member of your family will buy a new or old car sometime during the next 6/12/24 months? Certainly, Practically Certain (99 in 100); Almost Sure (9 in 10); Very Probably (8 in 10); Probably (7 in 10); Good Possibility (6 in 10); Fairly Good Possibility (5 in 10); Fair Possibility (4 in 10); Some quarter century passed before economists began to systematically collect and analyse probabilistic expectations data.

As recognized by Manski (1990), after mid-1960s, economist lost interest in the analysis of intentions data, leaving the study of buying intentions to market researchers. As a matter of fact, at that time, the National Academy of Sciences Panel on Survey Measurement of Subjective Phenomena had not economists as a member of the panel. Nevertheless, this author studied the relationship between stated intentions[3] and subsequent economic behaviour. His primary conclusion was that intentions data was not very useful and not too much should be expected from it as it does not predict consumers' behaviour accurately. Instead, subjective probabilities were superior for predicting outcomes.

In this same line of research other studies have shown that subjective probabilities have more predictive power than qualitative responses. According to Dominitzand Manski (1997a), the elicitation of probabilistic expectations should be preferred to qualitative questions because problems may arise in the interpretation of qualitative responses, since they are subject to large variation between individuals. In order to study the cross-sectional variation in income expectations, one-year-ahead income expectations from the Survey of Economic Expectations are used.

The respondents were questioned:

"What do you think is the percent chance (or what are the chances out of 100) that your total household income, before taxes, will be less than Y over the next 12 months?"

The authors used four different income thresholds Y. They found that much of the cross-sectional variation in the central tendency of income expectations is associated with realized income, and some of the cross-sectional variation in income uncertainty is associated with realized income, age, and employment status.

Hamermesh (2004) believes that there has been an upsurge in the use and analysis in economics of subjective outcomes lately because the results are important inputs to

---

Possibility (3 in 10); Slight Possibility (2 in 10); Very Slight Possibility (1 in 10);No Chance, Almost No Chance (1 in 100).

[3] "Intentions" are represented by answers to questions such as *"Do you intent to purchase an automobile?"* and if so *"Would it be certain, very likely or likely"*.

economists' basic research. For that reason he suggested that a hierarchy is needed in order to base the research on economists' comparative advantage—coherent models of underlying behaviour— instead of duplicating what other social scientists have already done many times over or, worse still, generate results that have absolutely no economic meaning. Specifically, the author proposed the following ranking from less to more desirable approaches to research using subjective measures:

1. Atheoretical statistical models examining subjective outcomes and explaining them by their subjective determinants;

2. Atheoretical statistical models examining the objective determinants of subjective outcomes that are not relevant inputs into describing economic behaviour;

3. Atheoretical statistical models examining the objective determinants of subjective outcomes that are important inputs into describing economic behaviour;

4. Theoretically based statistical models of important, albeit non-economic outcomes. There is no doubt many other areas besides time stress where we can analogize a subjective outcome of general interest to a concept in one of our models and derive how individual agents' maximizing behaviour affects the outcome; and

5. Theoretically based statistical models of the determinants of economically relevant subjective outcomes.

Hamermesh (2004) mentioned that economists contribute little if we merely engage in fancier empirical work and still less if we describe subjective outcomes by other subjective outcomes. The biggest contributions can be in adding economic theories that allow a better understanding of objective behaviour using subjective outcomes, or of the determinants of subjective outcomes; or in understanding subjective outcomes, such as expectations, that underlie objective economic behaviour.

In his very influential survey, Manski (2004) stated that subjective probabilities have compelling advantages relative to verbal questions. In the first place, probabilities provide a well-defined numerical scale when reporting them, allowing for interpersonal comparison. Also, subjective probabilities about different events allow for empirical assessment of internal consistency (Bayes Theorem, the Law of Total Probability, etc.). Finally, a researcher can easily reach conclusions about the correspondence between subjective beliefs and frequentist realities.

Manski (2004) stated that we had already learned enough to recommend that economists should abandon their antipathy to measurement of expectations. There is a critical need for basic research on expectations formation. The unattractive alternative to measurement is to make unsubstantiated assumptions.

2.2. LITERATURE ON SSP

The analysis of individual reports on survival probabilities has only become popular in recent years. Only since mid-1990s economists have realized that the quality of the decisions that individuals take related to saving and investment, consumption, wealth accumulation, retirement among others, depends primarily on the accuracy of individual's subjective probabilities about future events. In particular, individuals' own assessment on their mortality risk (SSP) plays a key role on late life decisions. For example, consumption and savings decisions of an individual are thought to depend on beliefs about future interest rates, the likelihood of dying, and the risk of substantial future medical expenditures.

Two of the first authors to consider expected longevity as an element of study were Hamermesh and Hamermesh (1983). They analysed the responses on life expectancy questions of two different samples: male PhD economists teaching at four year colleges or universities, and a random sample of males. The question, *"How old do you expect you will live to be?"* was designed to elicit subjective life expectancy.

A basic econometric analysis undertaken revealed that subjective life expectancy varied appropriately with personal characteristics, and that no significant differences were found among the two sample groups.

Next, Hamermesh (1985) analysed the responses of two groups, 650 white male academic economists and 975 males random chosen to a questionnaire design to reveal life expectancy, as above. But also, for the first time, the questionnaire included subjective probability of survival to ages 60 and 80. The author tested internal consistency comparing deviations of subjective from actuarial survival probabilities (from life tables) to deviations of subjective from actuarial life expectancy. The results suggest that people think about survival probabilities differently from the way they forecast life expectancy. It was also tested whether subjective life expectations equal actuarial and it was found that people's subjective horizon slightly exceeds actuarial horizon. Hamermesh (1985) concluded that people do not extrapolate past improvements in longevity when they determine their subjective horizon. Whether they are aware of the trends in improving longevity and do not consider them, or they are not aware, today's life tables are a good proxy of people's expectations today.

It should be mentioned that these two preliminary studies made use a non-representative sample of individuals. One sample consisted of people with above average education; the other was selected from only one geographic area. Both contained only white males. Therefore the studies had the drawback that results were not representative of the US population. Nevertheless, they built the foundations of future research on SSPs.

2.2.1. Hurd's contribution

It wasn't until mid-1990s when SSPs really became popular. This coincides with the fact that significant amount of data from surveys became available in the US by that time. Specifically, the Health and Retirement Study (HRS)[4] had fielded questions on expectations among a population-representative sample. Expectation questions included the probability of a double-digit inflation rate, depression, working full time past ages 62 and 65, or moving into a nursing home, for example. Additionally, the HRS also includes questions about subjective survival chances to age 75 and to age 85. This was the first

[4] The University of Michigan Health and Retirement Study (HRS) surveys more than 22,000 Americans over the age of 50 every two years. Supported by the National Institute on Aging (NIA U01AG009740) and the Social Security Administration, the HRS is a large-scale longitudinal project that studies the labor force participation and health transitions that individuals undergo towards the end of their work lives and in the years that follow.

Since its launch in 1992, the study has collected information about income, work, assets, pension plans, health insurance, disability, physical health and functioning, cognitive functioning, and health care expenditures. Through its unique and in-depth interviews with a nationally representative sample of adults over the age of 50, the HRS provides an invaluable, growing body of multidisciplinary data to help researchers address the challenges and opportunities of ageing.

time that economists were able to have access to a large database from a population-representative sample which included respondents' survival chances, but also a significant amount of information regarding socio-economic characteristics, health status, financial situation, etc.

The first economist that analysed the reported survival chances in the HRS was Hurd in a series of papers (Hurd and McGarry, 1995; Hurd and McGarry, 1997; and Hurd *et al.*, 2001). He popularized them and since then they have been extensively studied in the US. As mentioned in the introduction of this section, these papers focused on an initial inspection of SSPs and their determinants. Further contribution by this author is considered in the literature review of Chapter 2 and 3. He popularized them and since then they have been extensively studied in the US.

Hurd and McGarry (1995) inaugurated a series of studies that had SSPs as the main ingredient. The authors used the first wave of the HRS to seek for two objectives: (i) check external consistency of SSPs. They wanted to understand how the probabilities compare with probabilities found in external data; and (ii) check internal consistency of SSPs. They wanted to determine if they behave like probabilities.

Regarding external consistency, Hurd and McGarry (1995) compared the reported survival chances with life tables from 1988. They obtained that men substantially over-estimate the probability they will live to 85, and women under-estimate the probability they will live to 75. As a consequence, both over-estimate the conditional probability of living to 85 given alive at 75. They concluded that this result is coherent with the likely mortality patterns to be observed in the future. Because life tables from year 1988 are constructed with information about people that died in 1988, it might not reflect the mortality risk of people that are likely to die not today, but maybe 20 years after. In a sense, if mortality risk was stationary over time, life tables from any year would be a relevant comparator.

On internal consistency, the authors compared SSPs to age 75 and age 85 across individual level. Because each individual has a positive probability of dying between 75 and 85 should she live to 75, SSP to age 75 should be greater than that to age 85. They found that only 2.5% of the respondents reported higher probability of reaching 85 than reaching 75, and 9.2% reported both probabilities equal to 1. Overall they constitute

11.7% of the whole sample that may not have understood the nature of the questions. Hurd and McGarry (1995) stated that all variables derived from household interviews have inconsistencies and observational errors, and that in this case these inconsistencies were tolerable.

Additionally, the authors analysed how SSPs correlated with other variables using simple equations. Variables used included measures of socio-economic status, personal characteristics, risk factors, diseases, and self-assessed health status. They obtained that generally SSPs covary with these variables in the same way actual outcome vary with the variables.

The first well-known paper on individually reported SSPs using the HRS dataset left us with an encouraging conclusion: this measure of subjective probability had great potential use in models of inter-temporal decision making under uncertainty. Moreover, Hurd and McGarry (1995) examined the probability attached to work after 62 and 65 years old that the HRS respondents report. They also concluded that these probabilities will be of great interest in future models.

The next important contribution to the SSP literature came with Hurd and McGarry (1997) when they published a paper when they took one step forward and analysed the predictive validity of expected longevity. Given that in the previous paper the results were based on cross-sectional analysis, it was not possible to establish whether individual reports of subjective probabilities provided information about mortality or not. Moreover, it was not possible to determine if SSPs had any power to predict mortality or how they evolve as new information comes to the respondents. Using a panel data from the first and second wave of the HRS, the authors tried to answer these questions with the objective to increase the confidence to use SSPs in models of decision-making under uncertainty.

It was concluded that respondents modify appropriately their survival probabilities based on this new information. Furthermore, they examined how well SSP at wave 1 predicts actual mortality between wave 1 and wave 2. Using a probit specification, Hurd and McGarry (1997) concluded that indeed, reported survival chance predict actual mortality and that those who survived to wave 2 reported probabilities approximately 50% greater at baseline than those who died[5].

Hurd and McGarry (1997) also performed a prior analysis to SSPs drawn from the second wave of the HRS. For this matter, a cross-sectional OLS approach was followed. The authors considered a regression of SSP to age 75 and to age 85 on the wealth and income quartiles, health variables, educational level and parental mortality variables. They concluded that SSPs aggregate close to life table values and covary appropriately with known risk factors.

A very similar approach was used by Hurd *et al.* (2001) when analysing the predictors of mortality among the elderly and how SSPs evolved over time[6]. But this time they used the first two waves of the Asset and Health Dynamics among the Oldest-Old (AHEAD) survey. This study is a biennial panel survey of individuals born in 1923 or earlier and their spouses. As in Hurd and McGarry (1997), they found that in cross-section the SSPs were related to baseline health conditions and they were higher among survivors. Moreover, Hurd *et al.* (2001) concluded that SSPs predict actual mortality.

A statistical inspection of SSPs revealed that average survival probability declines with age. Surprisingly, unlike actual mortality, it was found that there was little systematic variation in SSPs as a function of wealth, income quartiles, and education bands. Hurd *et al.* (2001) stated that a possible reason for the lack of any pattern by wealth, income, or education is the rather high rate of non-response.

With respect to the determinants of SSPs, a cross-sectional OLS approach was followed. The authors considered a regression of SSP on the wealth and income quartiles, education bands, and other explanatory variables. Two different models were estimated, including and excluding health variables. It was found that eight of the 13 health variables were significant at the 0.05 level, and they are associated with a reduction in the SSPs of 9 to 25%. Hurd *et al.* (2001) concluded that based on these results, it would be expected that

[5] For further details on the analysis of the predictive validity and the evolution of SSPs performed in Hurd and McGarry (1997), refer to the literature review of Chapter 2.

[6] For further details on the analysis of the predictive validity and the evolution of SSPs performed in Hurd *et al.* (2001), refer to the literature review of Chapter 2.

SSPs predict actual mortality because of their association with the health conditions which, themselves, are associated with mortality.

2.2.2. Further contributions

Other authors have also contributed to the literature on SSP. Smith *et al.* (2001) used a more robust panel data to assess whether longevity expectations were a predictor of mortality. Concretely they used the first four waves of the HRS. They concluded that SSPs were reasonably good predictions of future mortality. Additionally, Smith *et al.* (2001) reached three important conclusions: (i) subjective beliefs about longevity are consistent with individuals' observed survival patterns. The observed deaths are signalled through the lower longevity expectations respondents report in earlier interviews; (ii) SSPs of those who died later display a downwards trend over time; and (iii) longevity expectations respond negatively both to serious, new health shocks and to increases in individuals' functional limitations.

In the same line of investigation Siegel *et al.* (2003) tried to assess whether self-rated life expectancy predicted actual mortality after controlling for measures of health, self-rated health, and socio-demographic characteristics. Using HRS and AHEAD surveys, they obtained that for the latter both self-rated life expectancy and self-rated health predicted mortality. For the HRS Siegel *et al.* (2003) found that self-rated life expectancy was not significantly associated with mortality when variables related with self-rated health were included.

The economic literature then focused its attention not on SSPs themselves, but rather as an input of models of inter-temporal decision making under uncertainty. A complete review of this economic literature is provided in Chapter 3.

2.2.3. SSP outside US population based surveys

As showed above all, the economic literature on SSPs at the beginning was based on American surveys like the HRS and AHEAD. In other countries the economic literature is limited and was developed later. This economic literature, however, has not reached the level and degree of analysis compared to the American based one. One of the main reasons for this is that surveys similar to the HRS and AHEAD were only released during the last years of the 1990s and in most cases during the first decade of the new century.

Van Doorn and Kasl (1998) used the Australian Longitudinal Study of Ageing[7] (ALSA) that collects data on a population representative sample of individuals over 70 years old. Unlike the HRS, respondents are asked to assess their survival chances for the next 10 years using a non-numerical scale (i.e. *very likely*, *likely*, *unlikely*, and *very unlikely*). The objective of the study was to determine whether global self-rated life expectancy and parental mortality predict mortality. Van Doorn and Kasl (1998) found that for men self-rated life expectancy does predict mortality, however for women the predicting power is null. Additionally they found that self-rated life expectancy is a better predictor of mortality among the old than the young. The reason for this is that we would expect that thinking about death is more relevant to the lives of older people who are probably experiencing losses of friends and relatives more frequently than younger people.

Other studies in Australia have also tested the relationship between family history or longevity and mortality though they do not use a population representative sample. Two of these studies do find a significant relationship (Borawski *et al.*, 1996; Pijls *et al.*, 1993).

The English Longitudinal Study on Ageing[8] (ELSA) is a survey that contains quantitative information on people's expectations for the future, and in particular respondents are

[7] The Australian Longitudinal Study of Ageing (ALSA) is Australia's first multi-dimensional population based study of human ageing. The ALSA commenced in 1992 (under the direction of the late Prof Gary Andrews), with 2087 participants aged 70 years or more. At Baseline, a comprehensive personal interview and assessment of neuropsychological and physiological functions was undertaken at each person's home, supplemented by self-completed questionnaires, biochemistry, and additional clinical studies of neuropsychology and physical function. Since then a further 10 waves have been completed (some being short telephone interviews), with Wave 11 being completed in 2010. The general purpose of the ALSA study is to gain further understanding of how social, biomedical and environmental factors are associated with age related changes in health and well-being of persons aged 70 years and over.

[8] The English Longitudinal Study of Ageing (ELSA) is an interdisciplinary data resource on health, economic position and quality of life as people age. ELSA is the first study in the UK to connect the full range of topics necessary to understand the economic, social, psychological and health elements of the ageing process. The aim of ELSA is to explore the unfolding dynamic relationships between health, functioning, social networks and economic position. It is in effect a study of people's quality of life as they age beyond 50 and of the factors associated with it. ELSA first wave was released in 2002. In 2008 the fourth wave was released.

asked about their SSPs. As in the case of the HRS, individuals are asked to give a percentage chance that they will live to age 75 or more on a scale between 0 and 100[9].

However, they have not called much attention from economist and therefore have not been analysed into detail as in the case of the US. Banks and Emmerson (2004) analysed longevity expectations and future retirement consequences. The idea is that because people now live longer, there is a concern that people will not incorporate the longevity improvements into their retirement decisions, and hence may not save "enough". The authors provide a descriptive analysis of SSPs and their relationship with other variables.

Concretely, they report that those with higher education and income report higher chances of living to 75, and those with worst self-reported health status, and with previous conditions such as heart attacks, along with smokers, report lower chances. They also compare SSPs with information from English life tables. They found that, on average, both men and women underestimate their survival chances of living to 75.

Moreover, even though women live longer than men, women only report chances only slightly greater than those for men. As a result, women underestimate their survival chances by considerably more than men[10]. Finally the authors conclude that there is clearly much more to be done into English longevity expectations, but they warned that expectations errors, particularly if they are more prevalent in some groups than others, could have important policy implications.

O'Donnell *et al.* (2008) used the first three waves of ELSA to test whether the timing of retirement was responsive to SSPs. They found that indeed there is a relationship between these two variables. Surprisingly they concluded that men with especially low survival expectation are less likely to retire. However they did find that retirement likelihood decreased as survival expectations increase.

The Bureau of National Health Insurance of Taiwan (BNHI) is a survey that includes a question on SSPs. This question is exactly the same as the one in the HRS. Liu *et al.* (2007)

---

[9] In ELSA, older groups get different age reference points: for example, respondents aged 65-74 are asked to report the chance to living to 85.

[10] According to Banks and Emmerson (2004) women underestimate their likely longevity more than men. Women aged 60-64 put their chances of reaching 75 at just 65%; in reality, according to life tables, the probability is more than 80%.

analysed the determinants of SSPs and examined how individuals update their longevity expectations as new health information is released. The analysis is based on individuals' subjective perception of longevity before and after their physical examinations, as well as their medical diagnoses and advice on health behaviour provided by physicians.

The authors found that males and married persons are more optimistic about their longevity, and that income is strongly correlated with higher SSP. However, education does not seem to affect expected longevity. Liu *et al.* (2007) also concluded that individuals who acquire new negative health information revise downwards their survival chances. In particular, the expected longevity declines with health shocks such as a heart attack and with abnormal test results.

A recent study by Steffen (2009) explores somewhat similar issues. Contrary to some of the papers presented above, this paper explores the determinants of Subjective Life Expectancy (measured in year) as oppose to SSPs (measured in percentages). However, interesting conclusions can be obtained from it. The study is based on the German database SAVE. The SAVE study is a national representative survey of sociological, psychological, and financial characteristics of German households. The study started in 2001 and has been conducted annually from 2005 on. It comprises more than 3,000 German households. For the study, the author used the 2005, 2006 and 2007 SAVE waves.

As mentioned above, the peculiarity of this paper is that it concentrates on Subjective Life Expectancy. In the SAVE study, this is surveyed in several steps. First, respondents are asked which age they think men and women of their age will reach on average. Next, respondents are asked whether they think they will live longer, shorter or about the same as average. Finally, respondents are questioned how many years they will live longer or shorter, depending on the previous questions.

In order to analyse the determinants of Subjective Life Expectancy, a pooled OLS estimation method was used and it was assumed that Subjective Life Expectancy is a function of variables related to sex, age, family situation, educational level, economic situation, lifestyle and health. Additionally, Steffen (2009) also analyse the influence of these variables on the reported average life expectancy and on the probability of thinking to live longer using a probit model.

The author concluded that by far the most important determinant of Subjective Life Expectancy is the individual health situation. When economic variables are included, their importance is rather small. Moreover, smoking significantly reduces Subjective Life Expectancy but this is mostly driven by a lower average life expectancy. Finally, the better educated a person is, the higher the average life expectancy.

Steffen (2009) also analyses how Subjective Life Expectancy updates with new information. Details can be found in the literature review of Chapter 2.

SHARE also incorporates a question on SSPs. This survey incorporates information on a population representative sample of Europeans aged 45 or over. Given that the first wave was only released in 2004 and the second wave in 208, European SSPs have not been analysed in detail by economists. Up to our knowledge there are not too many studies that make use of SHARE's expected longevity data.

Balia (2007) investigated the formation of expected longevity through smoking behaviour for Italian respondents of SHARE. She uses the first wave of SHARE to briefly compare SSPs with Italian life tables and reports the variation of them with respect to health, risk factors and socio-economic variables. The study proposed a recursive model for expected longevity, smoking duration and self-reported health taking into account unobservable individual –specific heterogeneity.

The author found two types of smokers in the Italian population. The first type is more addicted to tobacco and do not internalize fully the health effect of its consumption. The second type is less addicted and seems to take into account the health effect of tobacco. Additionally it was found that current and former smokers differ in the way they discount future consequences of tobacco consumption. The main drawback of this interesting study is that it only concentrates on Italian respondents.

Delavande and Rohwedeer (2008) analysed the differential mortality in Europe and the US based on SSPs. They propose a new methodology to produce estimates of differential mortality. Traditionally it is required rich panel data with large sample, but they propose to relate SSPs to variables of socio-economic status in cross-section. This exercise is done using the HRS, ELSA and SHARE survey. Delavande and Rohwedeer (2008) concluded that this methodology performs well enough to constitute a powerful alternative in the absence of comparable longitudinal data across countries to provide international comparison of differential mortality.

In an unpublished short manuscript from the CeMMAP Workshop on "Novel measurement methods for understanding economic behaviour" Hurd *et al.* (2008) presented an international comparison of SSPs. In this manuscript they mention that survival expectations influence important life-cycle decisions, in particular for older people. Additionally, structural models of life-cycle behaviour require reliable measures of expected longevity. The main objective of this manuscript is to compare response patterns to questions on SSPs in different surveys. For this matter, the authors use the HRS, the ELSA, and the Survey of Health, Ageing and Retirement in Europe.

They compared SSPs in SHARE with European life tables and found that females tend to underestimate their survival chances before age 75 and both males and females overestimate it after age 75. Also, when comparing average longevity expectation by self-reported health status in SHARE and the 2004 wave of the HRS, they obtained a coherent pattern. They concluded that probabilistic survival questions seem to work fine. Unfortunately no paper has been published by the authors on this short manuscript.

Post and Hanewald (2011) studied the relationship between longevity risk, SSPs, and saving behaviour using the second wave of SHARE. They concluded that individuals are indeed aware of their longevity risk and that this awareness affects their saving decisions.

2.3. OTHER EXPECTATIONS

It should be mentioned that the use of expectations in economic literature does not concentrate only on SSPs. Economists are also interested on individuals' expectations on bequest, retirement, investment returns, income, employment, fertility, crime victimization, among others.

Most of this literature is built on the following surveys:

- HRS: Apart from SSPs, respondents are asked to assess the chance of leaving bequest, investment returns and retirement at age 62 and 65;

- Bank of Italy Survey of Household Income and Wealth. This is a biennial survey that started in 1989. Individuals are asked to report subjective probability distributions for the growth rate of nominal labour earnings and pensions and for the rate of inflation over the next twelve months;

- Survey of Economic Expectations. It is a nationwide survey that examines how Americans in the labour force perceive their near-term economic future. It comprises 8 waves collected between 1994 until 2002.In all waves of SEE, respondents were asked to report expectations for crime victimization, health insurance, employment, and income. In some waves, they were asked about returns on mutual-fund investments and about their future Social Security benefits;

- National Longitudinal Survey of Youth 1997: It is a nationally representative sample of approximately 9,000 American youths who were 12 to 16 years old as of December 31, 1996. Round 1 of the survey took place in 1997.Youths are interviewed on an annual basis. It includes questions on expectations on schooling, fertility, crime victimization and arrest; and

- Michigan Survey of Consumers: Published every month, this survey is a consumer confidence index. Respondents are asked to report expectations on investment return, income and employment status.

In the next section we present in detail the data that will be used on our analysis and modelling of SSPs.

3. DATA

In order to investigate the consistency of SSPs and their determinants we have used data drawn from the first two and only waves publicly available of SHARE that provides information on Europeans aged 45 or over.

SHARE is a multidisciplinary and cross-national panel database of micro data on health, socioeconomic status and social and family networks[11]. The first wave of the panel was released in 2004 while the second became available in 2008. In 2011 a retrospective survey which focused on people's life histories was released. Even though this survey is not a proper additional wave of SHARE, it is commonly known as the third wave. Currently data for the next wave is being collected and is expected to be released at the beginning of 2013.SHARE is coordinated centrally at the Mannheim Research Institute for the Economics of Aging (MEA). It is harmonized with the HRS and the ELSA surveys.

Eleven countries have contributed with data to the 2004 SHARE baseline study. They are a balanced representation of the various regions in Europe, ranging from Scandinavia (Denmark and Sweden) through Central Europe (Austria, France, Germany, Switzerland, Belgium, and the Netherlands) to the Mediterranean (Spain, Italy and Greece). Further data has been collected in 2005-06 in Israel, though, due to the lack of complete information for this country, it has been excluded from the analysis. Two 'new' EU member states - the Czech Republic and Poland - as well as Ireland joined SHARE in 2006 and participated in the second wave of data collection in 2006-07.

Based on probability samples in all participating countries, SHARE represents the non-institutionalized population aged 50 and older. Spouses are also interviewed if they are younger than 50. In the first wave the survey consisted of a sample of 31,115 individuals (13,811 males and 17,304 females) and in the second wave 33,281 respondents participated (14,749 males and 18,532 females). A total of 18,741 individuals have participated in both waves (8,270 males and 10,471 females). In all three cases the majority of interviewees were aged between under 66 years old.

For the analysis undertaken we have considered individuals aged 45 up to 65 years old. After eliminating missing values and non-respondent items, the final sample used in the analysis consists of 15,058 respondents for the first wave and 16,485 for the second wave.

[11] Data collected include health variables (e.g. self-reported health, health conditions, physical and cognitive functioning, health behaviour, use of health care facilities), bio-markers (e.g. grip strength, body-mass index, peak flow), psychological variables (e.g. psychological health, well-being, life satisfaction), economic variables (current work activity, job characteristics, opportunities to work past retirement age, sources and composition of current income, wealth and consumption, housing, education), and social support variables (e.g. assistance within families, transfers of income and assets, social networks, volunteer activities).

We have found 8,448 individuals that have participated in both waves of the SHARE survey.

We concentrate on individuals aged under 66 for two reasons. In the first place, because of the way the question on SSPs is worded (*what are the chances that you will live to be age T or more?* where the target age, T, contained in this question is chosen conditional on the respondent's age), individuals in our final sample are asked about the chances of reaching a common target age, in this case 75 years old. For older respondents, the target age is set in five-year bands such that the distance from current to target age is between 10 and 15 years.The second reason is because the vast majority of longitudinal respondents are aged 65 or under. At the same time, individuals from this age interval are more likely to take retirement decisions, which in turn, will be relevant for the Chapter 3.

As mentioned in the previous section, SSPs is a rather new research topic in SHARE. There are not too many studies that rely on SHARE's expected longevity data. We can highlight the work of Balia (2007), Delavande and Rohwedeer (2008), Hurd *et al.* (2008) and Post and Hanewald (2011).

Even though expected longevity has not been extensively studied using SHARE data, since its release in 2004 SHARE has become increasingly popular among social scientists and researchers and particularly among labour economists. SHARE has been the data source for many books, book chapters, journal articles and working papers. Next, we present some of the most relevant publications that have used SHARE database.[12]

Given the substantial amount of data related with psychological health, it has been extensively studied using SHARE database. Abu-Rayya (2006) concluded that European elders participate less in social activities with increasing age to statistically significant degrees; younger elders tend to be more socially involved and older elders less so. Peytremann-Bridevaux and Santos-Eggimann (2008) findings suggest that depressive symptoms were associated with significantly greater use of all healthcare domains but not preventive services, with the exception of colorectal cancer screening. Sundström *et al.* (2009) found that the combination of living alone and having bad health is associated with

[12] A complete list of SHARE based publications can be found in: http://www.share-project.org/.

10 times higher odds of feeling lonely as compared with living together with someone and having good health.

General health studies have also become popular within researchers using SHARE data. Obesity in older Europeans was examined by Andreyeva (2006). This researcher concluded that health education and health promotion as well as taxes and subsidies to regulate health are important prevention and intervention tools the government can use to reduce obesity. Rueda and Artazcoz (2008) used logistic regression models to analyse gender inequalities in health status and in social determinants of health. They concluded that health inequalities persist among the elderly. Women have poorer health statuses than men and in both sexes the risk of poor health statuses increases among those with low educational attainment.

Solé-Auró and Crimmins (2008) studied health differences between immigrants and the native born populations aged 50 years and older in 11 European countries using SHARE database. The authors found that migrants generally have worse health than the native population.

SHARE based retirement and labour supply has also been a recurrent research area among economists. Blanchet *et al.* (2005) reviewed different retirement pathways from the first wave of SHARE. Specifically they analysed unemployment and disability as a form of pre-retirement and gradual retirement in Europe. In the same line, Carta (2007) analysed the causes of involuntary and voluntary early retirement in Italy.

Alternatively, Pagán (2009) studied the determinants of part-time labour supply of older workers with disabilities in Europe and concluded that being disabled has a positive effect on the probability of working on a part-time basis, although this effect varies by country. Similarly, Alavinia and Burdorf (2008) explored the relationship between unemployment, retirement and health and found that in many European countries poor health, chronic diseases, and lifestyle factors were associated with being out of the labour market. Crespo and Mira (2008) studied caring for parents and employment decisions for middle-aged women. The results show that the estimated effect of providing informal care to elderly parents on the probability of labour participation is negative and large.

Other subjects have also been studied using SHARE information. The relationship between old European and their adult children was examined by Fokkema *et al.* (2008). They concluded that the nature of parent-child solidarity differs between European countries. Geographical proximity, contact frequency and feelings of family care obligations exhibit the general north-south divide. Erlinghagen and Hank (2006) studied the determinants of the participation of older Europeans in volunteer work. Their analysis reveals a clear spatial pattern, with relatively high participation rates in Northern Europe and relatively low participation rates in Mediterranean countries, and shows that age, education, health and involvement in other social activities strongly influence an individual's propensity to engage in volunteer work.

Finally, an analysis of housing and accommodation patterns of Europeans aged 50 or over was conducted by Angelini and Laferrère (2008).

4. SUBJECTIVE SURVIVAL PROBABILITIES

SHARE dataset presents a variety of topics selected for their policy relevance for the age 50+ segment of the European population. Specifically, individuals' beliefs about future events are fundamental in forward-looking models because probability beliefs may affect household decisions.

SHARE respondents are asked several questions about their expectations regarding future events. For example, individuals are asked to attach probabilities to the likelihood of leaving or receiving a large bequest, government reducing pension, government rising official retirement age, improving of standard of living, and life expectancy. In addition, respondents aged less than 61 in the second wave are asked about the chances to work after age of 63.

The main object of interest of this paper is the individuals' subjective assessment of their mortality risk. The ideal scenario would be to obtain their subjective beliefs for different time horizons but SHARE contains only one such question worded as follows:

What are the chances that you will live to be age T or more?

Where the target age, T, contained in this question is chosen conditional on the respondent's age. As reported in Table 1, for respondents younger than 66, the target age is 75 but for older respondents, the target age is set in five-year bands such that the distance from current to target age is between 10 and 15 years. The question of SSPs is the ninth of 11 questions about expectations. This point is rather important because previous questions should have introduced respondents to the probabilistic format, and therefore, we would expect more correct answers.

For the purpose of this Thesis, we have selected individuals aged between 45 and 65, so that they are asked about a common target age: 75 years old. Individuals' responses range from 0 to 100 and we treat them as measures of SSPs.

As mentioned in Section 2, other surveys have also incorporated questions about individual's expectations about future events including SSPs. Particularly, in the HRS for the US, individuals are requested to use any number between zero and ten, where zero equals absolutely no change, and ten equals absolutely certain, to assess the probability of working, of housing purchase, job stability, financial help to family, housing prices, Social Security, the economy, and life expectancy.

In relation to SSPs, the main difference between HRS and SHARE questionnaire is that in the former, individuals are asked for the chances of living to 75 years old or more, but also they are asked about the chances of living to 85 years old. This difference was the main driver used by Hurd and McGarry (1995) to assess internal consistency of expected longevity. The authors quantified that around 70% of the answers satisfy the condition that the probability of reaching 75 years old is higher than the probability of reaching 85, but 2.5% of answers do not satisfy this condition, which can be interpreted as if those respondents may not have understood the nature of the question. They concluded that, even though there is some internal inconsistency, broadly speaking the observations of SSPs act as probabilities.

Other surveys that specifically ask respondents about expected longevity are the ELSA, where the question is essentially the same as the one in SHARE, and the BNHI where, on the contrary, the question is exactly the same as the one in the HRS. Contrary, ALSA respondents are asked to assess their survival chances for the next 10 years using a non-numerical scale (i.e. *very likely, likely, unlikely, and very unlikely*).

One of the main concerns in the expectations literature is the presence of focal point responses. Researchers are often sceptical about the validity of the responses and believe that individuals will only give the responses 0, 50, or 100%. However, Manski (2004) stated that generally, respondents use the full expanse of the 0-100% chance scale, typically rounding to the nearest 5%. Hurd and McGarry (1995) and Liu *et al.* (2007) also found presence of focal point responses and they argued that one interpretation of this result is that many respondents choose one of the three points according to whether they are rather confident, uncertain, or not at all confident about reaching age 75.

Figure 1 illustrates the frequency distribution of the subjective probabilities of surviving to age 75 for SHARE respondents in wave 1 and wave 2. It can be observed that indeed, reported life expectancy presents focal points. Interviewees do use the whole scale but around 55% of them report 50, 80 or 100%. It should also be mentioned that the percentage of refusals or "don't know" answers account for 7.5%. These results are common to the two waves in SHARE.

Nevertheless, we observe less severe focal point responses in SHARE compared with similar datasets. Results obtained by Hurd and McGarry (1995) in the HRS show that 8% of HRS respondents reported zero probability, around 22% reported 50% probability, and 23% reported 100% probability. Liu *et al.* (2007) using BNHI data from Taiwan, found that 18% reported zero probability, around 20% reported 50% chance, and 33% reported 100% chance. On the contrary, in the SHARE dataset we find less severe focal point responses. In particular we find that only around 2% of individuals report zero probability, approximately 22% report 50% probability, and roughly 18% report 100% chance.[13]

[13] In the first wave 1.96% of reported zero probability, 22.35% reported 50% and 17.97% answered 100%. In the second wave these figures were 2%, 21.63% and 17.82% respectively.

Table 2 reports SSPs for females and males in the different countries selected for the study together with the probability of reaching age 75 constructed from life tables[14]. A life table (also called mortality or actuarial table) shows, for every age, what is the probability that someone of that age will die before his/her next birthday, therefore the probability of reaching a specific age, in this case 75, can be easily derived. The Human Mortality Database (HMD) provides death rates and life tables for many countries[15]. For the comparison, we have used life tables from the year 2000.

Comparing survival expectations and life tables is reasonable if we suppose that mortality risk is stationary over time and there is no heterogeneity in the population, because life tables are constructed from observed deceases while individual expected longevity is related with future decease. Even though these two conditions are unlikely to be met, this comparison reveals several aspects that are worth mentioning.

In the first place, on average, females report higher SSP than males, which is coherent with mortality data, as females, on average, live longer than males do[16]. However, in southern countries like Italy, Greece and France average SSP is slightly higher for males than for females. This is due to the fact that for these countries the age distribution of the sample is very different by gender. In particular, there is a much higher proportion of males between 60 and 65 years old.

Direct comparison between males and females reveals differences not only geographically, but also by age. Table 3 presents the results. We can observe that females report higher SSP at ages ranging from 50 to 59 and at age 62[17]. Of course, there is not a single explanation to this observed statistical fact as it is likely that many facts are interacting, therefore a cross-sectional approach is needed.

---

[14] Reported SSP corresponds to the mean of wave one and two, except for Czech Republic, Poland and Ireland. For these countries, because they did not take part in the first wave, we report SSP in the second wave.

[15] Free access data from the website http://www.mortality.org.

[16] Wave 1: average SSP females = 68.69; average SSP males = 67.95.Wave 2: average SSP females = 68.38; average SSP males = 67.54.

[17] Comparison of Subjective Survival Probabilities for ages between 45 and 49 is biased because the sample size is too small.

Comparing SSP with national life tables -actuarial probabilities- also reveals two important aspects; on the one hand, females persistently underreport their probabilities of reaching 75 at every age and in every country but Denmark, while males on average over report their survival chances especially when they are younger. These results are coherent with those obtained by Hurd and McGarry (1995) and Liu *et al.* (2007). The former obtained concluded that men substantially overestimate the probability they will live to 85 while women underestimate the probability they will live to 75. Contrary, Viscusi and Hakes (2003) found that both white males and females underreport the probabilities of living to 75. Banks and Emmerson (2004) found that British women underestimate their likely longevity more than men as women aged 60-64 put their chances of reaching 75 at just 65%; in reality the probability is more than 80%.

We follow the approach proposed by Viscusi and Hakes (2003) and used by Liu *et al.* (2007) and estimate a linear regression of SSP on actuarial probability for the whole sample and for males and females separately[18]. Results are shown in Table 4. The outcome shows that the three slopes are well below one and, in the case of females, non-significant. These results imply that people's beliefs do not always coincide with objective survival chances, and this conclusion is more considerable in the case of females.

4.1. CROSS-SECTIONAL ANALYSIS

In this section we examine the relationship between an individual's reported SSP and her characteristics. The main objective of this section is to assess the consistency of subjective beliefs with respect to personal characteristics. In this sense, we would expect that, for example on average, smokers reported lower life expectancy than non-smokers, males lower than females, more educated individuals higher that less educated, healthy respondents higher than sick respondents, etc. because of epidemiological and mortality evidence.

[18] Table 4 presents the results using data for the first wave. Results using data for the second wave do not vary significantly.

The cross-sectional analysis is done separately for every of the two waves available in the SHARE database. This approach will allow us to examine the intertemporal consistency of the results obtained.

The relationship between reported survival chances and covariates can be expressed as follows:

$$SSP_i = \alpha + \beta X_i + \gamma H1_i + \delta H2_i + \eta S_i + \theta PM_i + \varepsilon_i \qquad (1)$$

where subscript i represents each respondent, X_i includes socio-demographic characteristics, $H1_i$ includes health condition variables, $H2_i$ includes health behaviour variables, S_i represents self-reported health, PM_i stands for parental mortality experience, and finally residual, ε_i, represents the unobservable determinants of the SSP.

This approach was initially suggested by Hurd and McGarry (1995) when analysing the determinants of SSPs in the first wave of the HRS. He proposed two alternative models: the first model excluded self-reported health status in the regression while the second model included it[19]. Hurd and McGarry (1997), Hurd *et al.* (2001) also suggested this approach when examine the determinants of SSPs in the second wave of the HRS and in the first two waves of the AHEAD survey. Additionally Liu *et al.* (2007) used the same regression function to study the relationships between SSPs and socio-demographic characteristics, self-reported health status and diseases, health conditions, and health behaviour using multiple regression for Taiwan[20].

---

[19] Hurd and McGarry(1995) included the following regressors: household income, wealth, age, marital status, gender, light physical activity, heavy physical activity, self-perceived health, race, smoking behaviour, drinking behaviour, educational level, high blood pressure, diabetes, cancer, chronic lung disease, heart problems, angina, congestive heart failure, stroke, arthritis and weight.

[20] Liu *et al.* (2007) used the following regression function: P = f (X, S, H1, H2, e) where the dependent variable P is the subjective probability of living to age 75 (P75) or 85 (P85).

The socio-demographic variables, represented by X, include age, gender, marital status, educational level, income, parental mortality, and whether the respondent lives with children. Subjective health status, S, is a vector of self-reported diseases and two measures of self-reported health status (current health compared with others of the same age and compared with the respondent's own health a year earlier). The vector H1 measures health conditions, including number of hospital admissions, number of visits to outpatient clinics, Hamermesh (2004) critized this methodology because he considers that using a subjective measure to proxy an objective determinant of a subjective outcome (as in the case of self-reported health status) can result in biased estimations. His logic is the following: if we are interested in the determinants of subjective outcome $S*$ by objective determinant B and controls X, and we have information on a large set of respondents indexed i, we wish to estimate:

$$S_i^* = \alpha X_i + \beta B_i + \varepsilon_i \qquad (2)$$

where ε_i is an error term. Because we do not observe B_i we have to rely on subjective proxy for it, BS_i. As any subjective proxy, BS_i may be a fairly good predictor of B_i but it will be affected by person-specific effect θ_i. This effect may include optimism or pessimism, for example. At the same time subjective response S_i^* will also be affected by a person-specific effect η_i. According to Hamermesh (2004) both person-specific effects are likely to be highly positively correlated.

We thus, have as observables:

$$BS_i = B_i + \theta_i \qquad (3)$$

$$S_i = S_i^* + \eta_i \qquad (4)$$

Then we are really estimating:

$$[S_i^* + \eta_i] = \alpha X_i + \beta[B_i + \theta_i] + \varepsilon_i \qquad (5)$$

insomnia, unhappiness, and obesity. Health behaviors, represented by H2, include exercise, smoking, drinking, and the habit of eating breakfast.

Hamermesh argues that given that it is expected that the correlation between η_i and θ_i is strongly positive, the estimation is likely to present bias. The author also considers that it is hard to believe that the positive correlations between the observed variables of interest that are induced by this problem are not large in nearly all cases.

We believe that even though there is a possibility that this problem may arise because of the inclusion of subjective measure (self-perceived health status) as dependent variable, it is also true that in our approach we also include relevant objective measures as health condition variables and health behaviour variables.

On top of that, this same approach has been extensively used. For instance, it has been sued by Hurd and McGarry (1995, 1997), Hurd *et al.* (2001), Liu *et al.* (2007), Balia (2007), Benitez-Silva and Ni (2008), Delavande and Rohwedeer (2008), among others for SSPs. This is a sign that Hamermesh's concern may not be a limitation for this type of methodology. We therefore consider appropriate to model SSP as a function of socio-demographic characteristics, health condition variables, health behaviour variables, self-reported health, and parental mortality experience.

Next, we describe in detail the set of variables used in the regression function to estimate the determinants of SSPs.

4.1.1. Socio-demographic characteristics

The characteristics taken into account in the analysis include age and age squared[21] in order to test non-linear relationship, gender (dummy variable equals one if male and two if female), and educational level defined using the seven categories of the ISCED international codes where 0 equals the lowest level of education and 6 represents the highest level[22]. We have also included a set of dummy variables for the marital status: married (which include legal marriage and also partnership), divorced, widowed and

[21] Age squared is divided by 100.

[22] ISCED CODE: 0 (no education), 1 (primary education or first stage of basic education), 2 (lower secondary or second stage of basic education), 3 (upper secondary education), 4 (post-secondary non-tertiary education), 5 (first stage of tertiary education), and 6 (second stage of tertiary education).

single.Other characteristics include the number of children, income quartiles[23], and finally, a set of dummy variables referring to the country where the respondent lives[24].

Previous studies have also considered socio-demographic variables as possible determinants of SSPs. For instance, Hurd and McGarry (1995) include household income, wealth, age, marital status, gender and race. Hurd and McGarry (1997) and Hurd *et al.* (2001) also include educational level. Finally, Liu *et al.* (2007) also includes a dummy variable con control for the presence of children in the household.

4.1.2. Health condition characteristics

A global measure of mental health was included: the EURO-D scale. It is constructed as the number of the following twelve symptoms that a respondent suffers: depression, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment, and tearfulness.

Several measures of physical health were also included: a dummy variable with the intention of controlling for the presence of a long term illness (1 if the respondent is suffering from a long-term illness, 0 otherwise) and the number of illnesses that the individual has suffered in the past or is currently suffering classified using medical criteria, into two categories, severe and non-severe[25].

Severe illnesses include a heart attack or any other heart problem, high blood pressure, high blood cholesterol, a stroke or cerebral vascular disease, diabetes, bronchitis or emphysema, malignant tumour, peptic ulcer, Parkinson disease, hip fracture or femoral fracture, and cancer in the brain, oral cavity, thyroid, lung, breast, oesophagus, stomach, liver, pancreas, kidney, cervix, endometrium, non-hodgkin lymphoma andleukemia. Non-severe illnesses include asthma, arthritis, osteoporosis, stomach ulcer, cataracts, and,

[23] We use the level of household income adjusted by purchasing power parity level in 2005 for waves one and two.

[24] Country dummies are used in both estimations (OLS and tobit) but not reported to save space.

[25] The difference between severe and non-severe illnesses is that the formers imply a significant decrease in life expectancy according to medical criteria.

cancers in some organs or part of the body like larynx, pharynx, prostate, testicle, ovary, colon or rectum, bladder and skin.

We have also defined the Global Activity Limitation Indicator (GALI) equals zero if the respondent's answer to the question "for the past six months at least, to what extent have you been limited because of a health problem in activities people usually do?" was not limited and equals one if the answer was severely limited or limited. The motivation to dichotomise this variable is the smaller numbers of severely limited when analysing data per country, gender and age groups.

Finally, two measures of the number of physical limitations were taken into account. First, activities of daily living (ADL) includes limitation when dressing, including putting on shoes and socks, walking across a room, bathing or showering, eating, such as cutting up your food , getting in and out of bed and using the toilet, including getting up or down.

Secondly, instrumental activities of daily living (IADL) includes limitation when using a map to figure out how to get around in a strange place, preparing a hot meal, shopping for groceries, making telephone calls, taking medication, doing work around the house or garden, and managing money, such as paying bills and keeping track of expenses.

All of the papers that we have reviewed that examine the determinants of SSP include health condition variables with different degree of detail. On the one hand, Hurd and McGarry (1995) only include the number of physical limitations. On the other hand, Hurd *et al.* (2001) considers the effect of specific illnesses, depression, cognition issues and physical limitations.

4.1.3. Health behaviour characteristics

This paper includes three classes of health behaviour variables: in the first place, smoking was included, but as opposed to previous studies, we have included dummy variables not only for two categories, smoker or non-smoker, but three: current smoker, former smoker, and non-smoker.

Consumption of alcohol behaviour was also included, using four dummy variables reflecting the frequency of the habit: never or on a daily, weekly or monthly basis. Finally, sport activity was defined as four if the answer to the question "how often do you engage in vigorous physical activity, such as sports?" was hardly ever or never, equals three if the answer was one to three times a month, equals two if once a week, and equals one if more than once a week.

Hurd and McGarry (1995) only include the smoking behaviour as a relevant determinant of SSP. Hurd and McGarry (1997) also considers the effect of drinking behaviour. Finally, Liu *et al.* (2007) also incorporates the effect of having breakfast every day.

4.1.4. Self-reported health

Individuals are asked to assess their own health into five categories: excellent, very good, good, fair, and bad; therefore a dummy variable has been constructed to reflect the different responses, where 1 equals excellent and 5 equals bad health[26].

All of the papers we have reviewed that examine the determinants of SSP include a self-reported health variable.

4.1.5. Parental Mortality experience

Previous studies as Hurd and McGarry (1995), Bloom *et al.* (2004) and Liu *et al.* (2007) have suggested that an individual's SSP is affected by parental mortality experience; also Feinstein (1993) stated that in the population the longevity of children increases with the longevity of their parents. We define two alternative models in order to control for this matter; the first one defines parental mortality equals one if the respondent has at least one parent dead and zero otherwise, irrespectively of the sex of the respondent and of the parent(s).

The second approach follows the Liu *et al.* (2007) view that men may be more influenced by the experience of their father's mortality, and women more by the experience of their

[26] Respondents in the first wave are asked to assess their own health twice. In the first question, known as US version, the available options are excellent, very good, good, fair and bad. In the second question, known as EU version, the options are very good, good, fair, bad and very bad. We have decided to use the US version because respondents in the second wave are only asked for this version.

mother's mortality, therefore, we interact gender and parental mortality experience in order to define 4 variables: male x father dead, male x mother dead, female x father dead and female x mother dead.

Table 5includes a brief description of all the variables used in the econometric analysis and their main statistics (mean and standard deviation in each wave)[27].

4.2. EMPIRICAL RESULTS. DETERMINANTS OF SSP

The results obtained for the Ordinary Least Square (OLS) estimation of the parameters, using a cross-sectional analysis for each of the two waves are presented in Table 6 and Table 7, respectively. The OLS specification has been previously used by Hurd and McGarry (1995, 1997), Hurd *et al.* (2001) and Liu *et al.* (2007)

Five models were specified for each wave and they differ in the explanatory variables included and in the treatment of parental mortality. In model 1we only include socio-demographic variables. In model 2 we also include both health condition and health behaviour variables. In addition, model 3 incorporated self-reported health status. Finally, in model 4 we also control for parental mortality experience irrespectively of the sex of the respondent and of the parent(s), while model 5 controls for "same-sex parental mortality experience" effect.

Similarly to Liu *et al.* (2007) we also report maximum likelihood estimates using a two-limit tobit model in Table 8 and Table 9 though the results are essentially the same as those in Table 6 and Table 7, correspondingly[28]. The motivation for using a tobit model is that reported survival chances should be limited to an interval [0-100] only and any value outside this interval would not make sense. Therefore we treat SSPs as a censored continuous variable. The censoring means that you don't have information below 0 or above 100, then using a two-limit tobit model (Long, 1997) could be sensible.

The two-limit tobit model, for respondents indexed i, would have the following form:

[27] Table 5 also includes a brief description and main statistics of other variables that are used in Chapter 2 and Chapter 3.

[28] In order to save space, Table 8 and Table 9 only report the results of model 4 and model 5.

$$SSP_i^* = \alpha + \beta X_i + \gamma H1_i + \delta H2_i + \eta S_i + \theta PM_i + \varepsilon_i \qquad (6)$$

$$SSP_i = \begin{cases} 0 & if \quad SSP_i^* < 0 \\ SSP_i^* & if \quad 0 \leq SSP_i^* \leq 100 \\ 100 & if \quad SSP_i^* > 100 \end{cases}$$

As shown in Table 6, we find a significant non-linear relationship between age and expected longevity for the first wave of SHARE (except when only socio-demographic variables are included), which is coherent with previous studies and with actuarial probabilities (life tables). Surprisingly, results for the second wave (Table 7) suggest non significance for age and age squared, though the signs of the coefficients coincide with those for the first wave.SSP analysis for Taiwan by Liu *et al.* (2007) also finds a significant non-linear relationship. Hurd and McGarry (1995) used a linear specification for the SSPs in the first wave of the HRS and found that age also has a positive significant effect on SSP, though its coefficient was smaller (0.4). Hurd and McGarry (1997) also used a liner specification for the second wave of the HRS and found similar results.

As we would expect *a priori*, European males report lower survival chances and this result is common to the two waves. It is also consistent with Hurd and McGarry (1995, 1997) finding for the US and with mortality data, as we know that males tend to live fewer years than females. On the other hand, Hurd *et al.* (2001) and Liu *et al.* (2007) found that survival expectations do not differ significantly by sex in Taiwan. Regarding marital status, our findings suggest that married, divorced and widowed report higher expected survival chances than singles for both waves, regardless of the model considered.

With respect to socio economic status, prior studies have estimated differential mortality by income and income inequality (Deaton and Paxton, 2004; Bommier *et al.*, 2006) and education (Feldman *et al.*, 1989; Lleras-Muney, 2005). The general conclusion is the existence of a strong correlation between the various measures of economic status and mortality. This could be related to the fact that more educated and richer individuals have access to a better health care.

We also find that more educated individuals report higher expected longevity. However, the effect of education decreases as more covariates are included. On average an extra level of education is associated with a range of 0.32-1.42 higher SSP which is consistent with mortality data. The same result was found by Hurd and McGarry (1995). Specifically, less than 12 years of education decreases SSP by 4.6% and more than 12 years of education increases SSP by 2.1% when self-perceived health status is not included in the regression. When it is included, the effect is not significant, though signs are coherent. Similarly, Liu *et al.* (2007) found no significant relationship but consistent signs. Hurd and McGarry (1997) found that having an education higher than high school on average increased expected longevity by 1.5%. Similar results were found by Hurd *et al.* (2001)

It is well known that richer individuals tend to live longer, and that is precisely what we observe from our results. We find positive significance on survival probability for the third and fourth quartile of income (with respect to the first one) for the first wave (except for model 1, where we also find a positive significant effect for the second quartile). For the second wave, we find positive significance only for the fourth quartile of income (except for model 1, where we also find a positive significant effect for the third quartile. Furthermore, results from model 5 suggest no significant effect for any quartile).Liu *et al.* (2007) found that personal disposable income is positively related to expected longevity. Contrary, Hurd and McGarry (1995, 1997), and Hurd *et al.* (2001) did not find a positive significant relationship.

When health variables (health conditions and health behaviour) are included in the analysis (model 2), generally, it is observed that the effect of socio-demographic characteristics diminishes. We find that mental health matters for expected longevity. For every additional Euro-D symptom, on average, reported SSP decreases between 1.4% and 1.8%, depending on the model. This result is common to the two waves and is coherent with epidemiological theory, as shown by Jacobson (2003) where she explores the physical consequences of depression stating that although the evidence of associations between depression and cardiovascular disease and diabetes is the strongest, data suggests links with osteoporosis, human immunodeficiency virus infection and the acquired immunodeficiency syndrome, stroke, and the prognosis in cancer. Likewise, Liu *et al.* (2007) also found that unhappiness reduces expected longevity by around 14%. Hurd *et al.* (2001) also found that depression decreases SSPs.

Another somewhat interesting result is that the presence of long-term illness produces a significant positive effect in the first and second wave, maybe due to the fact that the individual has already survived the illness for many years, so he/she is confident to reach age 75.

The number of illnesses that the respondent is suffering also has an effect on expected survival chances. We find that severe illnesses have a bigger effect than non-severe ones as one would expect. However, the marginal effect diminishes as self-perceived health and parental mortality variables are included. On average, every additional severe illness decreases SSP between2.7 and 1.9 in the first wave and between 2.2 and 1.5 in the second. Interestingly, non-severe illnesses have a significant effect only in the second wave (between -1.3 and -0.8).

On the subject of health behaviour variables, smoking is considered. According to Balia (2007) two alternative theories explain why individuals smoke. One theory defines smokers as irrational or myopic (Thaler and Sheffrin, 1981) so that they care more about present satisfaction than the future. The second theory (Becker and Murphy, 1988) states that smokers are instead, rational and forward looking, so that they internalize the effects of smoking. We find for both waves that current smokers report lower expected longevity, result that is consistent with previous studies and epidemiological data. For instance, Siddiqui *et al.* (2000) estimated that over the next 20 to 30 years, cigarette smoking will result in 10 million deaths annually on a worldwide basis, of which 70% will occur in developing countries. Interestingly, former smokers report higher SSP than people that have never smoked for both waves, regardless of the model considered. This may be due to the fact that former smokers have become more conscious about their health and habits and we are not controlling for that. Hurd and McGarry (1995, 1997) analysis suggests similar results. On the other hand, Liu *et al.* (2007) found no significant effect of smoking on expected longevity.

We have found that practicing sports, as we might expect, increases SSP, as shown by medical studies such as Hambrecht *et al.* (2000) and Tall (2002). Similarly Hurd and McGarry (1997) found similar results. However, Liu *et al.* (2007) did not find a significant relationship.

In epidemiological data, moderate drinking is associated with greater longevity than complete abstinence; examples can be found in the economic literature (Sharper, 1990; Boffeta and Garfinkel, 1990; Ellison, 1990), but also in medical literature, for instance, Saitz (2005) stated that moderate consumption of alcohol may be beneficial, but what constitutes "moderate" depends on age, sex, genetic characteristics, coexisting illnesses, and other factors. Particularly, weekly intakes of no more than five drinks for men or two drinks for women are associated with the lowest mortality. Evans and Bienias (2005) concluded that as compared with abstinence, low-to-moderate consumption of alcohol was associated with better cognition both cross-sectionally and over time in a study that involved more than 11,000 US nurses.

That is precisely what is observed from the results for the analysis undertaken in both waves of SHARE; Consuming alcohol whether it is on a daily, weekly or monthly basis, has a positive effect on the reported expected longevity, though weekly drinking represents the biggest effect. Hurd and McGarry (1995, 1997) also found that moderate drinking is related with higher expected longevity.

When an individual's self-reported health status is included in the analysis (model 3) it has the effect of diminishing the impact of health condition and health behaviour variables. It also produces a major positive effect on SSP; other things being equal, an individual with excellent health reports on average almost 19.5% more chances of reaching age 75 than someone claiming bad health in the first wave, and almost 14.5% in the second wave. Hurd and McGarry (1995, 1997) for the US and Liu *et al.* (2007) for Taiwan also found very similar results.

In relation to parental mortality experience, we find that its inclusion does not decrease considerably the impact of health condition, health behaviour and self-perceived variables. We find that respondents are negatively affected by the death of at least one parent in both waves. More importantly, the hypothesis of the "same-sex parent" seems to hold for SHARE respondents; this may be due to the fact that individuals have the view that the degree of the genetic influences from parents to offspring is sex dependent.

We find that males with a dead father report lower SSP than males with a dead mother in the two waves. Opposite to this, females with dead mother report lower SSP than females with a dead father in the first wave. Liu *et al.* (2007) also obtained evidence that this hypothesis holds for Taiwan. They found that being male (female) with a dead father (mother) reduces significantly SSP by 7.6% (4.8%) while being male (female) with a dead mother (father) does not significantly reduce SSP. Hurd and McGarry (1997) does not differentiate by sex but found that having a father or a mother alive increased expected longevity by 4.5% and 4.6% respectively.

The fact that results do not vary significantly across waves points towards the idea that subjective beliefs of SHARE respondents are reliable and that the results obtained are consistent across waves.

5. CONCLUSIONS

Expectations provide very important information that economists have historically disregarded. They were more interested in what people do rather than what people think. Fortunately, this has changed in the last 20 years and economists have realized that expectations are fundamental when testing basic theories of economics.

But before using expectations as an ingredient of economic theories as LCM and assessing whether individuals take into account their own expectations when taking economic decisions, it is fundamental to take one step backwards and analyse the consistency of those expectations. Specifically we concentrate on SSPs for SHARE database.

SSPs have previously been analysed. Hurd and McGarry (1995, 1997) used HRS cross-sectional data to analyse external and internal consistency of SSP. Hurd *et al.* (2001) used AHEAD to analysed the determinants of SSPs. Liu *et al.* (2007) used data drawn from the BNHI to examine the determinants of Taiwanese SSPs. But expected longevity drawn from SHARE has not been analysed in as much detail as in the HRS.

The objective of this Chapter was to assess if SSPs were: (i) consistent across waves; (ii) they average close to objective probabilities; (iii) consistent with mortality data and epidemiological data; and (iv) consistent with economic literature.

The fact that our results do not vary significantly across waves points towards the idea that subjective beliefs of SHARE respondents are reliable and that the results obtained are consistent across waves. Furthermore, our statistical analysis shows that SSPs present less severe focal point answers with respect to HRS and ELSA surveys.

With the help of life tables we have concluded that on average, European females persistently underreport their probabilities of reaching 75 while European males on average, over report their survival chances especially when younger; this could imply that individuals, especially females, are not always aware of their objective survival chances. However, this result should be taken with careful because direct comparison between SSPs and life table is only reasonable if we suppose that mortality risk is stationary over time and there is no heterogeneity in the population.

In order to test their consistency with mortality and epidemiological data, we have also undertaken a cross-sectional econometric analysis in order to examine the determinants of SHARE's SSPs. Our analysis has shown that reported life expectancy is coherent with mortality facts related with sex, income, education and country, but also with epidemiological data. We have found that both, physical and mental health matter for expected longevity and that subjective health status and healthy life-style have an effect on our perception of survival chances. Finally, we have shown that the "same-sex parent" hypothesis seems to hold for SHARE respondents.

We have also found that similar conclusions were obtained by Hurd and McGarry (1995) for the first wave of the HRS, Hurd and McGarry (1997) for the second wave of the HRS, Hurd *et al.* (2001) for the first two waves of the AHEAD survey, and by Liu *et al.* (2007) for Taiwan.

Overall, we have found that, using cross-sectional analysis, SSPs in SHARE do behave reasonably well. They average close to actuarial probabilities and they covary with other variables in the same way actual outcomes vary with the variables. Moreover, our results are in line with economic literature based on alternative data.

As mentioned in the Introduction, Hurd and McGarry (1997) stated that before using SSPs as inputs of models of decision-making under uncertainty we need to test their properties. Concretely, the authors mention three properties that should be satisfied:

1. Individuals understand questions about probabilities and accurately report their beliefs about the likelihood of future events;
2. Individuals adjust their reported probabilities in response to new information; and
3. The reported probabilities predict outcome.

This Chapter focused on the first property and concluded that SSPs elicited from SHARE do behave reasonably well. The next Chapter will focus on the next two properties.

6. REFERENCES

- Abu-Rayya, H. M. (2006). "Depression and Social Involvement among Elders". *The Internet Journal of Health*, 5(1): 8-8.

- Alavinia, S. M., Burdorf, A. (2008). "Unemployment and Retirement and Ill-Health: a Cross-Sectional Analysis across European Countries". *International Archives of Occupational Environmental Health*, 82: 39-45.

- Andreyeva, T. (2006). *An International Comparison of Obesity in Older Adults*, Virginia, Pardee RAND Graduate School: 157.

- Angelini, V., Laferrère, A. (2008). "Home, Houses and Residential Mobility". In: Börsch-Supan, A., Brugiavini, A., Jürges, H.,Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G., *Health, Ageing and Retirement in Europe (2004-2007) –Starting the Longitudinal Dimension*, Mannheim, MEA (Mannheim Research Institute for the Economics of Aging): 99-107.

- Balia, S. (2007). "Reporting Expected Longevity and Smoking: Evidence from Share". Centro Ricerche Economiche Nord Sud (CRENOS), *Working Paper,* 2007/05.

- Banks, J., Emmerson, C. (2004). "Not so Brief Lives: Longevity Expectations and Wellbeing in Retirement". In: Banks, J., Emmerson, C., Oldfield, Z., *Seven Ages of Man and Woman: A Look at Life in Britain in the Second Elizabeth Era*, London, Economic and Social Research Council: 28-31.

- Becker, G. S., Murphy, K. (1988). "A Theory of Rational Addiction". *Journal of Political Economy*, 96(4): 675-700.

- Benitez-Silva, H., Ni, H. (2008). "Health Status and Health Dynamics in an Empirical Model of Expected Longevity". *Journal of Health Economics,* 27(3): 564-584.

- Blanchet, D., Brugiavini, A., Rainato, R. (2005). "Pathways to Retirement". In: Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G., *Health, Ageing and Retirement in Europe - First Results from the Survey of Health, Ageing and Retirement in Europe*, Mannheim, MEA: 246-252.

- Bloom, D., Canning, D., Graham, B. (2003). "Longevity and Life-Cycle Savings". *Scandinavian Journal of Economics*, 105: 319-38.

- Bloom, D., Canning, D., Moore, M. (2004). "The Effect of Improvements in Health and Longevity on Optimal Retirement and Saving". *NBER Working Paper* 10919, National Bureau of Economic Research, Cambridge, MA, August.

- Boffeta, P., Garfinkel, L. (1990). "Alcohol Drinking and Mortality among Men Enrolled in an American Cancer Society Prospective Study". *Epidemiology*, 1(5): 343-348.

- Bommier, A., Magnac, T., Rapoport, B., Roger, M. (2006). "Droit a la Retraite et MortalitéDifférentielle". *Économie et Prévision*, 168: 1-16.

- Borawski E., Kinney J. M., Kahana E. (1996). "The Meaning of Olderadults' Health Appraisals: Congruence with Health Status and Determinant of Mortality". *Journal of Gerontology: Social Sciences,* 5 IB: S157-S170.

- Carta, E. (2007). *Causes of 'Involuntary' and 'Voluntary' Early Retirement in Italy. A Quantitative Analysis from SHARE*. University of Surrey.

- Crespo, L., Mira, P. (2008). "Caring for Parents and Employment Status of European Middle-Aged Women". In: Börsch-Supan, A. Brugiavini, H. Jürges, A. Kapteyn, J. Mackenbach, J. Siegrist, and G. Weber (eds.), *Health, Ageing and Retirement in Europe (2004-2007). Starting the Longitudinal Dimension*, Mannheim, MEA: 261-266.

- Deaton, A., Paxson, C. (2004). "Mortality, Income, and Income Inequality over Time in Britain and the U.S.". In: Wise, D. (ed.), *Perspectives in the Economics of Aging*, Chicago, University of Chicago Press: 247-280.

- Delavande A., Rohwedeer S. (2008). "Differential Mortality in Europe and the U.S.". *RAND Labour and Population Working Paper*, 613.

- Dominitz, J., Manski, C. F. (1997a). "Using Expectations Data to Study Subjective Income Expectations". *Journal of the American Statistical Association,* 92: 855-862.

- Ellison, R. (1990). "Cheers!". *Epidemiology*, 1(5): 337-339.

- Erlinghagen, M., Hank, K. (2006). "The Participation of Older Europeans in Volunteer Work". *Ageing and Society,* 26(4): 567-584.

- Evans, D., Bienias, J. (2005). "Alcohol Consumption and Cognition". *The New England Journal of Medicine*, 351: 289-291.

- Feinstein, J. (1993). "The Relationship between Socioeconomic Status and Health: A Review of the Literature". *The Milbank Quarterly*, 71: 279-322.

- Feldman, J., Makuc, D., Kleinman, J., Cornoni-Huntley, J. (1989). "National Trends in Educational Differences in Mortality". *American Journal of Epidemiology*, 128(8): 919-933.

- Fokkema, T., terBekke, S., Dykstra, P. A. (2008). *Solidarity between Parents and their Adult Children in Europe*. Amsterdam, Netherlands Interdisciplinary Demographic Institute, 112.

- Friedman, M. (1957). "The Permanent Income Hypothesis". In: Friedman, M., *A Theory of the Consumption Function*, Princeton, Princeton University Press: 20-37.

- Hambrecht, R., Wolf, A., Gielen, S., Linke, A., Hofer, J. (2000). "Effect of Exercise on Coronary Endothelial Function in Patients with Coronary Artery Disease". *The New England Journal of Medicine*, 321: 452-460.

- Hamermesh, D. S., Hamermesh, F.W. (1983). "Does Perception of Life Expectancy Reflect Health Knowledge?". *American Journal of Public Health*, 73(8): 911-914.

- Hamermesh, D. S. (1985). "Expectation, Life Expectancy and Economic Behaviour". *Quarterly Journal of Economics*, 100(2): 389-408.

- Hamermesh, D. S. (2004). "Subjective Outcomes in Economics". *Southern Economic Journal,* 71(1): 2-11.

- Hurd, M., McGarry, K. (1995). "Evaluation of the Subjective Probabilities of Survival in the Health and Retirement Study". *Journal of Human Resources*, 30: S7-S56.

- Hurd, M., McGarry, K. (1997). "The Predictive Validity of Subjective Probabilities of Survival". *The Economic Journal,* 112(482): 966-985.

- Hurd, M., Mcfadden, D., Merrill, A. (2001). "Predictors of Mortality among the Elderly". In: Wise, D. (ed.), *Themes in the Economics of Aging*, Chicago, University of Chicago Press: 171-198.

- Hurd, M.,Rohwedder, S., Winter, J. (2008). "Individual's Expectations about Longevity: An International Comparison". CeMMAP Workshop on "Novel Measurement Methods for Understanding Economic Behaviour". IFS, London, 3 July, 2008.

- Jacobson, S. (2003). "Physical Consequences of Depression". *The New England Journal of Medicine*, 350: 569-573.

- Juster, T. (1966). "Consumer Buying Intentions and Purchase Probability: an Experiment in Survey Design". *Journal of the American Statistical Association*, 61: 658-696.

- Liu, J., Tsou, M., Hammitt, J. (2007). "Health Information and Subjective Survival Probability: Evidence from Taiwan". *Journal of Risk Research,* 10(2): 149-175.

- Lleras-Muney, A. (2005). "The Relationship between Education and Adult Mortality in the United States". *Review of Economic Studies*, 72: 189–221.

- Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*. ThousandOaks, Sage Publications: 297.

- Manski, C. F. (1990). "The Use of Intentions Data to Predict Behaviour: a Best-Case Analysis". *Journal of the American Statistical Association*, 85: 934-940.

- Manski, C. F. (2004). "Measuring Expectations". *Econometrica*, 72(5): 1329-1376.

- Modigliani, F., Brumberg, R. H. (1954), "Utility Analysis and the Consumption Function: an Interpretation of Cross-Section Data". In: Kurihara, K. K. (ed.), *Post-Keynesian Economics,* New Brunswick, Rutgers University Press: 388-436.

- O'Donnell, O., Teppa, F., van Doorslaer, E. (2008). "Can Subjective Expectations Explain Retirement Behaviour?". *Working Paper,* 188, De Nederlansche Bank.

- Pagán, R. (2009). "Part-Time Work of Older Workers with Disabilities in Europe". *Public Health,* 123(5): 378-383.

- Peytremann-Bridevaux, I., Santos-Eggimann, B. (2008). "Healthcare and Preventive Services Utilization of Elderly Europeans with Depressive Symptoms". *Journal of Affective Disorders,* 105: 247-252.

- Pijls, L., Feskens, E. J. M., Kromhout, D. (1993). "Self-Rated Health, Mortality, and Chronic Diseases in Elderly Men: The Zutphen study. 1985-1990". *American Journal of Epidemiology,* 138: 840-848.

- Post, T., Hanewald, K. (2011). "Longevity Risk, Subjective Survival Expectations and Individual Saving Decisions". NETSPAR *Working Paper,* 043.

- Rueda, S., Artazcoz, L. (2008). "Health Inequalities among the Elderly in Western Europe". *Journal of Epidemiology and Community Health,* 62(6): 492-498.

- Saitz, R. (2005). "Unhealthy Alcohol Use". *The New England Journal of Medicine*, 352: 596-607.

- Sharper, A. (1990). "Alcohol and Mortality: A Review of Prospective Studies". *British Journal of Addiction*, 85: 837-847.

- Siddiqui, S., Ogbeide, R., Al Khalifa, I. (2000). "Smoking in a Saudi Community: Prevalence, Influencing Factors, and Risk Perception". *American Academy of Family Physicians 2000 Scientific Assembly*, 33(5): 367-370.

- Siegel, M., Bradley, E., Kasl, S. (2003). "Self-Rated Life Expectancy as a Predictor of Mortality: Evidence from the HRS and AHEAD surveys". *Gerontology*, 49(4): 265-271.

- Smith, K., Taylor, D., Sloan, F. (2001). "Longevity Expectations and Death: Can People Predict their own Demise?".*The American Economic Review*, 91(4): 1126-1134.

- Solé-Auró, A.,Crimmins, E.M. (2008). "Health of Immigrants in European Countries". *International Migration Review,* 42(4): 861-876.

- Steffen, B. (2009), "Formation and Updating of Subjective Life Expectancy: Evidence from Germany", *MEA Studies*, 8.

- Sundström, G., Fransson, E. Malmberg, B., Davey, A. (2009). "Loneliness Among Older Europeans". *European Journal of Ageing,* 6(4): 267-275.

- Tall, A. (2002). "Exercise to Reduce Cardiovascular Risk- How much is enough?". *The New England Journal of Medicine*, 347: 1522-1524.

- Thaler, R., Sheffrin, H. (1981). "An Economic Theory of Self-Control". *Journal of Political Economy*, 89(2): 392-406.

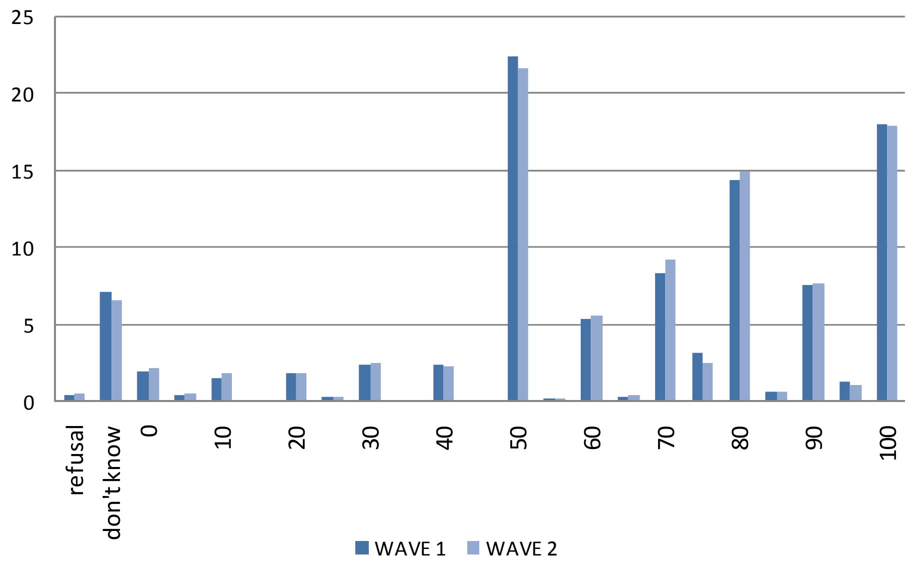
- Van Doorn, C., Kasl, S. (1998). "Can Parental Longevity and Self-Rated Life Expectancy Predict Mortality among Older Persons? Results from an Australian Cohort". *Journal of Gerontology: Social Sciences*, 53B(1): S28-S34.

- Viscusi, W., Hakes, J. (2003). "Risk Rating That Do Not Measure Probabilities". *Journal of Risk Research*, 6: 23-43.

- Wolfe, J. (1983). "Perceived Longevity and Early Retirement". *The Review of Economics and Statistics*, 65(4): 544-551.

7. TABLES AND FIGURES

FIGURE 1: DISTRIBUTION OF SSP



Source: own elaboration

TABLE 1: TARGET AGES IN SSP QUESTION

Respondent's age	**Target age**
45-65	75
66-70	80
71 – 75	85
76 – 80	90
81 – 85	95
86 – 90	100
91 – 95	105
96 – 100	105
101 – 105	110
106 and older	120

Source: own elaboration

TABLE 2: SSP AND LIFE TABLE BY COUNTRY AND GENDER

	Females		Males	
	SSP	**Life table**	**SSP**	**Life table**
Switzerland	73.0	84.3	72.2	71.8
Sweden	73.2	81.4	69.8	70.7
Spain	72.5	84.6	71.2	67.7
Netherlands	72.9	79.2	69.9	65.8
Italy	68.6	83.2	72.2	68.5
Greece	63.3	79.5	66.0	65.3
Germany	67.7	80.0	66.8	64.1
France	68.2	84.2	68.6	67.5
Denmark	75.9	73.4	70.5	62.7
Belgium	64.5	80.6	62.5	68.0
Austria	65.7	81.3	64.5	69.1
Czech Republic	50.4	75.9	49.6	59.6
Poland	55.9	75.5	52.3	62.5
Ireland	74.5	80.1	70.6	68.6

Source: own elaboration

TABLE 3: SSP AND LIFE TABLE BY AGE AND GENDER

	Females		Males	
Age	**SSP**	**Life table**	**SSP**	**Life table**
45	65.7	77.8	74.6	61.8
46	66.9	77.9	76.4	62.0
47	65.3	78.1	64.5	62.2
48	66.3	78.2	64.5	62.4
49	67.3	78.4	67.7	62.7
50	69.4	78.5	65.8	62.9
51	67.9	78.7	65.4	63.2
52	68.6	78.9	66.9	63.5
53	69.0	79.2	67.8	63.9
54	68.7	79.4	67.1	64.3
55	68.4	79.7	67.8	64.7
56	68.4	80.0	67.5	65.2
57	69.3	80.3	67.6	65.7
58	69.7	80.6	67.4	66.2
59	68.4	81.0	67.8	66.9
60	68.3	81.4	68.5	67.5
61	69.6	81.9	69.7	68.3
62	69.4	82.4	66.7	69.1
63	69.0	82.9	70.2	70.0
64	67.6	83.5	69.8	71.1
65	68.8	84.2	68.9	72.2

Source: own elaboration

TABLE 4: OLS REGRESSION SSP AND LIFE TABLE

SSP	Males		Females		All	
	Coefficient	**Std. Dev.**	**Coefficient**	**Std. Dev.**	**Coefficient**	**Std. Dev.**
Constant	0.361***	0.058	0.797***	0.070	0.628***	0.020
Life table	0.475***	0.086	-0.133	0.086	0.076***	0.026
R^2	0.005		0.001		0.001	

Note: Robust standard errors in parenthesis. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

TABLE 5: VARIABLES

		First wave		Second wave	
Variable	**Description**	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
SSP	Reported Subjective Survival Probability	68.46	25.08	67.92	25.56
Age	Age of the respondent	56.78	4.90	57.14	4.70
Gender	Dummy (1 if male; 2 if female)	1.55		1.56	
Married	Dummy (1 if married; 0 otherwise)	0.81		0.82	
Divorced	Dummy (1 if divorced; 0 otherwise)	0.08		0.08	
Widowed	Dummy (1 if widowed; 0 otherwise)	0.05		0.05	
Single	Dummy (1 if single; 0 otherwise)	0.06		0.05	
Austria	Dummy (1 if Austrian; 0 otherwise)	0.07		0.04	
Germany	Dummy (1 if German; 0 otherwise)	0.11		0.08	
Sweden	Dummy (1 if Sweden; 0 otherwise)	0.11		0.08	
Netherlands	Dummy (1 if Dutch; 0 otherwise)	0.12		0.09	
Italy	Dummy (1 if Italian; 0 otherwise)	0.09		0.08	
France	Dummy (1 if French; 0 otherwise)	0.10		0.08	
Denmark	Dummy (1 if Danish; 0 otherwise)	0.06		0.09	
Spain	Dummy (1 if Spaniard; 0 otherwise)	0.07		0.05	
Greece	Dummy (1 if Greek; 0 otherwise)	0.11		0.10	
Switzerland	Dummy (1 if Swiss; 0 otherwise)	0.04		0.05	
Belgium	Dummy (1 if Belgian; 0 otherwise)	0.14		0.10	
Czech	Dummy (1 if Czech; 0 otherwise)			0.08	
Poland	Dummy (1 if Polish; 0 otherwise)			0.08	
Education	Educational level (ISCED international codes where 0 is the lowest level and 6 the highest)	2.84	1.48	2.92	1.42
Income ('000)	Household income in Euros adjusted by PPP level in 2005	51.1	147.8	48.8	310.7
Wealth ('000)	Household wealth in million Euros adjusted by PPP level in 2005	366.9	1,045.2	315.4	1,082.1
Children	Number of children	2.09	1.30	2.12	1.24
Euro-D	Number of the mental health symptoms that a respondent suffers	2.07	2.10	2.06	2.12
Long-term illness	Dummy (1 if respondent suffers a long-term illness; 0 otherwise)	3.35		3.38	
Adl	Number of limitations in activities of daily living (Adl)	0.08	0.46	0.09	0.47

Iadl	Number of limitations in instrumental activities of daily living (Iadl)	0.12	0.51	0.12	0.51
Severe illness	Number of severe illnesses suffering	0.71	0.96	0.72	0.96
Non-severe illness	Number of non-severe illnesses suffering	0.33	0.61	0.33	0.61
GALI	Dummy (1 if health limits respondents; 0 otherwise)	0.33		0.34	
Never smoker	Dummy (1 if never smoker; 0 otherwise)	0.47		0.46	
Former smoker	Dummy (1 if former smokers; 0 otherwise)	0.28		0.26	
Current smoker	Dummy (1 if current smoker; 0 otherwise)	0.26		0.27	
Never drinking	Dummy (1 if never drinking ; 0 otherwise)	0.22		0.24	
Daily drinking	Dummy (1 if daily drinking; 0 otherwise)	0.22		0.19	
Weekly drinking	Dummy (1 if weekly drinking; 0 otherwise)	0.35		0.34	
Monthly drinking	Dummy (1 if monthly drinking; 0 otherwise)	0.21		0.23	
Sports	Equals 4 if practicing sports hardly ever or never, 3 if one to three times a month, 2 if once a week, and 1 if more than once a week.	2.28	1.29	2.33	1.30
S-perceived health	1 (excellent), 2 (very good), 3 (good), 4 (fair), and 5 (bad)	2.71	1.03	2.85	1.07
Male x father dead	Dummy (1 if respondent is male and his father is dead; 0 otherwise)	0.38		0.37	
Male x mother dead	Dummy (1 if respondent is male and his mother is dead; 0 otherwise)	0.28		0.28	
Female x mother dead	Dummy (1 if respondent is female and her mother is dead; 0 otherwise)	0.33		0.34	
Female x father dead	Dummy (1 if respondent is female and her father is dead; 0 otherwise)	0.45		0.47	
Parental mortality	Dummy (1 if respondent has at least one parent dead; 0 otherwise)	0.88		0.89	
Receiving inheritance	Reported probability of receiving an inheritance	22.7	34.8	23.6	32.1

Source: own elaboration

TABLE 6: SSP MODEL (OLS SPECIFICATION)

SSP wave 1	Model 1 (controlling for X)			Model 2 (X, H1, and H2)			Model 3 (X, H1, H2, and S)			Model 4 (X, H1, H2, S and parental mortality)			Model 5 (X, H1, H2, S and same sex parental mortality)		
	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.
Age	1.55	*	0.93	1.93	**	0.89	2.23	**	0.88	2.69	***	0.88	2.69	***	0.89
Age squared/100	-1.18		0.82	-1.42	*	0.79	-1.69	**	0.78	-2.04	***	0.78	-1.98	**	0.78
Gender	1.45	***	0.41	3.65	***	0.43	3.30	***	0.42	3.29	***	0.42	2.87	***	0.95
Married	4.73	***	0.98	3.16	***	0.92	3.27	***	0.92	3.28	***	0.92	3.25	***	0.91
Divorced	3.94	***	1.19	4.01	***	1.12	3.83	***	1.11	3.86	***	1.11	3.81	***	1.10
Widowed	3.88	***	1.32	3.57	***	1.26	3.38	***	1.24	3.40	***	1.24	3.44	***	1.24
Education	1.42	***	0.15	0.68	***	0.15	0.38	***	0.15	0.36	**	0.15	0.32	**	0.15
IncomeQ2	1.49	**	0.69	0.62		0.66	0.61		0.66	0.61		0.66	0.59		0.65
IncomeQ3	3.84	***	0.66	2.33	***	0.63	2.01	***	0.62	1.99	***	0.62	1.95	***	0.62
IncomeQ4	4.39	***	0.68	2.40	***	0.65	1.76	***	0.65	1.75	***	0.65	1.72	***	0.65
Children	-0.10		0.18	0.27		0.17	0.25		0.17	0.24		0.17	0.23		0.17
Euro-D				-1.83	***	0.11	-1.45	***	0.11	-1.45	***	0.11	-1.46	***	0.11
Long-term illness				-0.73		0.49	1.24	**	0.49	1.23	**	0.49	1.23	**	0.49
Adl				-1.63	**	0.65	-1.10	*	0.64	-1.10	*	0.64	-1.06	*	0.64

Iadl			-1.85 ***	0.53	-1.25 **	0.52	-1.27 **	0.52	-1.31 **	0.52
Severe illness			-2.76 ***	0.24	-1.98 ***	0.24	-1.96 ***	0.24	-1.93 ***	0.24
Non-severe illness			-0.22	0.38	0.42	0.38	0.41	0.38	0.38	0.38
GALI			-3.58 ***	0.53	-1.20 **	0.54	-1.20 **	0.54	-1.19 **	0.54
Former smoker			1.95 ***	0.48	1.90 ***	0.47	1.92 ***	0.47	1.93 ***	0.47
Current smoker			-2.34 ***	0.51	-1.92 ***	0.50	-1.88 ***	0.50	-1.88 ***	0.50
Daily drinking			2.08 ***	0.65	1.59 **	0.64	1.57 **	0.64	1.52 **	0.64
Weekly drinking			2.77 ***	0.58	2.13 ***	0.58	2.11 ***	0.58	2.06 ***	0.58
Monthly drinking			1.77 ***	0.62	1.40 **	0.62	1.40 **	0.62	1.34 **	0.62
Sports			-0.67 ***	0.16	-0.32 **	0.16	-0.33 **	0.16	-0.32 **	0.16
Self-perceived health					-5.00 ***	0.25	-4.94 ***	0.25	-4.90 ***	0.25
Parental mortality							-3.06 ***	0.60		
Male x father dead									-3.53 ***	0.78
Male x mother dead									-1.01 *	0.61
Female x mother dead									-2.88 ***	0.55
Female x father dead									-1.65 **	0.70
Constant	5.71	26.39	0.63	25.31	3.67	24.95	-8.20	25.01	-9.06	25.27
R-squared	0.034		0.112		0.135		0.137		0.139	
N	15,058									

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1. Source: own elaboration

TABLE 7: SSP MODEL (OLS SPECIFICATION)

SSP wave 2	Model 1 (controlling for X)			Model 2 (X, H1, and H2)			Model 3 (X, H1, H2, and S)			Model 4 (X, H1, H2, S and parental mortality)			Model 5 (X, H1, H2, S and same sex parental mortality)		
	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.
Age	0.51		0.93	0.49		0.89	0.66		0.89	1.07		0.89	0.94		0.89
Age squared/100	-0.32		0.82	-0.20		0.79	-0.33		0.78	-0.64		0.78	-0.47		0.79
Gender	1.52	***	0.39	3.61	***	0.40	3.31	***	0.39	3.34	***	0.39	2.51	***	0.87
Married	3.60	***	0.95	2.50	***	0.90	2.39	***	0.89	2.37	***	0.89	2.30	***	0.89
Divorced	1.94	*	1.13	2.21	**	1.07	2.04	*	1.06	2.03	*	1.06	1.90	*	1.06
Widowed	2.32	*	1.26	2.68	**	1.20	2.35	**	1.19	2.34	**	1.19	2.24	*	1.19
Education	1.37	***	0.15	0.65	***	0.14	0.43	***	0.14	0.40	***	0.14	0.37	***	0.14
IncomeQ2	0.52		0.63	-0.25		0.60	-0.23		0.60	-0.26		0.60	-0.31		0.60
IncomeQ3	2.13	***	0.62	0.67		0.60	0.57		0.59	0.55		0.59	0.51		0.59
IncomeQ4	3.52	***	0.63	1.50	**	0.60	1.00	*	0.60	0.99	*	0.60	0.94		0.60
Children	-0.15		0.17	-0.02		0.16	-0.04		0.16	-0.04		0.16	-0.04		0.16
Euro-D				-1.80	***	0.11	-1.49	***	0.11	-1.49	***	0.11	-1.50	***	0.11
Long-term illness				-2.59	***	0.48	-1.01	**	0.49	-1.02	**	0.49	-1.02	**	0.49
Adl				0.33		0.58	0.66		0.58	0.63		0.58	0.65		0.58
Iadl				-3.00	***	0.56	-2.79	***	0.56	-2.79	***	0.56	-2.78	***	0.56

Severe illness			-2.18	***	0.23	-1.58	***	0.23	-1.55	***	0.23	-1.53	***	0.23
Non-severe illness			-1.29	***	0.37	-0.80	**	0.37	-0.80	**	0.37	-0.81	**	0.37
GALI			-1.92	***	0.51	-0.20		0.52	-0.22		0.52	-0.25		0.52
Former smoker			1.00	**	0.45	1.01	**	0.45	1.06	**	0.45	1.06	**	0.45
Current smoker			-2.89	***	0.46	-2.57	***	0.46	-2.48	***	0.46	-2.44	***	0.46
Daily drinking			2.23	***	0.62	1.84	***	0.62	1.83	***	0.62	1.82	***	0.62
Weekly drinking			2.41	***	0.55	1.95	***	0.54	1.96	***	0.54	1.89	***	0.54
Monthly drinking			1.17	**	0.57	0.99	*	0.56	0.97	*	0.56	0.95	*	0.56
Sports			-0.73	***	0.15	-0.46	***	0.15	-0.46	***	0.15	-0.47	***	0.15
Self-perceived health						-3.70	***	0.24	-3.66	***	0.24	-3.60	***	0.24
Parental mortality									-2.84	***	0.55			
Male x father dead												-2.75	***	0.74
Male x mother dead												-2.43	***	0.60
Female x mother dead												-1.58	***	0.52
Female x father dead												-2.39	***	0.65
Constant	38.99	26.45	46.05	*	25.35	50.04	**	25.19	39.31		25.33	42.91	*	25.51
R-squared	0.104		0.178			0.191			0.192			0.193		
N	16,485													

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

TABLE 8: SSP MODEL (TOBIT SPECIFICATION)

SSP wave 1	Model 4 (X, H1, H2, S and parental mortality)			Model 5 (X, H1, H2, S and same-sex parental mortality)		
	Coefficient		Std. Dev.	Coefficient		Std. Dev.
Age	3.133	***	1.105	3.152	***	1.107
Age squared/100	-2.352	**	0.973	-2.298	**	0.974
Gender	3.921	***	0.527	3.470	***	1.238
Married	4.633	***	1.104	4.587	***	1.102
Divorced	5.016	***	1.340	4.958	***	1.339
Widowed	4.975	***	1.491	5.015	***	1.489
Education	0.285		0.183	0.239		0.183
IncomeQ2	0.770		0.769	0.750		0.768
IncomeQ3	2.316	***	0.760	2.271	***	0.759
IncomeQ4	2.049	***	0.785	2.014	**	0.784
Children	0.267		0.197	0.260		0.197
Euro-D	-1.699	***	0.129	-1.710	***	0.129
Long-term illness	1.561	**	0.622	1.565	**	0.622
Adl	-1.193	*	0.634	-1.146	*	0.633
Iadl	-1.471	**	0.579	-1.506	***	0.579
Severe illness	-2.201	***	0.275	-2.156	***	0.275
Non-severe illness	0.366		0.433	0.323		0.432
GALI	-1.238	*	0.654	-1.233	*	0.654
Former smoker	2.451	***	0.604	2.454	***	0.604
Current smoker	-1.670	***	0.610	-1.669	***	0.609
Daily drinking	1.635	**	0.775	1.576	**	0.775
Weekly drinking	2.324	***	0.707	2.271	***	0.706
Monthly drinking	1.620	**	0.747	1.546	**	0.746
Sports	-0.431	**	0.194	-0.421	**	0.194
Self-perceived health	-5.974	***	0.310	-5.927	***	0.310

Parental mortality	-3.430 ***	0.783		
Male x father dead			-4.066 ***	1.018
Male x mother dead			-1.079	0.773
Female x mother dead			-3.336 ***	0.700
Female x father dead			-1.874 **	0.891
Constant	-11.103	31.291	-12.397	31.549
Sigma	28.373	0.194	28.348	0.195
Log-likelihood	-60,040.7		-60,025.4	
N	15,058			

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

TABLE 9: SSP MODEL (TOBITS PECIFICATION)

SSP wave 2	Model 4 (X, H1, H2, S and parental mortality)		Model 5 (X, H1, H2, S and same-sex parental mortality)	
	Coefficient	Std. Dev.	Coefficient	Std. Dev.
Age	0.883	1.109	0.767	1.111
Age squared/100	-0.430	0.973	-0.257	0.975
Gender	3.795 ***	0.496	3.061 **	1.207
Married	2.869 ***	1.092	2.793 **	1.091
Divorced	2.547 *	1.301	2.402 *	1.300
Widowed	2.870 **	1.440	2.758 *	1.439
Education	0.313 *	0.178	0.276	0.178
IncomeQ2	-0.443	0.711	-0.500	0.711
IncomeQ3	0.658	0.715	0.605	0.715
IncomeQ4	1.255 *	0.737	1.199	0.736
Children	-0.014	0.192	-0.020	0.192
Euro-D	-1.735 ***	0.123	-1.742 ***	0.123
Long-term illness	-1.284 **	0.597	-1.285 **	0.597
Adl	0.906	0.592	0.918	0.592
Iadl	-3.289 ***	0.558	-3.282 ***	0.558
Severe illness	-1.795 ***	0.262	-1.766 ***	0.262
Non-severe illness	-0.963 **	0.410	-0.973 **	0.410
GALI	0.073	0.623	0.032	0.623
Former smoker	1.509 ***	0.569	1.516 ***	0.569
Current smoker	-2.789 ***	0.563	-2.739 ***	0.563
Daily drinking	2.001 ***	0.753	1.985 ***	0.752
Weekly drinking	2.073 ***	0.662	1.999 ***	0.661
Monthly drinking	1.035	0.676	1.009	0.675
Sports	-0.663 ***	0.184	-0.668 ***	0.184
Self-perceived health	-4.490 ***	0.292	-4.424 ***	0.291
Parental mortality	-3.121 ***	0.754		
Male x father dead			-2.877 ***	1.006

Male x mother dead			-2.878 ***	0.747
Female x mother dead			-1.700 **	0.664
Female x father dead			-2.875 ***	0.871
Constant	51.707	31.476	54.862 *	31.765
Sigma	27.999	0.183	27.967	0.183
Log-likelihood			-65,450.2	
N	16,485			

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

CHAPTER 2.
SUBJECTIVE SURVIVAL PROBABILITIES. HOW DO THEY EVOLVE AND DO THEY PREDICT MORTALITY?

1. INTRODUCTION

Usually economic models are based on forward-looking agents. These models assume that individuals take present decisions that will affect them now, but also in the future. These decisions are assumed to be based on their beliefs about the future. For example, decisions like consumption and saving are supposed to be dependent on expectations about nominal interest rates, inflation rates, medical expenses, requests, and ultimately on the likelihood of dying. However, due to the lack of information, it has been necessary to make assumptions about agents' expectations. On the one hand, macroeconomic models often assume rational expectations achieved by a trial-and-error mechanism where agents adjust their initial beliefs if they are incorrect. On the other hand, microeconomic estimations often use observable population probabilities. In both cases, these assumptions are rather strong and are likely not to hold in reality.

The LCM developed by Modigliani and Brumberg (1954) is a good example of the above mentioned. According to the LCM there is a well-defined link between the consumption plans of an individual and her income and income expectations as she passes from childhood, through the work participating years, into retirement and eventual decease. The LCM implies that economic agents maximize their lifetime utility. Because there is uncertainty how long "lifetime" means, one important ingredient of the model is the expectations of agent about how long they are likely to live.

However, due to the lack of date in the past, economists usually used the assumption that individuals determine their economic behaviour based the SSPs found in life tables. Hurd and McGarry (1997) mentioned that there are at least to objections to this assumption: (i) life tables may not be corrected even on average, because mortality risk is likely not to be stationary; and (ii) individuals in a population undoubtedly have differing subjective probability distributions and we can understand better individual behaviour if we have information about the individual-level distribution.

As in the previous Chapter, the main subject of interest in Chapter 2 is the individuals' expected longevity, also known as SSPs. Expected longevity is a fundamental factor that influences individual decisions like consumption, saving, purchase of life insurance and annuities, and labour supply among others. Despite their importance, only recently they have become of interest to social researchers, especially because of data availability and past reluctance. The latter was been one of the main blockers of the use of SSPs in particular, but the same applies to any expectations drawn from individuals. Nevertheless, since mid-1990s this view stated to change and economists increasingly started to use data on probabilistic expectations of significant personal events.

The analysis performed in Chapter 1 leads us to increase our confidence when using SSPs drawn from SHARE to estimate models of decision-making under uncertainty as the LCM and to test whether this model holds in real life. If this is the case we would expected that individuals do consider expected longevity when taking decisions like consumption, saving and wealth accumulation, and late life decisions like retirement as rational economic agents suggested by this hypothesis.

However, results obtained in Chapter 1 are not enough. Hurd and McGarry (1997) specified that before we can confidently use individual-level probability distributions, we need to understand and test their properties. The authors remark three properties that should be satisfied:

1. Individuals understand questions about probabilities and accurately report their beliefs about the likelihood of future events;
2. Individuals adjust their reported probabilities in response to new information; and
3. The reported probabilities predict outcome.

The first condition was extensively explored in Chapter 1 and it was concluded that SSPs in the SHARE database were coherent and behave reasonably well, therefore the first condition was fulfilled. In particular, we concentrated on a cross-sectional econometric analysis in order to examine the determinants of the SSPs. Our main results showed that cross-sectionally, expected longevity is coherent with mortality facts related with sex, income, education and country, but also with epidemiological data. We also found that both, physical and mental health matter and that subjective health status and healthy life-style have an effect on our perception of survival chances.

Therefore, we should go one step forward and study how SSPs evolve over time as new information comes to the respondent (second condition) and whether do they predict mortality (third condition). To our knowledge this exercise has never been done before for the SHARE dataset.

The overall goal of this Chapter is to give results that will increase our confidence that SSPs can be used in models of decision-making under uncertainty. This will be done testing the fulfilment of Hurd and McGarry (1997)'s second and third condition. In particular, this Chapter seeks three concrete objectives:

- Firstly, we want to examine if SSPs drawn from the SHARE database update in a systematic and reasonable manner with the arrival of new information. One would expect that as time evolves individuals update their beliefs about longevity. This could be because individuals perceive new medical, socio-economical, or personal information that shape initial beliefs.

  For example we would expect that, other things being equal, an individual reports lower expected longevity in the second wave if that individual has suffered a stroke between waves. But not only that, we would also expect to observe that the decrease in expected longevity is higher for those individuals that have suffered from an illness with lower survival rate, i.e. liver cancer, as oppose to an illness with higher survival rate, i.e. heart attack.

  For this matter we will study the determinants of the changes in SSPs from the first to the second wave of SHARE at an individual level. We will also determine if the observed update of SSP is coherent with epidemiological information and empirical evidence.

- Secondly, we want to examine whether people's beliefs are indeed correct. Concretely, we want to evaluate the predictive validity of SSP in the SHARE. In order to do so we test whether SSP adds any predicting power when forecasting decease. By controlling for observable risk factors we can find if the reported SSP incorporates reliable information on survival chances that would otherwise not be observed.

It should be mentioned that given that we do not observe many deceases in our sample as only two to three years have passed between the first and the second wave because our restricted sample (individuals aged between 45 and 65 years old) is rather young, the results may not be stable enough and they should be taken with careful. Alternatively we have analysed different characteristics in order to create a profile of people that did not survive to wave 2 and compare their expected longevity with that of those who survived.

- Finally, it would be interesting to compare our results with similar studies that use data from other surveys. Concretely, we will analyse and compare the results of six comparable studies:

  - Hurd and McGarry (1997) who utilized of the first two waves of the HRS to study how SSPs evolve over time and if they predict mortality;
  - Hurd *et al.* (2001) who used the first two waves of the AHEAD database to study how SSPs evolve over time and if they predict mortality;
  - Smith *et al.* (2001) who employed the first four waves of the AHEAD database to study the predictive validity of SSPs;
  - Benitez-Silva and Ni (2008) who used the first six waves of the HRS to study how SSPs update with new information;
  - Liu *et al.* (2007) who analysed data from the National Health Insurance program in Taiwan to study how SSPs evolve over time; and
  - Steffen (2009) who made use of data from the German database SAVE to study how SSPs update with new information.

The rest of the Chapter is structured as follows: Section 2 presents a literature review on how SSPs evolve over time as new information is released and the extent to how much SSPs predict mortality outcome. Section 3 describes data used in the analysis. Section 4 presents an overview of what has happened between the first wave of SHARE and the second wave. Specifically, we present the characteristics of those individuals who survived to the second wave and of those who died before it. Section 5 presents the econometric modelling details. Section 6 includes the results obtained from our analysis and in Section

7 we present the conclusions obtained. Finally, the Annex presents further analysis on the predictive validity of SSPs.

2. LITERATURE REVIEW

In this section we present a literature review on the area of SSP. The literature review presented in Chapter 1 provides a detailed overview of the history on how economists have treated data on individuals' expectations. In particular, in the case of SSPs, we mentioned that it was not until mid 1990's when they became popular when a significant amount of data from surveys became available in the US. This fact has two consequences: given that the interest of economists in expectations is recent, the literature in this field is not very extensive. Additionally, most of the literature has focused in the US, especially because the HRS was the first well-established national survey with several waves that included expectations among the questions that respondents are asked. In other regions like in Europe, Australia and in the UK, similar surveys were released later.

In this Chapter we concentrate on how SSP varies along the different waves of a survey and whether they are a good predictor of future mortality. Below we present the most prominent economic literature that deals with these questions. Some of the papers presented in this section were also presented in Chapter 1. However, in this Chapter our revision focuses on how individuals update their longevity expectations and their role predicting death, as oppose to the determinants of SSPs.

As in the case of the literature review in the first Chapter we initially present a review of the contribution of Hurd (Hurd and McGarry, 1997; and Hurd *et al.*, 2001) as he is one of the most prominent economists who analysed in detail the reported survival chances in the HRS. Further contributions of this author are also detailed in the literature review of Chapter 3. We then review the contribution of other authors.

2.1. HURD'S CONTRIBUTION

One of the first papers that tried to answer to the questions presented above was Hurd and McGarry (1997). This paper follows the paper published by these authors in 1995. The authors argued that because the Hurd and McGarry (1995)'s results were based on cross-section analysis, it was not possible to determine how SSPs evolved over time. For that reason they used data from the first two waves of the HRS to establish whether expected longevity has a significant power to predict mortality and to study how it evolves as new information comes to the respondent.

The first round of interviews was conducted in 1992 and the base line sample consisted of 12,652 individuals. The second wave was conducted in 1994 and included 11,492 of the original 12,652 individuals. In particular, Hurd and McGarry (1997) made use of the subsample of individuals aged 46-65 at the first interview which consisted of 11,090 respondents in the first wave.

The initial analysis performed involved reviewing the personal characteristics of the 11,090 respondents in the first wave by survivorship status in wave 2. Among the sample used 183 respondents died, 10,642 survived, and the status of 265 was unknown. The most important conclusion that the authors reached was that, on average, survivors reported much higher SSPs than those who died. Therefore, at least in a gross way, SSPs predict mortality. Additional findings when comparing the characteristics of survivors and non-survivors include that the latter were, on average, older and poorer. They also found that more men died. This is explained not only by the fact that mortality data suggest that females live longer, but also because of differences by sex in the age distribution of the sample. Moreover, smoking and disease prevalence was more common among non-survivors. Hurd and McGarry (1997) also compared SSPs across waves and concluded that wave 1 has a strong effect on wave 2.

However, the two main building blocks of the study are the analysis of changes in Survival Probabilities and mortality outcomes. For the former, the authors highlight the fact that there is considerable movement in SSPs at individual level and suggest that this may be due to unanticipated health shocks. This is because they assume that when individuals report their expected longevity they take into account their expectations about future events that would affect their survival chances, anticipated health shocks, for example. Therefore, unanticipated changes or new information is the primary source of variation of SSPs. Hurd and McGarry (1997) modelled changes in expected longevity as a function of changes in the survivorship of the respondent´s parents, spouses and siblings, and the onset of diseases. They concluded that changes in mortality status of parents do have a large effect on changes in survival probabilities, same as in the case of the dead of a spouse. In the case of siblings, no significant effects were found. They also concluded that respondents appropriately reduced their survival probabilities at a new diagnosis, particularly for conditions that do decrease life expectancy such as cancer, heart conditions and lung disease.

In the case of mortality outcome, Hurd and McGarry (1997) examined how well SSPs at wave 1 predict actual mortality after controlling for other observable variables. They used a probit model where the dependent variable is observed mortality as a function of expected longevity in the first wave and socio-economic status variables, health behaviours and disease conditions. The results shown suggest that SSP in the first wave has a significant and large effect on mortality. Additionally, mortality falls with income, wealth and physical activity. It was also found that cancer is the strongest predictor of mortality, followed by heart attack, and stroke.

Overall, Hurd and McGarry (1997) concluded that respondents modify appropriately their survival probabilities based on new information and health shocks. Moreover, SSPs predict actual mortality.

A similar approach was then followed by Hurd *et al.* (2001). They examined three questions: (i) whether SSPs vary in cross-section in a way that is appropriate with mortality data; (ii) how SSPs change in panel in response to new information such as the onset of an illness; and (iii) whether they predict actual mortality. Contrary to Hurd and McGarry (1997), they used data on the AHEAD survey.

The AHEAD survey is a biennial panel survey of individuals born in 1923 or earlier and their spouses. At the time of the study AHEAD contained two waves. The first wave was collected in 1993 and included 8,222 individuals. Wave 2 was fielded in 1995. Hurd *et al.* (2001) found 813 diseases between wave 1 and wave 2 and 45 individuals whose mortality status was unknown. From a visual inspection it was found that average wealth was much lower among those who died between waves. Furthermore, on average, survivors were more educated and had more income than non-survivors.

Regarding the first question examined, the authors presented a regression of SSPs as a function of wealth and income quartiles, education bands, and health condition variables. Eight of the 13 health variables are significant at the 0.05 level, and they are associated with a reduction in the SSPs of nine to 25% of the average probability. More details on the results obtained by Hurd *et al.* (2001) can be found in the literature review of the first Chapter.

As mentioned above, the second question they try to answer is how SSPs change in panel in response to new information such as the onset of an illness. Hurd *et al.* (2001) argued that the probabilities should change in response to new information that alters survival chances. Such information would be onset of a health condition that is associated with an increased risk of death. In the AHEAD data they found that among singles who had not had cancer prior to the baseline interview, 5.1% had a cancer between the waves. Among all singles, including those with a history of cancer prior to baseline, 5.5% had a new or initial cancer between the waves. Similar results were obtained for married individuals.

The authors used a regression of the change (wave 1 – wave 2) in the SSPs based on the onset of health conditions and other variables such as gender, marital status, spouse mortality and entering a nursing home. Hurd *et al.* (2001) found that, for the AHEAD respondents, a number of the conditions have negative coefficients indicating that onset reduces the SSPs, and cancer, high blood pressure, diabetes and depression have negative effects that are significant at the 5% level. Contrary to the finding of Hurd and McGarry (1997), they found that the death of a spouse increased SSPs. They argued that an explanation for the difference may be that at the ages of the AHEAD respondents the death of a spouse is preceded by a period of care that reduces the optimism of the caregiver.

Finally, they examined whether SSPs predict actual mortality. They presented cross-tabulated evidence that, for AHEAD respondents, SSPs have considerable explanatory power for mortality particularly in the low range. For example, the mortality rate among those who gave a zero probability of survival was about 0.13 compared with about 0.05 among those that gave a 0.50 probability of survival. Moreover, a probit estimation of mortality between waves was performed. It was concluded that SSP predicts actual mortality even after controlling for socio-economic indicators and the health conditions and self-perceived health status.

2.2. FURTHER CONTRIBUTIONS

In this subsection we present a review of further contributions. For the sake of clarity we firstly present a review of the literature that is based on US surveys, and then the literature based on surveys of other regions.

2.2.1. Literature based on US surveys

Smith *et al.* (2001) argued that many economists are convinced that when individuals are asked to formulate probabilities they are unlikely to do it correctly, but at the same time it is argued that people do have a reasonably articulated, but biased, internal scale for risk perceptions. As a result, the extent of bias in these subjective perceptions has been an important unresolved issue. The authors investigated the reliability of SSPs, testing whether they have any predictive power over individual mortality.

In order to perform the investigation, they used the first four waves of the HRS. What differentiates this study from earlier analyses (Hurd and McGarry, 1997; and Hurd *et al.*, 2001) is that it includes two additional waves of interviews that offer a more complete record to evaluate subjective beliefs. But more importantly, the main feature of the study is that it incorporated selection effects because the sample of individuals who are available for testing the predictive value of longevity varies.

The authors present a probit estimation on the observed mortality between wave 3 and wave 4 and between wave 3 and wave 2 (mortality model). Two different models are specified. The first model considered the role of longevity expectations, new health conditions between waves 2 and 3, changes in the number of reported limitations in activities of daily living between waves 2 and 3, smoking status in the preceding wave, and family income in the preceding wave. The second model does not consider smoking and income variables. As mentioned above, selection effects are incorporated. It takes into account for the fact that respondents must be alive in wave 3 in order to be eligible to be counted as a death between waves 3 and 4. This selection model includes the following as determinants of being alive for wave 3: the longevity expectation in wave 2, smoking status in wave 2, new health shocks between waves 1 and 2, earlier changes in activity limitations (between waves 1 and 2), family income in wave 2, and some demographic characteristics.

They found that all the hypothesized determinants of mortality are statistically significant and consistent with prior expectations. In particular, they obtained that in the mortality model higher SSPs in the previous wave decreases the probability of dying in the current wave. Moreover, in the selection model, they found that higher SSP in the previous wave is associated with higher survival rate in the current wave.

An interesting feature of the study presented by Smith *et al.* (2001) is that one of the purposes of the paper is to test whether the reported survival chances do reflect all the information that individuals have when eliciting them. In order to do so, they follow a two-step process: using the mortality model a predicted SSP is calculated. Then it is tested whether mortality in the next waves explains the predicted SSP. They concluded that the hypothesis that an individual's longevity expectation contains all the information that each individual knows about his/her survival chances should be rejected.

With this study Smith *et al.* (2001) reached two important conclusions. First, longevity expectations are reasonably good predictors of future mortality. They found that subjective longevity beliefs are consistent with individuals' observed survival pattern. In this respect, they found that deaths are signalled through lower SSPs in the previous waves. Second, longevity expectations are consistently updated with new health information. In this sense, they found that SSPs do respond negatively both to serious, new health shocks and to increases in individuals' functional limitations.

Using data from the first wave of the HRS and the first wave of the AHEAD survey, Siegel *et al.* (2003) investigated whether self-rated life expectancy predicted actual mortality after controlling for measures of health, self-rated health, and socio-demographic characteristics. In particular, they had two objectives: (i) replicate analyses from other research on the effect of self-rated health on mortality, and (ii) assess the effect of self-rated life expectancy on mortality, adjusting for a set of health measures and socio-demographic traits, and finally, adjusted for self-rated health status.

The incidence of mortality in Siegel *et al.* (2003)'s database is high. In the case of the AHEAD dataset, they found that 9% of men and 5% of women died between the first and the second wave. For the HRS dataset, these percentages were lower (4% and 2%, respectively).

The authors use a Cox proportional hazard specification in order to model the probability of dying before 2 years (for the AHEAD respondents) and 3 years (for the HRS respondents), conditional on being alive. They propose 3 hazard model estimated as a function of : (i) self-rated life expectancy and a set of covariates, (ii) self-rated health and a set of covariates, and (iii) self-rated life expectancy, self-rated health and a set of covariates.

Siegel *et al.* (2003) found that for the first specification, AHEAD respondents who expected to live longer were significantly less likely to die during the studies period, both in the case of men and women. However, for the HRS respondents no significant effect was wound. When analysing the effect of self-rated health on the hazard of death, they found that both HRS and AHEAD respondents who report a better health were significantly less likely to die during the studies period. Finally, when both self-rated measures are included, they found that for the AHEAD dataset both lower self-rated life expectancy and poorer self-rated health are predictive of mortality. In contrast, for the HRS respondents self-rated life expectancy is not predictive of mortality once self-rated health is included.

They concluded that, although self-rated life expectancy and self-rated health may be conceptually related, they have independent empirical effects on mortality. Also they suggested that the difference of results between AHEAD and HRS could be driven by the fact that respondents in the HRS are younger.

Benitez-Silva and Ni (2008) explored similar issues using the first six waves of the HRS. They stated that given that expected longevity is naturally a function of health, it is key to decide how health is measured. Moreover, in order to correctly model the transition of SSPs a correct measure of health dynamics should be applied. They argued that the dynamic component of health can be seen as an investment with the following characteristics: (i) individuals invest in their health through, for example, visits to health professionals for preventive care, changes in their diets, exercise, and changes in habits such as smoking or drinking. However, the outcomes of these investments are uncertain. (ii) Health depreciates through the natural aging process and through the worsening of chronic conditions. (iii) Health can deteriorate (or improve) significantly in rather discontinuous (and non-monotonic) ways as health shocks occur.

Two separate measure of health dynamics are presented: self-rated health changes (individuals are asked about the evolution of their health status) and computed health changes (health status in the latter wave compared with health status in the former wave). The authors argued that a measure of health dynamics based on computed health changes suffers from different drawback. On the other hand, self-rated health change is preferable because it incorporates all the relevant information.

When modelling changes in expected longevity, they considered two alternative specifications: (i) SSP in a given period as a function of SSP in the previous period and other variables including self-rated health changes, and (ii) the change in SSP as a function of these other variables including self-rated health changes. They concluded that expected longevity in the previous period does play an important role on today's expectation. Additionally, they found that using subjective measures of health is a preferred option in order to capture the dynamic effect on both individuals' SSPs and how they evolve over time.

2.2.2. Literature based on non-US surveys

With the objective of analysing and determining whether global self-rated life expectancy and parental mortality predict mortality in Australia, Van Doorn and Kasl (1998) used the ALSA dataset that collects data on a population representative sample of individuals over 70 years old. One interesting feature of ALSA is that, unlike the HRS, ELSA and SHARE, respondents are asked to assess their survival chances for the next 10 years using a non-numerical scale (i.e. very likely, likely, unlikely, and very unlikely). Using a weighted multiple logistic regression the authors explore 3-year mortality and concluded that for men self-rated life expectancy does predict mortality. However for women the predicting power is null. Additionally, they found that self-rated life expectancy is a better predictor of mortality among the old than the young. The reason for this is that we would expect that thinking about death is more relevant to the lives of older people who are probably experiencing losses of friends and relatives more frequently than younger people.

Another interesting and related study is the paper by Liu *et al.* (2007). This study addresses two different issues: first, the authors used pooled cross-sectional data to investigate the determinants of SSPs. Secondly, they used panel-structured data to explore how longevity expectations respond to new health information. For this matter, Liu *et al.* (2007) made use of data from the BNHI program in Taiwan. In particular, the study is based on 2 waves of data drawn from patients between 40 and 65 years old from the Mackay Memorial Hospital in Taipei, Taiwan, in 2001. Wave 1 was administered by in-person interview between July and December 2001 and included questions on socio-demographic characteristics, subjective health status, health behaviour, and longevity expectations. Wave 2 consisted of follow-up telephone interviews between two and three months later. Overall, seven hundred patients participated on both waves.

What is interesting from this study is that between waves, the patients are subject to physical examination, as well as medical diagnoses and advice on health behaviour provided by their physicians in order to understand how their longevity expectation reacts as new information is released. Specifically, seven tests are performed on patients (urinalysis, complete blood count, blood sugar, liver function, renal function, lipids, and uric acid). Moreover, physician advice includes advice on 'quitting smoking', 'quitting drinking', 'oral hygiene', 'weight control', and 'diet and nutrition', and diagnoses of six diseases (hypertension, thyroid disease, heart disease, hepatitis, hyperlipidaemia, gout) that were not self-reported in wave 1 of the 2001 survey.

Liu *et al.* (2007) highlight three limitations of the panel data used for the study. First, females are over-represented compared to national averages. Second, there is a potential selection bias issue because patients accepting to participate in the survey may be more health-conscious then those that did not want to participate. Finally, the interval between the two waves is of only two or three months, which limits the evolution over a longer period of SSPs.

As explained in the first Chapter, then dealing with the determinants of SSPs, they found that men tend to significantly overestimate their chances of living to 75 or 85, compared with life-table rates. In contrast, women tend to underestimate the probability of living to 75 and overestimate the chance of living to 85. The authors also found that males and married persons are more optimistic about their longevity, and that income is strongly correlated with higher SSP. However, education does not seem to affect expected longevity. Furthermore, Liu *et al.* (2007) concluded that the same-sex parent hypothesis hold for the Taiwanese respondents and that, consistent with actuarial data, the subjective probability of living to age 75 or 85 increases with age.

When exploring how longevity expectations respond to new health information, a Bayesian risk updating model is used in which the individual´s reported survival chance after receiving the results of the physical examination is a function of the reported survival chance before performing the examination and any new health information provided by the examination. Using an OLS specification, Liu *et al.* (2007) found that the coefficients of the reported survival chance before performing the examination were significantly greater than zero, which implies that prior beliefs do play an important role on posterior survival expectations. However, most importantly, they concluded that new information provided by the examination is almost three times as influential as the respondents prior beliefs in the case of survival chance up to 75 years old, and approximately ten times in the case of survival chance up to 85 years old.

Results presented by the authors provide further detail. In the case of the test outcomes, they found that abnormalities of lipid and liver function have a significant negative effect on SSPs. Furthermore, a recommendation about weight control is the only type of physician advice which significantly reduced SSPs and a diagnosis of heart disease has the strongest negative influence.

As mentioned in the literature review presented in Chapter 1, a recent study by Steffen (2009) uses the German database SAVE to explore how Subjective Life Expectancy (measured in year) updates as new information arise, as oppose to SSPs (measured in percentages).

Because Steffen (2009) considers as the object of interest Subjective Life Expectancy as the age at which respondents believe they will die, the standard Bayesian updating framework applied in most of the papers referred above cannot be directly applied in this case. He argues that this framework applied in this context would assume that the new Subjective Life Expectancy is to be the weighted average of the previous Subjective Life Expectancy and some new information. Therefore, it is not apparent why individuals would apply such a complex procedure when dealing with adjustment of absolute numbers.

The author simply assumes that new Subjective Life Expectancy is equal to the old one plus and adjustment addend (positive or negative). It is argued that adjustment of Subjective Life Expectancy between waves can be done for two reasons: age and health. The former is based on the idea that as people get older they adjust their prior beliefs because of growing optimism with age. The latter focuses on negative health shocks as unanticipated diagnosis of a severe illness or the unexpected worsening of medical conditions.

In order to analyse the role of these adjustments when individuals update their Subjective Life Expectancy, the method of estimation used is the OLS. The same econometric approach is also applied when analysing how the reported average Life Expectancy (respondents are asked which age they think men and women of their age will reach on average) varies across waves. Furthermore, a probit specification is used to analyse the role of the age and health adjustment to explain why a respondent used to believe that she will live longer than the average and now she does not (respondents are also asked whether they think they will live longer, shorter or about the same as average reported).

Steffen (2009) concluded that people do adjust their Subjective Life Expectancy on a regular basis. Moreover individuals update their Subjective Life Expectancy in a rational manner: a negative health shock leads people to adjust their Life Expectancy, while the reported average Life Expectancy remains unchanged as expected. He also mentions that these results raise the hope that a more precise general knowledge of the Average Life Expectancy can lead to a higher quality of individuals' economic decisions among economic agents.

In the next section we present in detail the data that will be used on our analysis of how SSPs evolve over time as new information is revealed to the respondent and whether SSPs predict actual mortality.

3. DATA

As in the previous Chapter, our analysis will be based on data drawn from SHARE database that is harmonized with the HRS and the ELSA.

The first wave of the SHARE database comprised a total of 31,115 respondents from Austria (1,893), Belgium (3,827), Denmark (1,707), France (3,193), Germany (3,008), Greece (2,898), Israel (2,598), Italy (2,559), Netherlands (2,979), Spain (2,396), Sweden (3,053), and Switzerland (1,004).

In 2008, with the release of wave 2, SHARE went into its longitudinal dimension. This brought new possibilities for data analysis as the same respondent is observed over time. According to SHARE, the efforts were conducted to re-contacting respondents from the first wave, but also to select a "refresher" sample. For the refresher sample the same sampling methods were used as in the first wave, only cohorts born in 1955 and 1956 were oversampled to keep the sample representative of the population 50 years old and older.

Overall, the second wave of SHARE consisted of 33,281 respondents from Austria (1,341), Belgium (3,169), Czech Republic (2,830), Denmark (2,626), France (2,968), Germany (2,568), Greece (3,243), Italy (2,983), Netherlands (2,661), Poland (2,467), Spain (2,228), Sweden (2,745) and Switzerland (1,462).

It should be mentioned that Israel has been excluded from our analysis because of the lack of complete information in the first wave and because currently there is no second wave available for this country.

The whole longitudinal sample which includes individuals that participated in both waves of SHARE, consist of 18,741 individuals, 8,270 males and 10,471 females. They are geographically distributed as following: 1,238 in Austria, 2,808 in Belgium, 1,249 in Denmark, 1,998 in France, 1,544 in Germany, 2,280 in Greece, 1,766 in Italy, 1,777 in Netherlands, 1,375 in Spain, 2,010 in Sweden and 696 in Switzerland. We observe 280 individuals aged under 50, 8,527 aged between 50 and 64 years old, 5.578 aged between 65 and 74 years old and 4,356 respondents aged 75 or more.

As in the previous Chapter, our analysis is restricted to individuals that are aged under 66 years old in the second wave[29] because of two reasons. Firstly, because of the way the question on SSPs is worded, individuals in our final sample are asked about the chances of reaching a common target age, in this case 75 years old. For older respondents, the target age is set in five-year bands such that the distance from current to target age is between 10 and 15 years. For example, if an individual is aged 66, he/she will be asked about survival chances over target age of 80 years[30].

Secondly, as highlighted above, the vast majority of longitudinal respondents are aged 65 or under. At the same time, individuals from this age interval are more likely to take retirement decisions, which in turn, will be relevant for the Chapter 3. After eliminating negligible missing values, non-respondent items and inconsistencies[31], we find that our sample will consist of 8,484 individuals aged from 43 to 65 years old in the second wave that have participated in both waves of the SHARE survey.

It is important to stress the fact that this age criterion is referred to the second wave. This is a rather important point because we are interested on individuals that are asked about their expected survival chances referred to the same target age in both waves. If we want to analyse the evolution of SSPs, it would not be sensible, for comparison reasons, to include a respondent who is asked about his/her SSP referred to a target age of 75 years old in the first wave, and about his/her SSP referred to a target age of 80 years old in the second wave. We might find this situation for individuals aged 63, 64 and 65 in the first wave (in 2003-04) and aged 66, 67 and 68 in the second wave (in 2006-07). Specifically, we have found 1,236 individuals in this situation that will not be taken into account for the analysis performed below.

---

[29] In the analysis performed in the first Chapter we defined our restricted sample as individuals aged between 45 and 65 years old. In this Chapter we have also included individuals aged between 43 and 44 (36 respondents) with the objective to increase the number of observations.

[30] More details can be found in Table 1 of Chapter 1.

[31] It should be highlighted that we have found very few inconsistencies regarding the target age of the SSP question. Some individuals aged under 65 have a target age different than 75 and some individuals aged over 65 have target age equals to 75. We have eliminated these observations.

4. WHAT HAPPENED BETWEEN WAVE 1 AND WAVE 2?

One of the advantages of having a panel data is that we can observe individuals over time, and therefore we can identify what has happened between waves. Concretely, it is interesting to know if a respondent has died between waves.

As show in Table 1, overall, 533 respondents from our original sample died between wave 1 and wave 2 out of 28,517 individuals[32]; therefore, the average mortality rate (number of individuals died between waves divided by total number of individuals in wave 1) is 1.87%. The table also shows that Belgium presents the lowest non-survival rate (0.94%), while Spain exhibits the highest rate (3.63%). This is surprising statistical result because Spain is among the countries with the highest life expectancy. However, given the low number of deceases, it is unlikely that this result would be consistent over time.

If we consider age groups for the respondents, results are quite sensible. As we would expect *a priori*, mortality rate increases with age. Specifically, all of the respondents aged under 50 survived, 0.64% of respondents aged 50 – 65 died, 1.86% of those aged 66 – 75 died, and 6.6% of individuals aged over 75 years old died between waves. Results are shown in Table 2.

For our sample, which is restricted to young elders aged under 66, mortality observed between wave 1 and wave 2 is quite low, as we would expect. Hurd and McGarry (1997) calculated the mortality rate for HRS respondents between wave 1 and wave 2. Their results suggest a higher mortality rate (1.69%). The difference could be explained partly by the fact that this mortality rate was calculated between 1991 and 1994, almost 13 years before SHARE, and it is expected that mortality risk is constantly decreasing.

One interesting exercise is to analyse the characteristics of those individuals that died between 2003 and 2007 in our sample of interest, respondents under 66 years old. We would like to know, for example, if they were already sick, whether they had low expected

[32] Because SHARE does not currently have information from Israel, we report the survival status of 28,517 individuals out of a total of 31,115. The difference corresponds to Israelis.

survival chances, what their socio-economic characteristics were. Next, we analyse these questions.

Table 3 shows the mean of some interesting variables by survivorship status in wave 2. The sample is comprised of individuals aged between 43 and 65 years old in the first wave, 14,977 of them lived to wave 2 and 81 of them did not survived. As mentioned in the Introduction, it is difficult to assess whether SSPs satisfy the third property suggested by Hurd and McGarry (1997) because we do not observe many deceases in our sample. Using a probit specification we find that SSPs do predict mortality. We believe that this result obtained should be taken with careful as they might not be stable enough. However, once more waves of the SHARE database become available, further research should confirm this. Annex 1 presents further details on the estimation.

Alternatively, we can inspect the SSP for those individuals that survived and for those that died. Figure 1 shows the cumulative distribution of SSPs by survivorship status. It can be observed that average SSP for those who survived is substantially higher than for those that died. We find that average SSP for the latter group is 10.3% lower. Similar results were found by Hurd and McGarry (1997), though they found a greater difference (20%). Again, part of the explanation for this difference is the timing when the first and the second wave of HRS and SHARE were done. Van Doorn and Kasl (1998) also report similar results for Australia.

As these authors stated, this results suggest that, at least in a gross way, that the SSPs predict mortality. Similarly, Börsch-Supan *et al.* (2008) clearly mentions that the most basic test that allows one to assess the predictive power of SSPs is to compare average survival probabilities between those respondents who survived from wave 1 to wave 2 and those who died. Average survival probabilities are much smaller for those who did not survive than for those who survived.

We have also analysed different characteristics in order to create a profile of people that did not survive to wave 2. It should be mentioned that this exercise does not imply, in any case, causality between the outcome and the relevant variables.

Our results show that, in our sample, people who died tended to be older. The average age of non-survivors exceeds that of survivors by almost 2.6 years. This result is coherent with information drawn from life tables and with epidemiological data. It is also coherent with other studies. Hurd and McGarry (1997) found an average age difference of 1.2 years between those who died between waves and those who survived. Hurd *et al.* (2001) and Van Doorn and Kasl (1998) also found, for the AHEAD database and for Australia, that mortality rate increases with age.

Hurd and McGarry (1997) and Hurd *et al.* (2001) found, for the US, that those who survived had higher income and are more educated. Our results, for Europe, are consistent with these findings. We found that average income for survivors was 13,980 Euros[33] higher than for non-survivors. We found that average ISCED international code for those who lived to wave 2 is 2.85[34] and 2.64 for those who did not survive.

Our analysis also points out that among those who died, the majority was males. Mortality data suggests that the mortality rate faced by females is lower, so this result is not surprising. Men comprise about 70% of the deceased and only 44% of the survivors. In the case of Hurd and McGarry (1997), these figures were 60% and 45% respectively. As we would expect, those who died between 2004 and 2007 reported worse health at baseline than the survivors. Average self-reported health for those who died was 3.53[35] while those who lived to wave two reported an average of 2.71. Hurd and McGarry (1997) found that the fraction in poor health in each of the two groups was 0.06 and 0.36 respectively. Van Doorn and Kasl (1998) obtained comparable results. Our results obtained also suggest that, on average, mental illness (EURO-D), smoking and lack of practicing sports is more common among fatalities.

[33] Income is measured in Euros corrected for Purchasing Power Parity 2005.

[34] ISCED CODE: 0 (no education), 1 (primary education or first stage of basic education), 2 (lower secondary or second stage of basic education), 3 (upper secondary education), 4 (post-secondary non-tertiary education), 5 (first stage of tertiary education), and 6 (second stage of tertiary education).

[35] Self-reported health is defined as follows: 1 (excellent), 2 (very good), 3 (good), 4 (fair) and 5 (bad). Therefore, a lower average self-reported health implies better subjective health.

Hurd and McGarry (1997) concluded that disease conditions all have the expected relationship: among those who died greater fractions had suffered different diseases[36]. Similarly, we find that the number of severe and non-severe illnesses suffered is lower for those who lived to wave two. Also, on average, the presence of a long-term illness and physical limitation (ADL and IADL) was more common among those who died.

Now that we have drawn a profile of those individuals that did not survive to the second wave of SHARE, it is interesting to analyse what has happened to those individuals that did survive and we observe them twice. As mentioned above 8,484 individuals aged from 43 to 65 years old in the second wave that participated in both waves of the SHARE survey.

The first object of interest is the evolution of their SSP. It is not straight forward to determine *a prior* if we should expect subjective expected longevity to increase or to decrease in the second wave with respect to the first wave. On the one hand, we could believe that the fact that these individuals have already survived one wave would reassess their odds of living to age 75 as being higher. Moreover, if survival expectations follow a life table, they must, by construction, increase with age. But, on the other hand, one could expect the reassessment to be downwards, as a consequence of a change in awareness of risk factor as new health problems and illnesses that are more likely to appear between waves.

Academic evidence in this matter is mixed. Hurd and McGarry (1997) found that more respondents gave a decline in the probability than gave an increase, and that average expected longevity decreased in the second wave of the HRS. Contrary, Smith *et al.* (2001) using four waves of the HRS found that longevity expectations are increasing and do not decline over the full period. Also Hurd *et al.* (2001) using AHEAD database found that SSPs increased between waves in all age bands.

Our analysis coincides with the evidence presented by Smith *et al.* (2001). Our results suggest that average SSPs in the second wave is slightly higher than for the first wave of

[36] Hurd and McGarry (1997) considered the following diseases: high blood pressure, diabetes, cancer, lung disease, heart condition, angina, congestive heart failure, stroke and arthritis. The prevalence of each of these diseases is higher for fatalities.

SHARE. This result is coherent with life tables. Table 4 present the results. It can be observed that the average SSPs in the second wave is 1.11% higher than in the first wave. Moreover, the standard deviation in the second wave is lower. At the same time, Figure 2 shows the distribution of the change of expected longevity (wave 2 – wave1). It can be observed that positive changes are slightly more common than negative changes. In particular, around 13% of the respondents reported a SSP 10% higher in the second wave, while approximately only 10% of the respondents reported a SSP 10% lower in the second wave. It should also be mentioned that almost 35% of respondents report a similar (in the interval ±5%) SSP in the second wave compared to their answers in the first wave.

It is also interesting to observe the stability of the SSP responses across waves at individual level. Table 5 shows that around 37% of respondents reported a higher expected longevity in the second wave, while around 32.5% reported lower SSP compared with the first wave. As mentioned above, this result is opposed to that of Hurd and McGarry (1997). These authors also found that 26.9% of respondents gave the same response in both waves, mostly at 50% and 100%. Our analysis concluded that a higher percentage of respondents reported the same SSP (30.2%), mostly at other values different than 0%, 50% or 100%.

Table 6 shows the average SSP response in the second wave conditional on reported SSP in wave 1. The overall impression is that expected longevity in the second wave is, in some way, conditioned to expect longevity reported 3-4 years ago, in the first wave. Hurd and McGarry (1997) found similar outcomes. This result will be very helpful when modelling SSP's transition in the next section.

Apart from previous expectations, subjective longevity is expected to change in response to new information that alters survival chances. Such information would include the onset of a health condition associated with an increased risk of death.

Table 7 presents the incidence of new health conditions between wave 1 and 2 for all respondents. Thus, for example, among those who not had a heart attack prior to the baseline interview (7,944 individuals), 211 suffered one in between the first and the second wave (2.66%).

It can be observed that high blood pressure is the most common new health condition associated with an increased risk of death. 11% of those who did not suffer from this illness in wave 1 contracted it between both waves. Also significant is the onset of high cholesterol and arthritis; 9.6% and 8.5% of respondents developed the respective illnesses. Heart attack, diabetes, osteoporosis and lung disease are also common. Among contracted cancer, breast for females, and prostate for males are relatively common (0.5% and 0.43% correspondingly). The prevalence of colon or rectum cancer is also noteworthy (0.25%).

The onset of new health conditions that we find is much lower than the results obtained by Hurd *et al.* (2001). Both results are not directly comparable because these authors use the AHEAD database which includes individuals aged 70 or over. A higher prevalence of illnesses among older population is to be expected.

Finally, we also find evidence that, on average, survivors present higher ADL and IADL limitations and improvement of mental health[37]. We also report evidence that average self-report health has decreased among survivors[38].

Our inspection of what has happened between wave 1 and wave 2 of SHARE reveals several aspects that are worth summarizing. We find out that those who did not survive to the second wave, on average, reported lower expected longevity, were older, less educated and poorer, reported more severe and non-severe illnesses were physically limited, and suffered from a long-term illness.

On the other hand, we conclude that those who did survive to wave 2, on average, increased their SSP; though, expected longevity in the second wave seems to be conditioned by previous expectations. We also find evidence pointing towards the onset of new health conditions associated with an increased risk of death, particularly related to high blood pressure, cholesterol and arthritis. Finally, survivors tend to undercut their subjective health status.

[37] Data inspection shows that the average number of ADL limitations has increased 0.003 and the average number of IADL limitations has increased 0.015. Also shows that the average EURO-D factor has decreased by 0.13. This result implies that, on average, mental health has improved.

[38] Data inspection shows that average change of self-reported health status is 0.144. This result implies an average decrease in health status because of the way variable is defined.

We believe that these results seem sensible from the mortality and epidemiological point of view and are coherent with academic literature. They are a good starting point to consider when modelling SSP's transition in the next section.

5. ECONOMETRIC MODELLING

In the first Chapter we examined the determinants of SSPs and we concluded that they vary in a reasonable manner with a number of observable variables like socio-economic characteristics, health conditions, health behaviour, self-reported health and parental mortality. However, this type of variation does not imply causality. We therefore, focus on the determinants of the evolution of SSPs, and we model the evolution of expected longevity.

5.1. METHODOLOGY

As mentioned in the Introduction, the main objective of this Chapter is to understand how the reported expected longevity changes between the first and the second wave in the SHARE database. Expected longevity can change for several reasons. We assume that when individuals report their expected longevity they take into account their current set of information which includes their expectations about future events, like changes in health status. In that case, we will observe changes in expected longevity when unexpected events or new information affects the individual's initial assessment. Those unexpected events or the new information may be related to health shocks, like the onset of new illnesses, or the death of a parent, for example.

In order to model how SSPs change over time we follow the methodology proposed by Benitez-Silva and Ni (2008) in order to capture the role that new information plays. The authors state that individuals integrate new and old information with weights attached to each of these sources when updating their expectations in the following manner:

$$SSP_{t+1,i} = SSP_{t,i} + \gamma(\omega_{t+1} - E[\omega_{t+1}|\Omega_t]) \qquad (1)$$

where respondent are indexed i and time as t, SSP represents the individuals' reported SSP; ω symbolizes exogenous changes that could affect expected longevity. These changes are specified in section 5.2. **Ω** characterized the set of information. Therefore, the term in parenthesis is the difference between the observed information and the expectation of this information individuals had in the previous period. The parameter γ is interpreted as the weight of the unanticipated component of new information on changes in expectations. Implicitly, we are assuming that new information is not anticipated, and therefore not embedded in the initial individuals' longevity expectations. But at the same time, as can be observed in Table 6, there is persistency based on previous expectations.

As suggested by Benitez-Silva and Ni (2008), if we were able to observe all the elements in Equation 1, one could estimate the equation simply by an OLS procedure. But it is clear that we do not observe the expectation of future exogenous characteristics given the current set of information. We therefore propose the following specification:

$$SSP_{t+1,i} = \alpha + \beta SSP_{t,i} + \gamma\omega_{t+1,i} + \epsilon_{t+1,i} \qquad (2)$$

where $\epsilon_{t+1,i}$ is error term.

Alternatively, we also consider the approach followed by Hurd and McGarry (1997) for the HRS and Hurd *et al.* (2001) for the AHEAD dataset. These authors consider a specific case of the Benitez-Silva and Ni (2008) approach where β=1. Therefore they focus on modelling the changes in SSPs as a function of changes in observable characteristics, as follows:

$$\Delta SSP_{t+1,i} = \alpha + \gamma\omega_{t+1,i} + \epsilon_{t+1,i} \qquad (3)$$

Additionally, we have also found that Smith *et al.* (2001), using data from the HRS, test both specifications (Equation 2 and Equation 3).

We have also reviewed a somewhat similar approach followed by Liu *et al.* (2007) for the National Health Insurance program in Taiwan. They consider a risk-updating model (Viscusi, 1985; Sloan *et al.*, 1999; Smith *et al.*, 2001; and Viscusi and Hakes, 2003). In particular, the individual's expected longevity at a certain moment, SSP_t is thought to be the result of a weighted average of her previous expected longevity, SSP_{t-1} and the unobserved risk equivalent, S_t , of new health information as shown below:

$$SSP_t = \frac{\theta SSP_{t-1} + \gamma S_t}{\theta + \gamma} \qquad (4)$$

The unobserved risk equivalent (S_t) of new health information is expected to be a function of observable characteristics (Z_t)

$$S_t = f(Z_t) \qquad (5)$$

Therefore, using a linear approximation, Liu *et al.* (2007) propose a model of the form:

$$SSP_{t+1,i} = \alpha + \beta SSP_{t,i} + \sum_{j=1}^{k} \delta^j Z_{t+1,i}^j + \varepsilon_{t+1,i} \qquad (6)$$

where respondent are indexed i and time as t, and the vector Z represents k exogenous changes that could affect expected longevity.

Clearly, this approach is equivalent to the one we are following in this Chapter.

5.2. VARIABLES USED IN THE ANALYSIS

As shown above, the whole modellization of the transition of SSPs across waves is based on past expectations and exogenous changes related to new information. We, therefore, present the variables to be used with the aim of capturing those exogenous changes.

In order to perform the econometric analysis pointed out above, we have taken advantage of the significant amount of information that SHARE database provides. Specifically, we have made use of a set of variables that will be helpful when modelling the transition of expected longevity between wave 1 and wave 2. It should be mentioned that the variables selected are based on those used in Chapter 1. Therefore, further details on the specification of these variables can be found there[39]. Also, it is important to specify that we use the evolution of relevant variables for our analysis. We consider this evolution denoted as the variable in the second wave minus the variable in the first wave.

The first variable of interest is SSPs in wave 1 and wave 2. A detailed statistical analysis was performed in Section 4. Apart from this variable, we have also included some measures of changes in socio-demographic characteristics. Most of these characteristics are expected to remain constant or we do not expect to observe significant variations between waves, for example, country of birth, gender, marital status, educational level and number of children. We therefore, only include the following socio-demographic characteristics: change in respondent's age and change in the respondent's income[40].

While Hurd and McGarry (1997) and Liu *et al.* (2007) didn't include any socio-demographic variable, Hurd *et al.* (2001), Smith *et al.* (2001) and Steffen (2009) include change in age and Benitez-Silva and Ni (2008) include changes in marital status and changes in the number of grandchildren.

[39] Table 5 in Chapter 1 presents a brief description of the variables used in the econometric modelling. In this Chapter we use the change in the variable.

[40] We have finally used a measure of the respondent's relative change in income using change in quartile distribution between wave 1 and wave 2. Results do not vary significantly if absolute income change is used instead.

The average age of individuals that form our restricted sample has increased from 55.6 years old in the first wave, to 57.9 years old in wave 2, as expected. On the other hand, average income has decreased within our sample from 51,597 to 44,239 Euros[41] in the second wave. The reason for this could be related to the fact that some individuals in our sample may have retired or decreased the number of hours worked between waves, so that their overall income decreases. SHARE's full documentation of wave 2 provides further evidence. Grouping countries in Nordic (Sweden and Denmark), Central European (France, Belgium, the Netherlands, Germany, Switzerland and Austria) and Southern (Italy, Spain and Greece), the general finding is that, except for Southern countries, individuals who retire from work have a sizeable reduction of their incomes – accounting for PPP's – greater than individuals who work in both waves. The largest drop is found for Central European countries: large and significant drops of 20% or more are estimated for the Netherlands and Belgium. The second largest, and significant, drop is estimated for Nordic countries, while a small and insignificant drop is estimated for Southern European countries. Chapter 3 will analyse in detail occupational status' transitions.

A priori we believe that, based on the evidence presented in Section 3, and in line with results obtained by Smith *et al.* (2001) and Hurd *et al.* (2001), age would have a positive effect on the expected longevity in wave 2. We also believe that, based on the cross-sectional results obtained in Chapter 1, increasing income would also have a positive effect on SSP in 2006/07.

As in the previous Chapter, we have incorporated the evolution of a set of objective health measures from the first wave to the second wave. Specifically, we included how EURO-D has changed. EURO-D is a global measure of mental health. It is constructed as the number of the following twelve symptoms that a respondent suffers: depression, pessimism, suicidality, guilt, sleep, interest, irritability, appetite, fatigue, concentration, enjoyment, and tearfulness. We would expect that if mental health deteriorates (the EURO-D scale in wave 2 is higher than in the first wave), expected longevity would also deteriorate.

[41] Euros in 2005, adjusted by Purchasing Power Parity (PPP).

The next variable of interest is the evolution of the presence of a long-term illness. Our restricted sample shows that around 11% of individuals contracted a long-term illness between wave 1 and wave 2, and around 13% of the respondents recovered from it between waves. The rest of the sample did not modify their status.

It is not straightforward to determine *ex ante* the expected effect of the evolution of long-term illness will have on the evolution of SSPs. We should recall the somewhat surprising result obtained in Chapter 1, when we found that, cross-sectionally, having a continuing illness was related with higher expected longevity. One could believe that if someone healed of it, that individual would increase their survival chances because that person has survived a long-standing illness. But, on the other hand, the health of that individual could have deteriorated as a consequence of the illness so that expected longevity decreased.

We have also included the evolution of a set of variables related with physical limitation: the GALI indicator[42] and the number of limitations related with ADL and IADL as suggested by Hurd and McGarry (1997), Hurd *et al.* (2001) and Benitez-Silva and Ni (2008). The evolution of the GALI indicator between wave 1 and wave 2 takes values -1, 0 or 1 depending on whether an individual has become physically limited, has not changed her status, or has overcome previous physical limitations, respectively. Positive values of changes in ADL and IADL are related to an increase in the number of activities that a respondent cannot perform.

Other objective health measures were incorporated. Specifically, the evolution of the number of severe illnesses suffered by an individual between wave 1 and wave 2 includes information related to the onset of health problems like a heart attack or any other heart problem, high blood pressure, high cholesterol, a stroke or cerebral vascular disease, diabetes, bronchitis or emphysema, malignant tumour, peptic ulcer, Parkinson disease, hip fracture or femoral fracture, and brain, oral cavity, thyroid, lung, breast, esophagus, stomach, liver, pancreas, kidney, cervix, endometrium, non-Hodgkin lymphoma, leukemia cancer. Also, the evolution of the number of non-severe illnesses suffered by an individual includes the onset of health problems like asthma, arthritis, osteoporosis, stomach ulcer,

[42] The GALI indicator equals zero if the respondent's answer to the question "for the past six months at least, to what extent have you been limited because of a health problem in activities people usually do?" was not limited and equals one if the answer was severely limited or limited.

cataracts, and, larynx, pharynx, prostate, testicle, ovary, colon or rectum, bladder, skin cancer. It is important to mention that because of the way the SHARE questionnaire is done, individuals answer whether they are currently suffering, or *have suffered* these illnesses. Therefore, we can only observe the onset of new health problems, but not if someone has overcome them. Still, we believe that this information is very useful for the analysis. We would expect that the onset of new health conditions will be related with decreasing subjective survival chances.

Similarly to the first Chapter, we have also made use of health behaviour variables. In particularly, we have included changes in the consumption of alcohol and sporting activity behaviour. As in the previous Chapter, variables are defined according to the frequency of the habit: (1) daily, (2) weekly, (3) monthly or (4) never in the case of consumption of alcohol behaviour and (1) more than once a week, (2) once a week, (3) one to three times a month or (4) hardly ever or never for sporting activity. Therefore, a positive value for the change in the consumption of alcohol and sporting activity implies a decrease in frequency between waves. Given the evidence that we found in the Chapter 1 and medical evidence, we would expect to find that increases in the frequency for both variables will be related with increasing SSPs between waves.

Smoking behaviour was also considered in the first Chapter. Balia (2007) stated that two alternative theories explain why individuals smoke. One theory defines smokers as irrational or myopic (Thaler and Sheffrin, 1981) so that they care more about present satisfaction than the future. The second theory (Becker and Murphy, 1988) states that smokers are instead, rational and forward looking, so that they internalize the effects of smoking. Balia (2007) used Italian data from the first wave of SHARE and found that current and former smokers differ in the way they discount future consequence of tobacco consumption on health and mortality risk. Based on this finding and the results obtained in the first Chapter we believe that changes in smoking behaviour might play a role in the transition of expected longevity. Unfortunately, in our restricted sample, we only observe 10 individuals that quit smoking between waves and 3 individuals that started smoking in the second wave[43]. We, therefore, do not consider smoking transition.

Following the approach followed in Chapter 1, we have also included information on parental mortality events between waves. As suggested by Hurd and McGarry (1997), the effect of a parent's death may operate through both biological and psychological mechanisms. If the parent died of a cancer that is known to have a genetic link, the child might correctly reassess his/her own life expectancy. However, a parent's death may also affect the respondent's expected longevity because it reminds him/her of his/her own mortality.

We incorporate parental mortality experience taking into account the sex of the respondent and the sex of the dead parent. This approach was called "same-sex parent" in the previous Chapter. Given the results obtained there, it would be sensible to think that if someone loses a parent between waves, it will have a negative effect on the evolution of expected longevity, and this effect could be bigger for males with their father's death between waves and for females with their mother's death between waves.

Our comparable studies also incorporate changes in health variables. Hurd and McGarry (1997) use a set of 9 concrete diseases. Hurd *et al.* (2001) also incorporated mental health variables. Furthermore, Benitez-Silva and Ni (2008) included changes in health behaviour such as smoking, practicing sports, and drinking. Smith *et al.* (2001) and Steffen (2009) included a single measure of health shock between waves.

Additionally, most of the comparable studies include changes in parental mortality when modelling how SSP updates over time. However, only Hurd and McGarry (1997) includes the possible "same-sex parent" effect.

[43] For those 10 individuals that quit smoking, the average difference between SSP in the second wave and in the first wave is positive (18.2). For those 3 individuals that started to smoke, this average difference is negative (-10).

5.2.1. Subjective Health Status

We next focus our attention on the last set of variables to be used in our econometric analysis: subjective health status. This variable is rather important to determine SSPs cross-sectionally, as shown in the previous Chapter, where we found that an individual with excellent health reports on average almost 19.5% more chances of reaching age 75 than someone claiming bad health in the first wave, and almost 14.5% in the second wave.

We believe that this variable requires a somewhat different approach when analysing its evolution from the first wave to the second wave of SHARE. We therefore follow the methodology proposed by Benitez-Silva and Ni (2008) based on HRS data.

Given that we expect that the evolution of expected longevity is a function of the evolution of health, it is key to decide how health is to be measured. Health can be measured by objective and subjective measures. As shown above, for the former we have included several instruments like onset of a set of severe and non-severe illnesses, the evolution of GALI, ADL and IADL indicators, and even health behaviour variables like drinking and sporting frequency.

Regarding subjective health measures, previous studies have shown that self-rated health plays a significant role in explaining individuals' actual mortality (Mossey and Shapiro, 1982; Idler and Benyamini, 1997; Miilunpalo *et al.*, 1997; Lee, 2000; Burstrom and Fredlund, 2001; Mete, 2005). They all found that self-rated health is an independent predictor of mortality, even controlling for objective measures of health. Ferraro and Kelley-Moore (2001) used panel data to argue that self-rated health can be considered a dynamic measure of health.

Benitez-Silva and Ni (2008) ask themselves whether self-rated health - while controlling for other objective health measures- captures all the important information regarding health that could help understand longevity expectations. They argue that a measure of *health dynamics* should be of value in understanding how individuals are assessing future states of the world.

SHARE dataset, as in the case of HRS and ELSA, contains two alternative questions related to the evolution of subjective health status. The first one is a "*computed* self-reported health change". For this question, in each wave, individuals are asked to assess their own health in each wave into five categories: excellent, very good, good, fair, and bad[44]. Therefore, the evolution of subjective health status can be computed as the difference between self-reported health status in the second wave and in the first wave.

Alternatively, SHARE dataset also includes a "self-reported health change" variable: individuals are asked: *"compared with your health when we talked with you in [previous interview], would you say that your health is much better, better, about the same, worse or much worse"*. Following Benitez-Silva and Ni (2008), we argue below that using differences in self-reported health status can be problematic since they can fail to capture meaningful health changes among respondents and, therefore, self-reported health change is preferred.

Self-reported health status can be thought as a continuous (latent) variable but it is discretized in a particular way (excellent to poor category). Therefore, two arguments can be raised against the idea that self-reported health status alone should be used in empirical models when trying to control for subjective health dynamics: heterogeneity and information loss. We next examine these two arguments.

Heterogeneity in reporting health status across respondents over-time can result in different self-reported health level for individuals with the same latent health status, because of lack of clear cut differences between categories proposed to the respondents. Figure Figure 3 presents panels (a) to (d) where we illustrate the inconsistencies that could arise.

Panel (a) shows a situation where an individual clearly knows his/her cut points, and the health categories are clearly differentiated. Most of the empirical studies assume that all individuals are of this type.

Panel (b) presents the case where individuals can distinguish between the different health categories but the cut points may shift over time, possible due to change in health reference group (peer effect[45]) or even changes in their own understanding of health.

[44] As in Chapter 1 we have decided to use the US version of the self-perceived health question because respondents in the second wave are only asked for this version.

[45] Self-reported health status is likely to be the result of a comparison between a reference group. This means that someone might report that they are in good health, while what they really mean is that they are

Therefore, although this type of individuals does not have problems distinguishing between categories, self-reported health status might change even if health status does not vary.

Panel (c) exemplifies a situation where individuals can not clearly distinguish between the different health categories. These individuals have what we call *grey areas*, like the one between $H_{3,l}$ and $H_{3,r}$. If the actual health status falls within these areas, an individual cannot clearly categorize it and the answer would be arbitrary. For example this is the case for individual I_1. Her real health status (between $H_{3,l}$ and $H_{3,r}$) is the same in both periods but she can report to have good or very good health in each period. Therefore, the evolution of her self-reported health can show that her health has improved (from good to very good), or has worsened (from very good to good) or has stayed constant. Clearly only the latter option is the correct.

Panel (d) depicts an extreme situation where grey areas might be changing over time for a given individual. We can therefore observe that self-reported health status changes even if health status does not change and that self-reported health status remains constant even if health status changes.

Information loss can arise even if individuals do not change their way to evaluate personal health status. For example, in panel (a), individual I_2 has a real health status $I_{2,t}$in period t, and either $I_{2,t+1}$ or $I'_{2,t}$ in period t+1. Even though her real health status has changed, the computed self-reported health change will show that her health has remained constant. For this reason it is expected that "self-reported health change" variable is a better alternative to use when controlling for subjective health dynamics. Because the question is related to personal health, peer effects are mitigated. Also the cut point shift issue is avoided as the assessment of the health change is not category-specific.

in good health compared with the reference group.The theory of reference groups has a long tradition in Sociology, and more recently with some empirical applications, for example in Bank *et al.* (1990).

6. RESULTS

As mentioned above we have modelled the evolution of SSPs from wave 1 to wave 2 in two different fashions: (i) we have considered as the independent variable the reported expected longevity in the second wave. This is determined as a function of the reported expected longevity in the first wave and variables that capture changes in personal characteristics between waves. (ii) we have also considered as the independent variable the change in the reported expected longevity between waves (wave 2 – wave 1). This is determined as a function of variables that capture changes in personal characteristics between waves.

For both type of specifications we included four different models with the objective of capturing the different effects that may arise. Specifically, model 1 controls for socio-demographic characteristics and parental mortality. Model 2 also incorporates objective health variables. We group the onset of illnesses into two categories: severe and non-severe[46]. Model 3further incorporates self-reported health change. Finally, model 4 also includes a disaggregation of the illnesses in order to study the effect of the onset of each one.

In all the cases we have estimated the relationship using an OLS approach consistent with previous studies (Hurd and McGarry, 1997; Hurd *et* al., 2001; Smith *et al.*, 2001; Liu *et al.*, 2007; Benitez-Silva and Ni, 2008). We also report robust standard deviations. Results are shown in Table 8 and in Table 9.

6.1. SPECIFICATION 1.MODELLING SSP IN WAVE 2

Next we present the results when we consider as the independent variable the reported expected longevity in the second wave as a function of the reported expected longevity in the first wave and variables that capture changes in personal characteristics between waves.

[46] The difference between severe and non-severe illnesses is that the formers imply a significant decrease in life expectancy according to medical criteria. See section 4.1.2 in Chapter 1 for further details.

We also present the results of comparable studies and contrast them with our results. Liu *et al.* (2007) for Taiwan, and Benitez-Silva and Ni (2008) and Smith *et al.* (2001) for the HRS also use a similar specification when modelling the transition of expected longevity.

The first interesting result that we observe is that in all four models the effect of SSP in the first wave is positive and quite significant. Following the information shown in Table 6, it was expected to observe persistency in SSPs along waves. For Taiwan Liu *et al.* (2007) found similar results. They found that SSP in the first wave had a positive and significant effect on that of the second wave, irrespectively of the other variables included. Interestingly they also found a similar positive and significant effect for SSPs up to age 85, though the effect is much smaller[47]. For the HRS, Benitez-Silva and Ni (2008) found that SSP in the previous wave had a positive and significant effect on current expected longevity. Similar results were also found by Smith *et al.* (2001).

As we expected *a priori*, based on the evidence presented in Section 3, age has a positive effect on the expected longevity in wave 2, however we do not significance. Our results are in line with results obtained by Smith *et al.* (2001) and Hurd *et al.* (2001). Additionally, we have found that increase in income quartile is associated with increase in expected longevity, though no significant effect is found[48]. This result is in line with that of the first Chapter where we found that cross-sectionally, income determines SSP. Interestingly, none of the studies revised consider changes in income when modelling the evolution of expected longevity.

Regarding parental mortality, overall, we find mixed results. In the first model we find that same-sex parent hypothesis seems to hold for females but we are not able to find significance for males. Interestingly, when self-reported health change is incorporated (model 3 and model 4), we find a significant negative effect for male respondents whose mother has died. Other studies present mixed evidence; Smith *et al.* (2001) does not concentrate in changes in parental mortality but in current parental mortality status. The authors found that irrespectively of the sex of the respondent, having a mother alive

[47] They authors found a coefficient between 0.235 and 0.239 depending on the specification. In the case of SSP related to a target age of 85 years old, the coefficient ranges from 0.089 to 0.091. In every case the coefficients are significant.

[48] Alternatively we have also considered the change in total income as the independent variable but the results do not vary.

increases expected longevity. Alternatively, Benitez-Silva *et al.* (2008) concluded that changes in parental mortality between waves do not affect expected longevity. We were not able to obtain any study that tests the "same-sex parent" hypothesis.

Model 2 incorporates objective health variables in the analysis. Our results also indicate that the evolution of mental health has a significant effect on the evolution of expected longevity. Concretely, we have used the EURO-D indicator, which is a global measure of mental health. We obtain that the onset of one additional mental health symptom between wave 1 and wave 2 on average decreases reported SSP between 0.54 and 0.67. As in the case of income, none of the studies revised consider changes in mental health as a possible determinant of the evolution of expected longevity.

We have also considered the evolution of the presence of a long-term illness. As mentioned in Section 4, *a priori* it is not straightforward to determine the effect that the evolution of long-term illness will have on the evolution of SSPs. One can think of two different effects. On the one hand, the onset of a long-term illness could increase expected longevity (as suggested in Chapter 1) because that person has already survived some years. But at the same time one can think that onset of a long-term illness could have a negative effect because health deteriorates. Similarly, one could believe that if someone healed of a long-term illness, she would increase her survival chances because she has survived to a long-standing illness. But, on the other hand, the health of that individual could have deteriorated as a consequence of the illness so that expected longevity decreased.

Our results indicate that the onset of a long-term illness has a negative effect on the evolution of SSPs (therefore, healing of a long-term illness has a positive effect on expected longevity), though we do not find that this effect is significant irrespectively of the model specification.

The results drawn from our model suggest that the effect of an increase in the number of limitations related with ADL is different than that of IADL. A similar result was found in Chapter 1, using a cross-sectional approach. We find that if a person starts suffering from an additional IADL limitation between wave 1 and wave 2, her SSP in the second wave decreases significantly compared to that of the first wave. Contrary, an additional ADL limitation does not significantly affects expected longevity in the second wave. For the

HRS, Benitez-Silva and Ni (2008) found changes in ADL limitation not related with muscles do have a significant negative effect on SSP. Smith *et al.* (2001) also found that additional number of activity limitations has a significant negative effect on expected longevity.

We have also found that changes in the GALI indicator have a non-significant negative effect the evolution of expected longevity. The same result is found for decreases in the frequency of consumption of alcohol. Regarding this last results, it should be said that Benitez-Silva and Ni (2008) also conclude that starting to drink has a positive effect on SSPs, though they did not find significance. Finally, we find that increases in the frequency of sport activity between wave 1 and wave 2 are significantly related with higher expected longevity in the second wave. For Taiwan Liu *et al.* (2007) found that physician advice on quitting smoking, drinking, oral hygiene, weight control and diet and nutrition decreased expected longevity in the second wave. Additionally, Smith *et al.* (2001) concluded that smoking also decreases expected longevity.

When considering changes in illnesses that an individual has suffered we have group illnesses in two categories (severe and non-severe illnesses) and, alternatively, we have considered changes in some specific illnesses. Our results indicate that an increase in the number of illnesses between wave 1 and wave 2 decreases SSPs in the second wave. Irrespectively of the model specification we find that an additional non-severe illness produces a greater effect than an additional severe illness. Smith *et al.* (2001) also found that a health shock between waves negatively affects expected longevity but the authors only find significance for health shocks between wave 2 and wave 3, but not for health shocks between wave 1 and wave 2. Similarly, Liu *et al.* (2007) concluded that the number of health shocks between waves has a negative effect on expected longevity but this effect is not significant.

On specific illnesses our results suggest that the onset of a liver cancer decreases expected longevity by 24%. This result is in line with the empirical evidence. According to the American Cancer Society, the overall five-year relative liver cancer survival rate for 1995-2001 was only 9.0%[49]. Our results also indicate that, on average, the onset of a brain

[49] The liver cancer survival rate is based on the relative liver cancer survival rate, which measures the survival of the cancer patients in comparison to the general population to estimate the effect of cancer.

cancer significantly reduces expected longevity by 16.6%. Again, this result seems sensible, as the American Cancer Society estimated that the overall five-year relative brain cancer survival rate for 1995-2001 was only 33%.

As expected, we also find that, on average, suffering a stroke decreases 10% SSP. In the case of a heart attack the decrease is lower (4.3%). Empirical evidence presented by the American Heart Association in its Heart Disease and Stroke Statistics (2007) mentions that in 2004 the overall death rate for stroke was 50%. The National Centre for Health Statistics in the US estimated that if all forms of major Cardio Vascular diseases were eliminated, life expectancy would rise by almost seven years.

Surprisingly, we find that a newly diagnosed breast cancer actually increases life expectancy by around 10%. This result could be explained by the fact that survival rate for this type of cancer has increased substantially in the last three decades, and it is expected to continue to increase. As reported by Coleman *et al.* (1999) and Richard (2008) for women diagnosed with breast cancer in 2001-2006, five-year relative survival rates have reached 82% (England only) compared with only 52% thirty years earlier in 1971-75. Additionally, Rachet *et al.* (2009) estimated that breast cancer survival rate in 2001-2006 for women aged 50 and 59 (this age range corresponds to the majority of the females in our sample) is almost 90%.

Other studies have also estimated the effect of specific illnesses in expected longevity. Liu *et al.* (2007) concluded that the onset of new health conditions between wave 1 and wave 2 (hypertension, thyroid disease, heart disease, hepatitis, hyperlipidaemia and gout) and newly abnormal health test outcomes decrease SSP in the second wave. Surprisingly, Benitez-Silva and Ni (2008) concluded that newly diagnosed high blood pressure increases expected longevity.

As expected, we have found that an individual's self-reported health change has a prominent role in the evolution of SSPs. More importantly, when this variable is incorporated, the effect of objective health variables diminishes. We find that moving

The American Cancer Society estimated that 18,510 men and women (12,600 men and 5,910 women) were diagnosed with liver cancer and that 16,200 men and women died of liver cancer in 2006.

from poor to excellent self-reported health change, on average, increases expected longevity by around 12%. Benitez-Silva and Ni (2008) found similar results.

6.2. SPECIFICATION 2.MODELLING THE CHANGE IN SSP

We present the results obtained when we consider as the independent variable the change in the reported expected longevity (the difference between expected longevity in the second wave and that of the first wave) as a function of variables that capture changes in personal characteristics between these waves.

As in the case of the previous specification, we also present the results of comparable studies and contrast them with our results. Benitez-Silva and Ni (2008), Hurd and McGarry (1997) and Smith *et al.* (2001) for the HRS, and Hurd *et al.* (2001) for the AHEAD also use a similar specification when modelling the transition of expected longevity.

Firstly, we find that the change in age of the respondent negatively affects the change in expected longevity but not significantly. The same result was found by Smith *et al.* (2001). As in the previous specification, we find that changes in income do not significantly determine changes in SSPs, regardless of the model considered.

Regarding parental mortality we observed that when we control for the "same-sex parent" hypothesis, we conclude that males are more affected by father's death and females are more affected by mother's death. Smith *et al.* (2001) and Benitez-Silva and Ni (2008) do not distinguish between the sexes of the respondents and found that newly father's death and newly mother's death do not significantly affect changes in expected longevity. Hurd and McGarry (1997) concluded that the "same-sex parent" hypothesis does not hold for the HRS but found that the age at dead of the father and mother does determine expected longevity.

Our results also suggest that an additional depression symptom, the onset of a long-term illness, an additional IADL limitation, positive changes in the GALI indicator, and decreases in the frequency of sport activity have a negative effect on expected longevity. Hurd *et al.* (2001) also found that the onset of depression decreases expected longevity;

additionally, Smith *et al.* (2001) concluded that an increase in the number of activity limitations decreases SSPs.

As in the previous specification we consider changes in illnesses that an individual has suffered grouped in two categories (severe and non-severe illnesses) and, alternatively, we have considered changes in some specific illnesses. Our results suggest that an increase in the number of illnesses between wave 1 and wave 2 decreases SSPs in the second wave, though we do not find significance. On specific illnesses, we find that suffering a stroke and the onset of a liver cancer significantly decreases expected longevity. Surprisingly, our results suggest that newly diagnosed high cholesterol increases SSPs, probably because people are more willing to modify their lifestyle after the diagnose. Hurd and McGarry (1997) found that a newly diagnosed cancer decreases expected longevity. Hurd *et al.* (2001) for the AHEAD found that a newly diagnosed cancer, high blood pressure, lung disease, and diabetes decreases expected longevity. Contrary, Benitez-Silva and Ni (2008) concluded that diabetes increases expected longevity.

Similar to the previous specification, when individual's self-reported health change is incorporated, the effect of objective health variables diminishes. Interestingly, we only find significance of self-reported health change for model 3. When we disaggregate severe and non-severe illnesses (model 4), it becomes non-significant.

7. CONCLUSIONS

In the previous Chapter we concentrated on a cross-sectional econometric analysis in order to examine the determinants of the SSPs in the SHARE database. We concluded that from the cross-sectional point of view, SSPs were coherent and behave reasonably well. But the result obtained in Chapter 1 is not enough it we want to use SSPs to estimate models of decision-making under uncertainty. According to Hurd and McGarry (1997) before we can confidently use individual-level probability distributions, additionally we need to analyse whether individuals adjust their reported probabilities in response to new information and whether the reported probabilities predict outcome.

Regarding the latter, we have found that in our database that only 81 respondents have died between waves out of 15,058 individuals. Therefore the results obtained should be taken with careful as they might not be stable enough. In any case, we found that SSPs do predict mortality. Moreover, we find that average SSP for those that survived is substantially higher than for those that died.

In this Chapter we have also study how SSPs evolve over time as new information comes to the respondent. We address how SSP updates with new information, whether this is medical, socio-economical, or personal. And we need to address whether this update is coherent with empirical evidence and with the evidence presented in other studies based on similar surveys.

Our results suggest that expected longevity shows persistency along waves. Past expected longevity has a significant weight in present expected longevity. We have also found that the evolution of mental health has a significant effect on the evolution of expected longevity and that the effect of an increase in the number of limitations related with ADL is different than that of IADL. Our results also suggest that increases in the frequency of sport activity between wave 1 and wave 2 are significantly related with higher expected longevity in the second wave.

Our results indicate that health shocks, an increase in the number of illnesses between wave 1 and wave 2, decreases SSPs in the second wave. Specifically, we find that the onset of a liver cancer decreases expected longevity by 24%, the onset of a brain cancer significantly reduces expected longevity by 16.6%, suffering a stock decreases 10% SSP. In the case of a heart attack the decrease is lower (4.3%). Surprisingly we find that a newly diagnosed breast cancer actually increases life expectancy by around 10%. This result could be explained by the fact that survival rate for this type of cancer has increased substantially in the last 3 decades, and it is expected to continue to increase. Additionally, we found that an individual's self-reported health change has a prominent role in the evolution of SSPs and that the "same-sex parent hypothesis" seems to hold.

Overall, we have found that the evolution of SSPs is coherent with epidemiological evidence and with previous studies done based on other surveys. We conclude that individuals adjust their reported expected longevity in response to new information and that this adjustment is consistent. Additionally, we have found that SSPs predict mortality.

Therefore, we can conclude that this Chapter provides results that increase our confidence that SSPs can be used in models of decision-making under uncertainty.

8. REFERENCES

- American Heart Association (2007). "Heart Disease and Stroke Statistics-2007 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee". *Circulation: Journal of the American Heart Association,* 115: e69-e171.

- Balia, S. (2007). "Reporting Expected Longevity and Smoking: Evidence from Share". Centro Ricerche Economiche Nord Sud (CRENOS), *Working Paper,* 2007/05.

- Bank, B., Biddle, B., Slaving, R. (1990). "What Do Students Want?".*The Sociological Quarterly,* 33(3): 321-335.

- Becker, G. S., Murphy, K. (1988). "A Theory of Rational Addiction". *Journal of Political Economy*, 96(4): 675-700.

- Benitez-Silva, H., Ni, H. (2008). "Health Status and Health Dynamics in an Empirical Model of Expected Longevity". *Journal of Health Economics,* 27(3): 564-584.

- Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (2008). *Health, Ageing and Retirement in Europe (2004-2007). Starting the Longitudinal Dimension*. Mannheim, MEA: 366.

- Burstrom, B., Fredlund, P. (2001). "Self-Rated Health: Is it as Good a Predictor of Subsequent Mortality among Adults in Lower as Well as in Higher Social Classes?".*Journal of Epidemiology and Community Health*, 55: 836-840.

- Coleman, M., Babb, P., Damiecki, P. (1999). *Cancer Survival Trends in England and Wales, 1971-1995: Deprivation and NHS Region*. London, The Stationery Office, 695.

- Ferraro, K. F., Kelley-Moore, J. (2001). "Self-Rated Health and Mortality among Black and White Adults; Examining the Dynamic Evaluation Thesis". *Journal of Gerontology: Social Science,* 56B(4): S195-S205.

- Hurd, M., McGarry, K. (1997). "The Predictive Validity of Subjective Probabilities of Survival". *The Economic Journal,* 112(482): 966-985.

- Hurd, M., Mcfadden, D., Merrill, A. (2001). "Predictors of Mortality among the Elderly". In: Wise, D. (ed.), *Themes in the Economics of Aging*, Chicago, University of Chicago Press: 171-198.

- Idler, E. L., Benyamini, Y. (1997). "Self-Rate Health and Mortality: A Review of Twenty-Seven Community Studies". *Journal of Health and Social Behaviour*, 38(1): 21-37.

- Lee, Y. (2000). "The Predictive Value of Self Assessed General, Physical, and Mental Health on Functional Decline and Mortality in Older Adults". *Journal of Epidemiology and Community Health*, 54: 123-129.

- Liu, J., Tsou, M. Hammitt, J. (2007). "Health Information and Subjective Survival Probability: Evidence from Taiwan". *Journal of Risk Research,* 10(2): 149-175.

- Mete, C. (2005). "Predictors of Elderly Mortality: Health Status, Socioeconomic Characteristics and Social Determinants of Health". *Health Economics,* 14: 135-148.

- Miilunpalo, S., I. Vuori, P. Oja, M. Pasanen, H. Urponen (1997). "Self-Rated Health Status as a Health Measure: The Predictive Value of Self-Reported Health Status on the Use of Physician Services and on Mortality in the Working-Age Population". *Journal of Clinical Epidemiology*, 50(5): 517-528.

- Mossey, J., Shapiro, E. (1982). "Self Rated Health: a Predictor of Mortality among the Elderly". *American Journal of Public Health*, 72(8): 800-808.

- Rachet, B., Maringe, C., Nur, U., Quaresma, M., Shah, A., Woods, L. M., Ellis, L., Walters, S., Forman, D., Steward, J., Coleman, M. P. (2009). "Population-Based

Cancer Survival Trends in England and Wales up to 2007: an Assessment of the NHS Cancer Plan for England". *The Lancet Oncology*, 10(4): 351-369.

- Richard, M. A. (2008). "Trends and Inequalities in Survival for 20 Cancers in England and Wales 1986-2001: Population-Based Analyses and Clinical Commentaries". *British Journal of Cancer*, 99(1): S1-S120.

- Siegel, M., Bradley, E., Kasl, S. (2003). "Self-Rated Life Expectancy as a Predictor of Mortality: Evidence from the HRS and AHEAD surveys". *Gerontology*, 49(4): 265-271.

- Sloan, F., Taylor, D., Smith, K. (1999), "Are Smokers Too Optimistic?". In: Grossman, M. and Hsieh, C. R. (Ed.), *Economic Analysis of Substance Use and Abuse: The Experience of Developed Countries and Lessons for Developing Countries*, Northampton, Edward Elgar Publishers: 103-134.

- Smith, K., Taylor, D., Sloan, F. (2001). "Longevity Expectations and Death: Can People Predict their own Demise?".*The American Economic Review*, 91(4): 1126-1134.

- Steffen, B. (2009), "Formation and Updating of Subjective Life Expectancy: Evidence from Germany", *MEA Studies*, 8.

- Thaler, R., Sheffrin, H. (1981). "An Economic Theory of Self-Control". *Journal of Political Economy*, 89(2): 392-406.

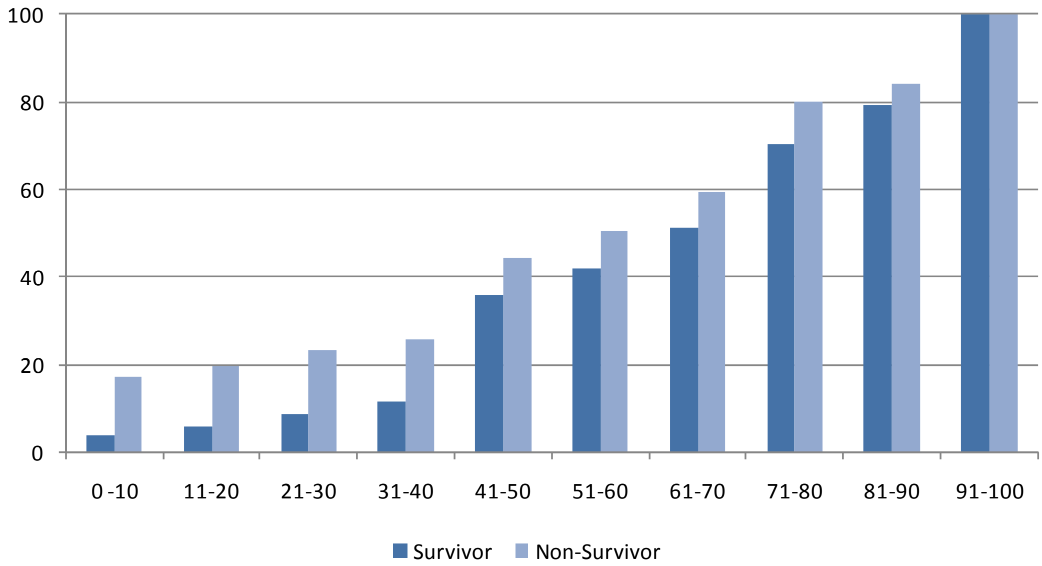
- Van Doorn, C., Kasl, S. (1998). "Can Parental Longevity and Self-Rated Life Expectancy Predict Mortality among Older Persons? Results from an Australian Cohort". *Journal of Gerontology: Social Sciences*, 53B(1): S28-S34.

- Viscusi, W. (1985). "Are Individuals Bayesian Decision Makers?". *American Economic Review*, 75(2): 381-85.

- Viscusi, W., Hakes, J. (2003). "Risk Rating That Do Not Measure Probabilities". *Journal of Risk Research*, 6: 23-43.

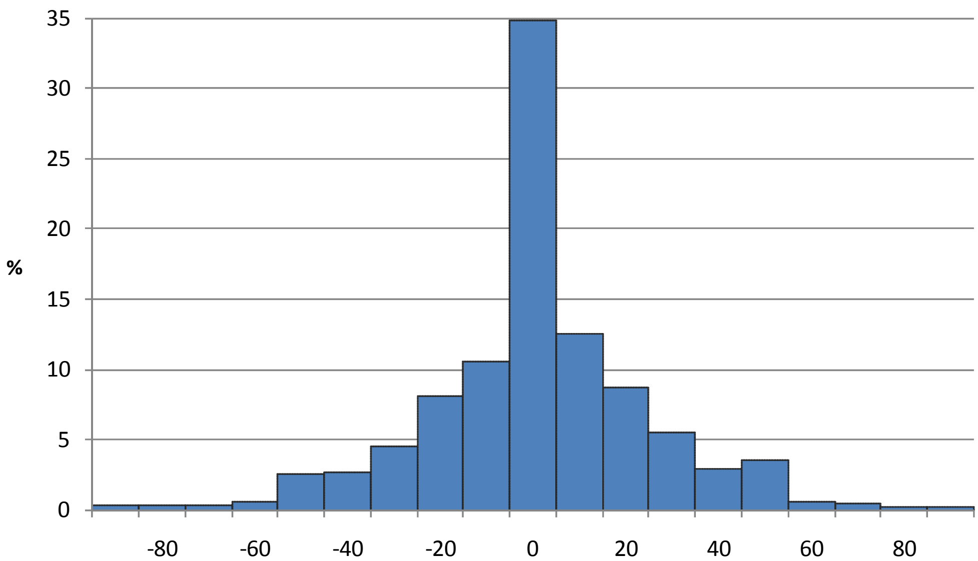
9. TABLES AND FIGURES

FIGURE 1: CUMULATIVE DISTRIBUTION OF SSP



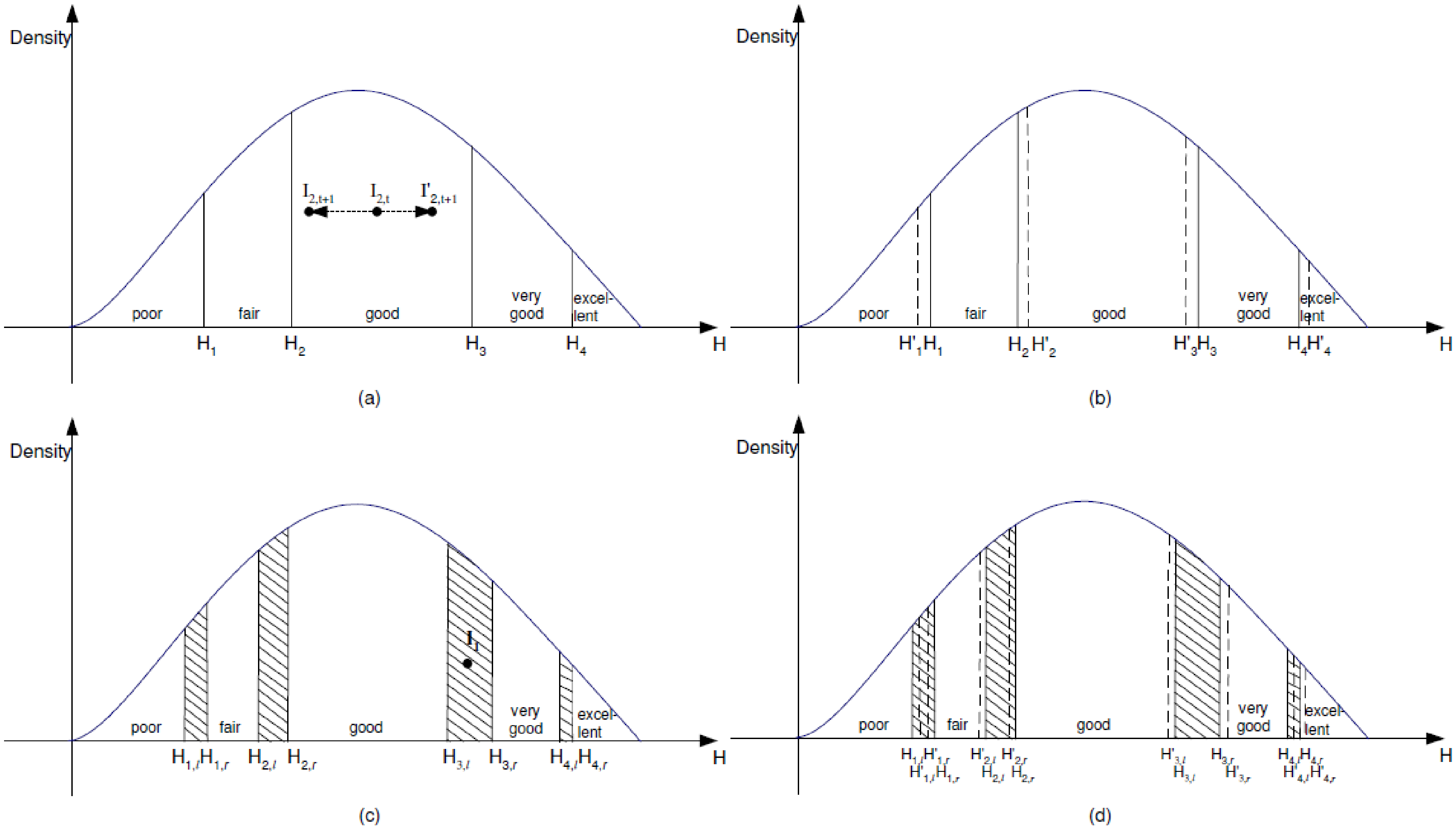
Source: own elaboration

FIGURE 2: CHANGE IN SSP BETWEEN WAVES (WAVE 2 – WAVE 1)



Source: own elaboration

FIGURE 3: SELF-REPORTED HEALTH STATUS



Source: Benitez-Silva and Ni (2008)

TABLE 1: MORTALITY RATE BY COUNTRY (WHOLE SAMPLE)

Country	Total	Died	Mortality rate
Austria	1,893	36	1.90%
Belgium	3,827	36	0.94%
Denmark	1,707	50	2.93%
France	3,193	53	1.66%
Germany	3,008	46	1.53%
Greece	2,898	49	1.69%
Italy	2,559	56	2.19%
Netherlands	2,979	49	1.64%
Spain	2,396	87	3.63%
Sweden	3,053	57	1.87%
Switzerland	1,004	14	1.39%
TOTAL	28,517	533	1.87%

Source: own elaboration

TABLE 2: MORTALITY RATE BY AGE GROUP (WHOLE SAMPLE)

Age group	Total	Died	Mortality rate
Under 50	1,179	0	0.00%
50- 65	15,697	100	0.64%
66 – 75	7,081	132	1.86%
Over 75	4,560	301	6.60%
TOTAL	28,517	533	1.87%

Source: own elaboration

TABLE 3: VARIABLES BY SURVIVORSHIP STATUS

	Died between waves		Survived to wave 2	
Variable	**Mean**	**St. Dev.**	**Mean**	**St. Dev.**
SSP	58.3	31.6	68.5	25.0
Age	59.4	4.1	56.8	4.9
Income ('000)	37.2	29.9	51.1	147.7
Male	0.70	0.45	0.44	0.49
Married	0.67	0.47	0.81	0.39
Self-reported health	3.53	1.04	2.71	1.03
Education	2.64	1.45	2.85	1.47
Euro-D	2.74	2.62	2.07	2.1
Long-term illness	0.64	0.48	0.41	0.49
ADL	0.35	1.17	0.08	0.45
IADL	0.48	1.31	0.12	0.5
Severe illnesses	1.31	1.23	0.71	0.95
Non-severe illnesses	0.65	0.87	0.33	0.6
GALI	0.65	0.47	0.33	0.47
Current smoker	0.41	0.49	0.25	0.43
Sports	2.79	1.33	2.27	1.29
Observations	81		14,977	

Source: own elaboration

TABLE 4: SSP IN THE TWO WAVES

SSP	**Wave 1**	**Wave 2**
Mean	69.08	70.19
Std. Dev.	24.41	23.86
Quartile 1	50	50
Quartile 2	70	75
Quartile 3	90	90
Quartile 4	100	100
Observations	8,484	

Source: own elaboration

TABLE 5: SSP COMPARISON IN THE TWO WAVES

Probability comparison	SSP	
Wave2 > Wave1	3,162	37.27%
Wave2 < Wave1	2,762	32.56%
Wave2 = Wave1	2,560	30.17%
Both prob. = 0	29	0.34%
Both prob.= 50	833	9.82%
Both prob. = 1000	742	8.75%
Both prob. other values	956	11.27%
Observations	8,484	

Source: own elaboration

TABLE 6: SSP IN WAVE 2 CONDITIONAL ON SSP IN WAVE 1

SSP wave 1	Mean SSP wave 2
0	35.7
10	40.5
20	46.1
30	52.3
40	55.7
50	61.7
60	65.4
70	70.7
80	75.1
90	80.4
100	83.9

Source: own elaboration

TABLE 7: ONSET OF NEW HEALTH CONDITIONS

New health condition		Not suffering in wave 1	New cases in wave 2	Incidence between waves
Heart attack		7,944	211	2.66%
High blood pressure		6,458	711	11.01%
High cholesterol		6,884	661	9.60%
A stoke		8,338	67	0.80%
Diabetes		7,938	201	2.53%
Lung disease		8,182	164	2.00%
Asthma		8,119	117	1.44%
Arthritis		7,201	613	8.51%
Osteoporosis		8,070	242	3.00%
Malignant tumour		8,169	112	1.37%
Stomach ulcer		8,064	119	1.48%
Parkinson		8,468	10	0.12%
Cataracts		8,323	108	1.30%
Hip fracture		8,408	29	0.34%
CANCER	Brain	8,476	5	0.06%
	Oral cavity	8,483	2	0.02%
	Larynx	8,480	2	0.02%
	Other pharynx	8,481	3	0.04%
	Thyroid	8,476	2	0.02%
	Lung	8,472	9	0.11%
	Breast	4,639	23	0.50%
	Esophagus	8,482	1	0.01%
	Stomach	8,478	2	0.02%
	Liver	8,481	4	0.05%
	Kidney	8,480	4	0.05%
	Prostate	3,706	16	0.43%
	Ovary	4,743	6	0.13%
	Cervix	8,461	5	0.06%
	Endometrium	8,470	3	0.04%
	Colon or Rectum	8,460	21	0.25%
	Bladder	8,479	6	0.07%
	Skin	8,459	11	0.13%
	Non-Hodgkin lymphoma	8,476	7	0.08%
	Leukemia	8,478	3	0.04%

Note: The table excludes pancreas and testicle cancer as their incidence was zero. Source: own elaboration

TABLE 8: RESULTS SPECIFICATION 1 (OLS ON SSP WAVE 2)[50]

SSP wave 2	MODEL 1		MODEL 2		MODEL 3		MODEL 4	
	Coefficient	Std. Dev.	Coefficient	Std. Dev.	Coefficient	Std. Dev.	Coefficient	Std. Dev.
SSP wave 1	0.47 ***	0.01	0.47 ***	0.01	0.46 ***	0.01	0.46 ***	0.01
Change in age	0.19	0.38	0.20	0.37	0.27	0.38	0.27	0.38
Change income quartile	0.06	0.21	0.10	0.21	0.06	0.21	0.05	0.21
Change Euro-D			-0.66 ***	0.13	-0.55 ***	0.13	-0.54 ***	0.13
Chance long-term illness			-0.79	0.50	-0.58	0.51	-0.58	0.50
Change ADL			-0.04	0.69	0.23	0.68	0.38	0.68
Change IADL			-3.02 ***	0.67	-2.73 ***	0.67	-2.45 ***	0.65
Change severe illness			-1.28 **	0.50	-0.88 *	0.51		
heart attack							-4.37 **	1.78
high blood pressure							-0.02	0.89
high cholesterol							0.68	0.87

[50]Model 1: controls for socio-demographic characteristics and parental mortality ("same-sex parent").

Model 2: also incorporates changes in objective health variables (EURO-D, long-term illness, ADL, IADL, severe and non-severe illnesses, GALI, drinking behaviour and practicing sport) in the analysis.

Model 3: further includes self-reported health change.

Model 4: additionally, includes individual illnesses.

a stoke										-10.84	***	3.08
lung disease										-3.92	**	1.92
malignant tumour										-5.01		3.21
brain cancer										-16.62	**	7.43
breast cancer										10.76	*	5.57
liver cancer										-24.00	***	4.98
Change non-severe illness				-1.69	**	0.67	-1.16	*	0.67			
asthma										-2.01		2.18
Arthritis										-0.29		0.99
osteoporosis										-3.47	**	1.50
stomach ulcer										-2.11		2.28
prostate cancer										-8.49		8.11
Change GALI				-0.72		0.51	-0.41		0.52	-0.33		0.52
Change drinking				-0.41		0.31	-0.30		0.31	-0.25		0.31
Change sports				-0.42	**	0.17	-0.42	**	0.17	-0.40	**	0.17
Reported self-p. health change							2.96	***	0.39	2.73	***	0.39
male x father dead	-1.90		1.51	-1.66		1.47	-2.14		1.48	-2.27		1.48
male x mother dead	-1.97		1.35	-2.08		1.32	-2.32	*	1.32	-2.14	*	1.31
female x mother dead	-2.29	*	1.26	-2.11	*	1.24	-2.55	**	1.26	-2.50	**	1.25
female x father dead	0.08		1.45	0.23		1.42	0.38		1.43	0.53		1.40
Constant	37.67	***	1.19	38.01	***	1.20	38.43	***	1.21	38.74	***	1.20
N	8,484			8,484			8,321			8,321		
R2	0.228			0.242			0.252			0.257		

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1.
Source: own elaboration

TABLE 9: RESULTS SPECIFICATION 2 (OLS ON THE CHANGE IN SSP)[51]

SSP wave 2 – SSP wave 1	MODEL 1		MODEL 2			MODEL 3			MODEL 4		
	Coefficient	Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.	Coefficient		Std. Dev.
Change in age	-0.69	0.44	-0.69		0.44	-0.62		0.44	-0.61		0.44
Change income quartile	0.14	0.25	0.19		0.25	0.16		0.25	0.15		0.25
Change Euro-D			-0.98	***	0.15	-0.96	***	0.15	-0.95	***	0.15
Chance long-term illness			-1.28	**	0.61	-1.27	**	0.61	-1.29	**	0.61
Change ADL			-0.82		0.83	-0.68		0.83	-0.65		0.83
Change IADL			-2.59	***	0.74	-2.60	***	0.75	-2.40	***	0.74
Change severe illness			-0.14		0.59	-0.06		0.60			
heart attack									-1.92		2.24
high blood pressure									0.06		1.06
high cholesterol									1.89	*	1.06
a stoke									-8.84	***	3.43
lung disease									0.05		2.30
malignant tumour									-4.78		3.54
brain cancer									-3.06		10.44
breast cancer									8.35		6.34

[51] Ibid.

liver cancer							-15.13 **	7.22
Change non-severe illness			-0.39	0.78	-0.33	0.78		
asthma							0.08	2.58
arthritis							-0.40	1.19
osteoporosis							-1.93	1.69
stomach ulcer							1.96	2.74
prostate cancer							-10.27	9.41
Change GALI			-1.11 *	0.62	-1.14 *	0.63	-1.08 *	0.63
Change drinking			-0.43	0.37	-0.35	0.37	-0.31	0.37
Change sports			-0.53 **	0.21	-0.59 ***	0.21	-0.58 ***	0.21
Reported self-p. health change					0.81 *	0.45	0.62	0.45
male x father dead	-3.26 *	1.74	-2.97 *	1.69	-3.25 *	1.69	-3.32 *	1.70
male x mother dead	-1.86	1.64	-1.95	1.61	-2.37	1.60	-2.22	1.60
female x mother dead	-4.56 ***	1.44	-4.41 ***	1.42	-5.00 ***	1.42	-4.91 ***	1.42
female x father dead	0.15	1.63	0.28	1.60	0.54	1.61	0.67	1.59
Constant	3.02 ***	1.07	3.01 ***	1.07	2.97 ***	1.08	2.94 ***	1.07
N	8,484		8,484		8,321		8,321	
R2	0.002		0.018		0.021		0.023	

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1.
Source: own elaboration

10. ANNEX

In order to test whether SSPs predict actual mortality we use a probit specification. As mentioned above, because we only observe 81 deceases out of 15,058 individuals in our restricted sample, the results obtained in this Annex should be taken with careful as they might not be stable enough. Once new waves are collected in the SHARE database we will be able to observe more deceases and the results would be more robust.

We model observed mortality between wave 1 and wave 2 as a function of SSP in wave 1. We also include other covariates such as age, gender, marital status[52], educational level, income quartiles, number of severe and non-severe illnesses, smoking behaviour[53], parental mortality[54], and self-rated health status.

Other studies have also used a similar approach. Hurd and McGarry (1997) and Siegel *et al.* (2003) used SSPs, self-rated health status, socio-demographic characteristics and health variables. Van Doorn and Kasl (1998) also included parental mortality and Smith *et al.* (2001) also included smoking behaviour. Finally, Hurd *et al.* (2001) included similar variables but instead of using SSPs, the authors used their deviation from Life Tables.

Most of these studies highlight the fact that both SSPs and self-rated health status independently do have a significant role predicting mortality. However, they mention that it is interesting to analyse how that role is diminished when both variables are included simultaneously. For this reason we present three different models. The first model uses SSPs, socio-demographic variables and parental mortality to model mortality. The second model adds objective health variables (number of severe and non-severe illnesses). Finally, the last model also includes subjective health status. The results are shown in Table 10.

We have found evidence that in models 1 and 2 SSPs do significantly predict mortality. Lower expected survival chance is associated with higher probability of decease. However, when subjective health status is included, the effect of SSPs diminishes. Similar results

[52] Dummy variable equals to one if the respondent is married, 0 otherwise.

[53] Dummy variable equals to one if the respondent is currently smokes, 0 otherwise.

[54] Dummy variable equals to one if at least one of the respondent's parents is dead, 0 otherwise.

were found by Hurd and McGarry (1997), Hurd *et al.* (2001) and Siegel *et al* (2003). Contrary, Van Doorn and Kasl (1998) and Smith *et al.* (2001) did find that both measures have a combine significant effect on the probability of dying. Hurd and McGarry (1997) argued that even though self-assessed health status reduces the effect of SSP, self-assessed health status should not replace SSP in models of economic behaviour. We conclude that, even though the results are not stable enough, SSP does predict mortality. However, once more waves become available, further research should confirm this statement.

TABLE 10: DETERMINANTS OF MORTALITY, WAVE 1 TO WAVE 2 (PROBIT ON MORTALITY)[55]

Mortality = 1	MODEL 1		MODEL 2		MODEL 3	
	Coefficient	Std. Dev.	Coefficient	Std. Dev.	Coefficient	Std. Dev.
SSP	-0.005 ***	0.002	-0.003 *	0.002	-0.001	0.002
Age	0.041 ***	0.009	0.038 ***	0.009	0.040 ***	0.009
Gender	-0.382 ***	0.087	-0.402 ***	0.092	-0.409 ***	0.093
Married	-0.291 ***	0.088	-0.274 **	0.091	-0.277 ***	0.093
Education	-0.025	0.028	-0.009	0.028	0.005	0.028
Income Q2	0.017	0.115	0.045	0.118	0.059	0.120
Income Q3	-0.110	0.121	-0.094	0.126	-0.054	0.127
Income Q4	0.023	0.114	0.065	0.120	0.115	0.121
Parental mortality	-0.017	0.160	-0.038	0.163	-0.074	0.165
Severe illnesses			0.119 ***	0.035	0.070 **	0.036
Non-severe illnesses			0.158 ***	0.057	0.104 *	0.060
Smoking			0.245 ***	0.089	0.230 ***	0.090
Self-perceived health					0.213 ***	0.053
Constant	-3.777 ***	0.538	-4.004 ***	0.590	-4.841 ***	0.601
N	15,058		15,058		15,058	
Pseudo R2	0.0679		0.0966		0.1164	

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

[55] Model 1: includes SSPs, socio-demographic variables and parental mortality.
Model 2: also includes the number of severe and non-severe illnesses.
Model 3: also includes self-perceived health status.

CHAPTER 3.
THE EFFECT OF SUBJECTIVE SURVIVAL PROBABILITIES ON LABOUR SUPPLY DECISIONS

1. INTRODUCTION

As mentioned in the Introduction of the first Chapter, expectations are important inputs to economists' basic research. In particular, expectations on longevity play a significant role in the LCM, which we define below. However, before analysing the role of expectation in any economic model, it is essential to take one step backwards and analyse the expectations themselves. This is a rather important point as suggested by Hurd and McGarry (1997). They specified that before we can confidently use individual-level probability distributions, we need to understand and test their properties. They remarked three properties that should be satisfied:

1. Individuals understand questions about probabilities and accurately report their beliefs about the likelihood of future events;

2. Individuals adjust their reported probabilities in response to new information; and

3. The reported probabilities predict outcome.

The first two Chapters have concentrated on examining how SSP drawn from SHARE fulfil these conditions. Chapter 1 focused on the first condition proposed by Hurd and McGarry (1997). Specifically, we studied the determinants of SSP and its consistency with mortality and epidemiological data. We have found that SSPs in SHARE present less focal point (0%, 50% and 100%) answers with respect to other surveys like the HRS and ELSA. Moreover, using cross-sectional analysis, we concluded that SSPs do behave reasonably well. They average close to actuarial probabilities and they covary with other variables in the same way actual outcomes vary with the variables.

The second and third conditions proposed by Hurd and McGarry (1997) were analysed in Chapter 2. With respect to the second condition, we studied how SSPs evolve over time as new information comes to the respondent, whether this is medical, socio-economical, or personal. Our results suggest that expected longevity shows persistency along waves. Past expected longevity has a significant weight in present expected longevity. We also found that the evolution of mental health has a significant effect on the evolution of expected longevity. Our results also suggest that increases in the frequency of sport activity between wave 1 and wave 2 are significantly related with higher expected longevity in the second wave. Moreover, our results indicate that health shocks, an increase in the number of illnesses between wave 1 and wave 2, decreases SSPs in the second wave. We conclude that individuals adjust their reported expected longevity in response to new information and that this adjustment is consistent.

Regarding the third property, the second Chapter analysed whether expected longevity predicts actual mortality. Because we do not observe many deceases in our sample from wave 1 to wave 2, the results obtained should be taken with careful as they might not be stable enough. Once new waves are collected in the SHARE database we will be able to observe more deceases and the results would be more robust. In any case, we found that the average SSP for those that survived is substantially higher than for those that died. Furthermore, an econometric approach confirmed that SSP does predict mortality.

Overall, the first two Chapters have confirmed that SSPs in SHARE do fulfil the three properties proposed by Hurd and McGarry (1997). Therefore, we can conclude that the results obtained increase our confidence that SSPs can be used in models of inter-temporal decision models under uncertainty. In particular, the LCM provides an excellent framework to test whether economic agents take into account longevity expectations when taking economic decisions.

The LCM is the main theory for understanding household inter-temporal optimization. According to the LCM there is a well-defined link between the consumption plans of an individual and her income and income expectations as she passes from childhood, through the work participating years, into retirement and eventual decease. The hypothesis was developed by Modigliani and Brumberg (1954) in line with the Permanent Income Hypothesis developed by Friedman (1957) which states in its simplest form, that the choices made by consumers regarding their consumption patterns are determined not by current income but by their longer-term income expectations.

The LCM is based on the assumption that a rational economic agent maximizes lifetime utility. This lifetime utility is given by consumption and leisure. However, there is, obviously uncertainty on when someone will die, and here is where longevity expectations could play a significant role.

The utility maximizing agent will derive paths of consumption, saving, leisure, labour supply, and wealth from the dynamic optimization problem defined over a horizon, T. But how long is T? This is a very important element of the solution.

What is the role of SSPs in the LCM? In theory, a direct consequence of the LCM is that individuals who expect to live exceptionally longer will retire at a later age than individuals who expect to die early, as suggested, among others, by Attanasio (1998), Bloom *et al.* (2003), Hurd *et al.* (2004), Kalemli-Ozcan and Weil (2010), Perry (2005), Delavande *et al.* (2006), Salm (2006), Bloom *et al.* (2006, 2007), O'Donnell *et al.* (2008), and Cocco and Gomes (2012). The reason why those who expect to be long-lived will retire later is because they will need more wealth to finance more years of retirement, as income from a pension is substantially lower than labour income. This Chapter examines whether this theoretical premise holds in real life for SHARE respondents.

The objective of this Chapter is to examine whether SSPs play a significant role on labour supply decisions of Europeans aged under 65 using the first and the second wave of the SHARE database. We would expect to observe that individuals who expect to die earlier will retire sooner than individuals who expect to live longer, as suggested by several authors.

We believe that understanding how expected longevity affects labour supply decisions is a crucial topic for policymakers. In an environment of high unemployment rate, low fertility rate and increasing life expectancy there is increasing financial pressure on Social Security systems across Europe and most of European governments have already increased the official retirement age or are seriously considering doing so. Policymakers should be interested in understanding how to incentive people to work longer in a situation characterized by increasing longevity, so that the sustainability of the system is secured. As mentioned by O'Donnell *et al.* (2008), understanding how survival expectations influence retirement is becoming increasingly important in the context of continued increase in life expectancy and the perception that this threatens the financial sustainability of pension systems.

The rest of the Chapter is structured as follows: Section 2 presents a detailed literature review. Section 3 presents the data used in the analysis and a descriptive analysis. The econometric modelling strategy is shown in Section 4 and the results are presented in

Section 5. Section 6 includes a sensitivity analysis and finally, the conclusions are shown in section 7.

2. LITERATURE REVIEW

In this section we present a literature review relevant for this Chapter. We have therefore focus on works that use micro databases principally. We have organized this literature review into two main blocks. We start with the relevant works that consider the implication of the LCM on variables such as saving, consumption, retirement and bequests. We then explore the literature on the relationship between longevity expectations and retirement behaviour in the LCM context.

2.1. LCM, SAVING, AND CONSUMPTION

According to the LCM, expected longevity should play a prominent role on many economic decisions that economic agents take. The economic literature has typically focused on decisions such as saving and consumption. Below we present a detailed literature review on these issues. Overall, we find that economic literature has concluded that saving and consumption behaviour of economic agents tends to be in line with the predictions of the LCM. However, it should also be considered that uncertainty could affect economic decisions.

2.1.1. LCM and saving

In a very brief study Gan *et al.* (1998) questioned whether SSPs influence, or at least vary with, economic behaviour. In particular, they used the first wave of the AHEAD database to look at the relationship between stated saving behaviour and beliefs about survival. The link between these variables is the LCM that predicts that saving rate should be positively related to life expectancy. Saving behaviour is captures using qualitative information from the AHEAD database, as individuals are asked whether they are net savers, zero savers or net dissevers.

The authors used a simple ordered probit to ask whether, in addition to wealth and income, saving rates varies with respect to SSP. They presented three different estimations for the whole sample, couples, and single member households. For the whole sample, and for couples, Gan *et al.* (1998) found that saving behaviour does respond to SSP. Contrary, in the case of single member households, they obtained no significant effect but the correct sign.

The authors concluded that there is a significant positive correlation between SSP and saving. Moreover, combined with the optimistic bias about survival that increases with age, it gives one explanation for the fact that saving rates do not fall as rapidly with age as a classical LCM would suggest.

Bloom *et al.* (2003) investigated empirically whether changes in longevity play an important role in determining national savings, and in particular if the observed increase in saving rate in East Asia can be explained based on life expectancy increases. They developed a theoretical model where the age structure of the population is taken into account. They argued that in a stable population, net saving rate should be zero, as young adults save for retirement and retirees dis-save. However, this is not the case in a transitional population. Unfortunately, the authors ignore any effects of the uncertainty about the timing of death, even though they recognized that the effects may be important. They constructed a model of aggregate savings that includes life expectancy as an argument and estimated the parameters using cross-country panel data.

They concluded that improvements in health and longevity are likely to have a large impact on life-cycle behaviour. They argued that increases in longevity tend to increase the relative length of retirement, therefore increasing the need for retirement income which has the effect of higher saving rates among the working population, which explains the saving boom in East Asia during 1950-1990.

According to the LCM, consumption shows a smooth pattern along the lifespan of an individual if no unanticipated shocks occur. However, Hurd and Rohwedder (2003) pointed out that British and US households apparently reduce consumption at retirement and the reduction cannot be explained by the LCM. This is the so-called retirement-consumption puzzle.

With the help of the HRS and the Consumption and Activities Mails Surveys (CAMS), they explore the consumption patterns of individuals before and after retirement. However, due to the nature of the data, the study has a major drawback because the authors compare anticipated spending change before retirement and realized change for different individuals.

They argued that the decline in consumption observed after retirement is apparently due to the cessation of work-related expenses and the substitution of home production for market-purchase goods and services.

Groneck*et al.* (2011) developed a theoretical model based on a standard LCM adjusted with a Bayesian learning model under ambiguity on life expectation using HRS data. The motivation of the study was to shed light into the so-called "old-age dissaving puzzle" which is related to the fact that households still hold a large fractions of assets at very old age.

Apart from intended bequests, the authors argue that the reason for this might be the existence of bounded rationality and bounded self-control. The former refers to the inability of individuals to have perfect foresight as the problems to solve may be too complicated. The latter highlights the possibility of individuals having the right intentions or beliefs but not having the willpower to carry out the appropriate changes in behaviour.

They specify three non-rational behaviours that clearly contradicts standard rational behaviours, and therefore a standard rational agents approach would not address: (i) people seem to undersave for retirement when they are younger; (ii) people seem to hold large amounts of assets at the end of their life; and (iii) people do not have perfect foresight and they seem to revise their plans when older.

They found that, compared with a rational agent, an agent who has ambiguity over her subjective survival beliefs consumes more when younger, and therefore undersaves. On the other hand, consumption is lower at older ages. The consequence of this behaviour is that asset accumulation is much lower than in the case of a rational agent. They concluded that the model developed is can be helpful when explaining the observed "old-age dissaving puzzle".

Using the first wave of the SHARE database, Post and Hanewald (2011) examined whether individuals are aware of longevity risk, and if so, whether this awareness affects their actual saving behaviour. For this matter, they subdivide the sample into groups of individuals who can be expected to have homogeneous survival rate. Next, they regress the dispersion of individuals' subjective survival estimates on the dispersion observed in actual survival rates together with control variables. Given that they obtained a positive significance, it was concluded that respondents are aware of longevity risk.

Next, they addressed the question whether respondent adjust their saving behaviour in response to their awareness of risk. They found that the observed awareness of risk translate into an increased dispersion of saving outcomes. However, they did not find an increase on precautionary savings.

Another interesting aspect to be considered when analysing saving decisions is the role of bequests. People may decide to save more and/or accumulate more wealth because they plan to leave bequests to their relatives. Unfortunately, not many papers have examined the role of bequests on late life decisions using national databases.

Hurd and Smith (2001) used two waves from the HRS and the AHEAD survey to study the role of inheritances and bequests in shaping household decisions on wealth accumulation. In particular, they explored bequests by comparing actual bequests made by individuals upon their deaths, with their previously stated bequest intentions. This is done with the intention of examining the reasons individuals might revise their bequest expectations. The authors found that the reasons for between-wave revisions of the subjective bequest probabilities include changes in SSPs, household wealth, out-of-pocket medical expenses and widowing.

Following a similar approach Gan *et al.* (2004) investigated whether subjective expectations about future mortality affect consumption and bequests motives. They used data from the second, third and fourth waves of the AHEAD sample. It is argued that the first wave of AHEAD is excluded because good evidence that it underreported asset holdings. They found out that bequest motives are small on average pointing towards the idea that most bequests are involuntary or accidental. However, the effects of planned bequests are not negligible.

Both Hurd and Smith (2001) and Gan *et al.* (2004) coincide in the fact that LCM should also take into account bequests to explain late-life decisions such as retirement.

2.1.2. LCM and consumption

Consumption behaviour and its relationship with the LCM has also been analysed previously in the economic literature. Below we examine some of the most relevant contributions. It is generally concluded that economic agents adjust their consumption pattern with their survival chances as the LCM would suggest.

Salm (2006) investigated how subjective mortality expectations and heterogeneity in time and risk preferences affect the consumption and saving behaviour of elderly. For this matter he made use of the fifth and sixth wave of the HRS.

The dependent variable is defined as the real growth rate between waves of consumption. Consumption is measured as the sum of annual expenditures on non-durable goods, which include food, gas, clothing, dining out, vacations, tickets to events and hobbies. In the case of the explanatory variables, subjective mortality risk is calculated using annual life table information.

Salm (2006) stated that determining whether consumption growth decreases with higher mortality expectations is an alternative and more direct test of the LCM. In this respect, He found that consumption and saving choices vary with subjective mortality rates as suggested by the LCM. In particular, he estimated that consumption expenditure on non-durable goods decreases as a result of an increase in subjective mortality rates.

The study also analysed the impact of individuals' risk aversion and discount rates on consumption growth. The idea behind is that more risk averse agents are less willing to accept consumption volatility and that individuals with higher discount factor would allocate more fund to present consumption. The author concluded that different answers in survey questions about time and risk preferences reflect differences in actual saving and consumption behaviour.

Perry (2005) analysed the relationship between SSPs and consumption growth. He argued that according the LCM, a lower survival expectations should lead to more positively sloped consumption trend. In order words, those who think they are more likely to survive will have less consumption growth over time, and those who think they will die sooner won't need to save much and will consume more today.

In order to perform the analysis Perry (2005) uses the first 12 waves of the HRS. Given that the HRS does not provide full information on consumption, he estimated it based on strong assumption. Moreover, he used SSP and life table information to estimate mortality risk in the next year for all the respondents. The estimation approach consisted on an OLS approach.

A very interesting aspect of the study is that it allowed for the possibility of people not understanding the SSP question, or alternatively, people not being able to behave rationally as the LCM would suggest due to low-cognitive ability. For this reason the author presented sensitivity analysis based on a restricted sample of cognitive-able individuals, selected based on HRS questions.

The results presented by Perry (2005) showed that there is weak evidence that subjective survival chances have an effect on consumption growth. He obtained the correct sign but non-significance. This result is also observed for the high-cognitive subsample, but the effect of SSP, even though still not significant, was bigger. Finally, after the disappointing results, he point out towards the possible measurement error contained in the consumption data.

Biró (2011) used the two waves of the SHARE database to analyse the adjustment of consumption expenditures of elderly people if the subjective longevity changes. She explained that in consonance with the LCM, the effect of mortality hazard on the expected consumption growth is negative: those who have higher hazard plan lower consumption level in the future, and consume more in the present.

An interesting feature of this study is that it focused not on SSPs directly, but on subjective mortality hazards derived from SSPs in the SHARE database. As reported by Biró (2011), using the level or the change of SSP in the empirical models could lead to unreliable results. This is because of the survey design as the difference between the target age and the current age of the respondents varies across ages. The reported survival probability can change not only if the subjective life expectancy changes, but also if the target age in the probability question changes. For that reason, the author estimated mortality hazard rates using a similar approach as Gan *et al.* (2005).

The author also argued that mortality hazard is likely to be endogenous in the consumption model due to the presence of measurement error and unobserved variables. Therefore, mortality hazard is instrumented by the death of a sibling. It is stated that the instrumenting strategy hinges on the idea that the death of a sibling influences SSP, and such an event is not likely to have a direct effect on the consumption expenditures of the elderly people.

Based on a cross-sectional regression of the first differenced consumption on mortality hazard, she found that those who have positive wealth holdings adjust their consumption expenditures upwards if the subjective mortality hazard increases. In particular, at age 60 if the expected remaining lifetime decreases by 4 years then it is estimated that annual expenditure on food increases by 200-220 Euros per year.

2.2. SSP AND RETIREMENT

The literature has extensively analysed the role of SSP in the LCM using micro datasets. In particular, many authors have studied the role of expected longevity on retirement decisions of individuals, especially using the HRS database. The most prominent pieces of work are presented below.

It is commonly accepted that this economic literature builds on the work of Wolfe (1983). Using data from the American Social Security's work history sample, this author demonstrated that people that anticipate will live longer retire later using three different models: an actuarial, an economic, and a statistical model. The actuarial model compared the net present value of the retirement income for an early retiree and for a normal retiree. Based on assumptions about the annual pension, interest rate and penalization for early retirement, Wolfe (1983) estimated that individuals who expect to live less than 79 year are better off opting for early retirement. On the other hand, those who expect to live more than this borderline age should delay retirement until age 65. The economic model considers the income-leisure choices that an economic ration agent would make. The author concluded that greater perceived longevity should be associated with greater labour supply between 62 and 65 years old. Finally, Wolfe (1983) used a statistical model using data from the US Social Security's Continuous Work History Sample (CWHS) to conclude that perceived longevity plays a role in the decision when to begin receipt of Social Security retirement benefits.

Waldron (2001) examined the relationship between retirement age and subsequent mortality for men with a retirement age of at least 62 who have lived to at least age 65. This study uses data from the 1973 American Current Population Survey matched to Social Security administrative records. The author uses a probit specification to estimate the impact of retirement age on mortality. The sample consists of 10,938 retired males and the dependent variables is defined as a binary variable that equals 1 if an individual dies in the observation period (1973-1997) and 0 if the individual is alive at the end of the observational period (1997). Additionally, several other explanatory variables are included as age, birth cohort, marital status, race, education, among others.

The results found suggest that men retiring early are likely to die sooner than men retiring later. In particular, retiring at age 62 increases the probability of dying by 32% in the next 25 years relative to those retiring at age 65. Furthermore, less-educated male workers retire early may face higher mortality risk than other groups of workers because of lower educational level is associated with higher mortality risk, but also because early retirement is associated with higher mortality risk, as well. It was also concluded that men retiring at age 62 have a higher mortality risk than retiring between ages 62 and 3 months and 62 and 11 months, ages 63, 64 and 65.

It should be mentioned that this studies does not address the reason why there is this negative correlation between retirement age and mortality. However, Waldron (2001) presents two alternative hypotheses as an explanation which would require future research. The "quasi-disability" hypothesis argues that the reason why there is a negative correlation is because it is likely that a subgroup of workers exist who choose to take retirement benefits at age 62 and have poor health, but not as bad as to qualify for disabled benefits.

The second hypothesis stresses the fact that self-perceived survival chances predict actual mortality, as argued in the second Chapter. Therefore, rational economic agents that expect to die earlier (and effectively, on average they do) will retire earlier, as suggested by the life-cycle hypothesis.

Hurd *et al.* (2004) studied the effect of SSPs on retirement and on the claiming of Social Security benefits. They claim that according the LCM, individuals who expect to live longer will retire later. Moreover, those who expect to be long-lived will view the increase in Social Security benefits that results from claiming benefits at 65 rather than at 62 as being financially advantageous. The authors estimated that delaying one year Social Security claiming would imply an increase of approximately 8% in benefits.

Hurd *et al.* (2004) used the first four waves of the HRS (1992, 1994, 1996, and 1998) to estimate the probability of retirement as a function of SSP, eligibility of pension, age, wealth, wage rates, socio-economics variables, and health variables using a probit specification. They presented two separate estimations for individuals aged 53-61 at wave t+1 and for those over 62. They argued that this is convenient given the likely differing effect of pension eligibility and Social Security claiming. The analysis of retirement is based on retirement hazards, where it is estimated, conditional on labour force participation in wave 1, 2 or 3 (t), what is the probability of been retired in the next wave (t+1)?. Additionally, the determinants of the probability of early claiming is estimated is also estimated using a probit specification for individuals aged under 62, and for those aged over 62, a bivariate probit is used to estimate the probability of retirement and claiming benefits.

A rather important point is the definition of labour force participants and retirees. The authors define individuals in the labour force if they report themselves as full-time or part-time workers or unemployed. On the other hand, retirees are identified if they were in the labour force and in the following wave they report themselves as retired, partially retired, disabled or not in the labour force.

The results showed by the Hurd *et al.* (2004) points out towards an interesting conclusion. They argued that the effect of SSP on the retirement probability before age of 62 is negligible. However, after age 62 workers with very low survival probability do leave the labour force earlier than those with higher survival expectations. They also claimed that health indicators, particularly among the youngest, have large effects.

Regarding Social Security claiming, their analysis concluded that individuals with very low subjective survival chances do claim earlier, but the majority of workers claim Social Security benefits as soon as they are eligible.

One of the most important puzzles of the LCM is why retirement has increased among elders, even though mortality has decreased in the last century. One would expect to observed decreasing retirement rates. Kalemli-Ozcan and Weil (2010) has shed light on this issue. They constructed a model in which agents make labour/leisure choices over their lifetimes subject to uncertainty about their date of death.

They argued that two different effects on retirement exist as a consequence of changes in mortality. The "uncertain effect" by which the optimal plan of individuals would be to work until they die. This effect can be present in high mortality environments where individuals who saved up for retirement would face a high risk of dying before they could enjoy their planned leisure, so they continue working. On the other hand, they presented the "horizon effect". This effect point towards the hypothesis that rising life expectancy would lead to later retirement, as the LCM suggests. The idea is that as mortality falls, individuals will find it optimal to plan and save for retirement.

The authors argued that, starting from high levels of mortality, the "uncertain effect" could outweigh the "horizon effect". However, the effect of falling mortality on labour supply complements several other effect of falling mortality. Kalemli-Ozcan and Weil (2010) highlighted three of them. The growth of public pension programs such as Social Security could incentive workers to retire. Furthermore, increases in wages allow people to save more money during working years to be able to finance retirement years. And finally, they argued that changes in the production technology could make older workers less able to perform certain tasks compared to younger workers, and then the labour force participation of older people should decrease.

Similarly to Hurd *et al.* (2004), Delavande *et al.* (2006) analysed the extent to which an individual's survival expectations influence the decision of retirement and Social Security claiming. They claimed that the effects estimated by Hurd *et al.* (2004) might be too small because of measurement error in SSP which causes inconsistent coefficients.

Interestingly, they presented four different estimation strategies. In the first case, the independent variable is SSP (as in the case of Hurd *et al.*, 2004). Secondly, they used instrumented subjective survival values as independent variable. Thirdly, they used instrumented objective survival values based on life tables. And finally, they used both the instrumented subjective survival values and the difference between the instrumented subjective and objective survival probabilities.

Following the same estimation approach as Hurd *et al.* (2004), Delavande *et al.* (2006) found that, when SSPs are instrumented in order to correct measurement error, among people who are still working at age 62, those who expect to live longer are statistically more likely to delay claiming. In particular, they obtained that an increase of 5% in the predicted survival probability to age 75 leads to a 1.9% decrease in the proportion of individuals still working at age 62 who claim benefits before age 64. They also found no significant effect of expected longevity on retirement behaviour.

It was concluded that the effects of subjective probability beliefs on behaviour are almost completely obscure unless measurement error is taken into account.

Bloom *et al.* (2007) explored the proposition that expected longevity affects retirement decisions and accumulated wealth. They argued that according to the LCM, individuals who expect to live longer will retire later and accumulate more wealth in order to finance more years of retirement income. This issue is examined using data drawn from the first wave of the HRS with a subsample of individuals aged between 51 and 61 in 1992.

One of the main features of the study is that the authors instrumented SSP due to possible measurement error and also to address the potential endogeneity of SSP that could arise if SSPs are affected by wealth and retirement decisions. They used parental mortality experience (current age, or age at death of parents) as an instrument.

When analysing the effect of SSP on retirement decisions, they defined individuals using the same definition as Hurd *et al.* (2004). They defined as in the labour force if individuals are full-time or part-time workers or unemployed. Individuals not in the labour force (retired, partially retired, disabled, or not in the labour force) are defined as retired.

Additionally, Bloom *et al.* (2007) used the reported SSPs to age 75 as measure of expected longevity. They claimed that using SSPs to age 85 instead, do not modify the results.

The analysis is performed for singles and couples and for men and women. In addition to SSPs, it was included additional explanatory variables such as age, race, education, number of children, a variable reflecting the financial planning horizon, self-reported health status, a dummy variable reflecting health problems that might limit working, and income. Furthermore, for couples, SSP of the spouse was also included in order to test for joint-decision making.

The authors found that there was no evidence that higher SSP decreased the probability of retirement regardless of sex, couple status, control variables and estimation method (probit or IV probit). They argued that the reason for this result could be that retirement decisions are driven by institutional constraints and incentives.

In the case of the effect of SSP on wealth, a similar approach is followed and they obtained that among singles there is no evidence that higher SSPs increase the wealth of the respondents. However, for couples, they found a significant effect suggesting that a 10% increase in the husband's (wife's) SSP would result in approximately 27,600 USD (32,700 USD) increase in household wealth.

This very interesting study has one important drawback. The data used is cross-sectional which creates serious problems to the estimation. Ideally, the one would want to estimate the effect of subjective survival chances on retirement *when* the decision of retirement is taken. Following Bloom *et al.* (2007)'s approach we observe a retired individual and her SSP at a given period of time. However, the individual could have retired many years before. Therefore the results should be taken with caution.

Bloom *et al.* (2006) replicated the analysis performed in the US using HRS data, but using SHARE and ELSA data. When estimating the effect of SSPs on the probability of retirement and on wealth accumulation, the only difference with the previous study is that the variable reflecting the financial planning horizon and income are omitted.

In the case of SHARE, Bloom *et al.* (2006) found no significant evidence that higher SSPs decrease the probability of retirement of increase wealth accumulation regardless of sex, couple status, control variables and estimation method. For the UK, ELSA database analysis showed no evidence that SSPs affect retirement decision. However, they did find evidence of a significant effect on wealth accumulation among singles and couples, as suggested by the LCM.

Unfortunately, this study suffers from the same important drawback than the previous one.

Fischer and Sousa-Poza (2006) investigated the determinants of early retirement in Europe using the first wave of the SHARE database. They argued that understanding early retirement and its motivations could assist the formulation of policies that might encourage the return of younger retirees to active employment. This is especially important nowadays given that population is ageing and fertility rates are lower than 50 years ago.

The authors stated that this study contributes to this field of research by (i) jointly testing individual and institutional factors of early retirement; and (ii) using microdata from 10 countries. It is precisely these two features that are a novelty as it provides ample variation in institutional settings and macroeconomic conditions.

Fischer and Sousa-Poza (2006) used a probit specification in order to study the determinants of early retirement. In particular, the probability of early retirement is thought to be a function of personal and job characteristics and macroeconomic and institutional factors. In the case of personal and job characteristics, they included gender, marital status, education, tenure, firm size, firm industry and supervision of employees. On the other hand, real average replacement rate (the ratio between last earnings and first pension), pension wealth accrual (penalization for early retirement), the OECD indicator of employment protection legislation, the level of unemployment and real GDP per capita were included as macroeconomic and institutional variables.

It was concluded that pension systems' generosity is a very important factor influencing early departure from the labour market. They also found that the pension wealth accrual rate and average replacement rate also affect retirement decisions. Moreover, higher national wealth is associated with a lower probability of early retirement, while a growing unemployment rate triggers more early retirement.

On an individual level, it was obtained that married female workers and those with a better education are more likely to retire early. Surprisingly, the probability of early retirement is lower for people close to the statutory pension age. Furthermore, job characteristics, firm size, and industry also influence retirement decisions.

O`Donnell *et al.* (2008) presented a very interesting study where they examined whether expectations of survival influence the timing of retirement. They made use of the first three waves of the ELSA database.

The authors model the probability of retirement between consecutive waves as a function of survival expectations, and other covariates reported in the first of the waves, such as wealth, health, socio-economic variables, earnings, and pension arrangements. Taking the reported value of the first of the waves is rather important according to O`Donnell *et al.* (2008) as this avoids bias from reverse causality (the possibility that retirement itself causes a revision of survival expectations). The empirical strategy consists on the estimation of a pooled probit.

According to the authors, retirement cannot be unambiguously identified as, for example, some individuals may report themselves as retired but only work a few hours per week but they still receive retirement pension. In order to tackle this problem they took a conservative approach and model retirement as going from *working* (if an individual reports as employed or self-employed) status to *inactive* (if an individual is fully or partially retired, unemployed, permanently sick or disabled). Additionally, they test an alternative retirement transition, which is going from *working* status to *retired* (partially or fully) only.

An interesting feature of this paper is that it is argued that using SSP might not be the correct measure of survival expectations to be used in a model of retirement because, even though it may be taken as a proxy for life expectancy, the interpretation is impeded by the fact that the information provided by SSP about survival expectations is contingent on the respondent's proximity to the target age. It is suggested that the relevant parameter should be the extent to which an individual's expectations are above or below the actuarial averages, what they called individual mortality factor. In any case, the paper also tests for the inclusion of standard SSP instead in the model.

O`Donnell *et al.* (2008) also argued that SSP may contain a considerable degree of measurement error due to unfamiliarity of responses and the tendency to give focal point answers. They argued, based on previous studies (Delavande*et al.*, 2006; Bloom *et al.*, 2006; and Bloom *et al.*, 2007) that parental mortality and smoking behaviour can be used as valid instruments.

They found that there is a significant concave relationship between mortality expectations and retirement. Individuals that are pessimistic about their survival are least likely to retire. However, as expectations improve, the retirement probability first rises, but then falls over most of the distribution of survival expectations. These results are robust to the inclusion of SSP as oppose to mortality factor for males. But for females, only a linear significant relationship is found. They also obtained that for males, there is no evidence that SSP are endogenous to retirement, but for females this might not be true.

Cocco and Gomes (2012) studied the optimal consumption and saving choices in a LCM. They developed a theoretical model where individuals receive a stochastic labour income each period and decides how much to consume and save. Individuals form their own current survival probabilities but are unaware of future trends in mortality when taking the decision of consumption and saving, but form an expectation. Interestingly the model allows for endogenous retirement, so that individuals choose when to retire.

However, the main feature of the theoretical model is that individuals can invest in two financial assets. The first is a riskless asset and the second it what the authors called longevity bonds, a financial asset whose return is correlated to the shocks on survival probabilities, thus providing the investor with a perfect hedge against longevity risk.

Even though the model developed is not tested with micro data as it is theoretical, some very interesting features are highlighted. Cocco and Gomes (2012) are able to replicate the life-cycle profiles and found that early in the life households are liquidity constrained and consumption follows income very closely, but agents accumulate wealth. As agents approach retirement labour income decreases and consumption exceeds income, so that agents start consuming out their accumulated wealth.

More importantly they obtained that agents do react to shocks to life-expectancy. In particular, agents save more and retire later as a response to improvements in life expectancy. The authors estimated that, according the their calibration of the model if life expectancy at age 65 is 75 years old, agents would save 36.4% of their income and none of them would work at age 65. However, if life expectancy at age 65 increases to 85 years old, agents would save 45.7% of their income and 93% of them would work at age 65.

The study then analyse the effect of a longevity bond. They argued that even though agents react to improvements in longevity saving more and retiring later, they would benefit further from investing in this financial asset, and the benefits can be economically very significant.

Overall, we have found that the economic literature does consider the links between economic decisions such as saving, consumption and retirement, and longevity expectations in the same way the LCM suggests.

3. DATA

As in the previous two Chapters, we have taken advantage of SHARE database. Specifically, we make use of the first (and only) two waves of SHARE available at the moment. For the analysis undertaken we have considered individuals aged 45 up to 65 years old. After eliminating missing values and non-respondent items, the final sample used in the analysis consists of 15,058 respondents for the first wave and 16,485 for the second wave. We have found 8,448 individuals that have participated in both waves of the SHARE survey.

Given that SSPs have been extensively examined in the first two Chapters, in this Chapter the focus is on the reported labour situation of the respondents. The Employment and Pension module of SHARE provides substantial information on the labour situation of respondents, labour income, job characteristics, together with retirement data, such as date of retirement, reasons for retirement, and retirement pensions.

Individuals are asked about their current labour status and they are given 5 exclusive options: (i) retired; (ii) employed or self-employed (including working for family business); (iii) unemployed; (iv) permanently sick or disabled; and (v) homemaker. Additionally, for individuals that don't report themselves are employed or self-employed, they are asked if they have done any paid work in the last four weeks and if they are temporarily away from work[56]. Furthermore, if individuals are reported as permanently sick or homemaker, they are asked if they have ever done a paid work. As described below, we use these questions to classify individuals according to their true labour status.

The Employment and Pension module also provides information on the total number of hours worked per week by respondents, regardless of their reported labour status. This information will also be interesting when testing the role of SSPs on labour supply decisions. It might be the case that Europeans choose alternative retirement pathways such as decreasing the number of hours worked gradually, instead of a one off retirement decision.

Below, we present the different definitions of retirement that we consider in the analysis. Additionally, we present a descriptive analysis of the labour status of SHARE respondents.

3.1. RETIREMENT DEFINITION

In this Chapter we are interested in analysing the role of SSPs on retirement decisions. It is therefore crucial to understand what we understand as retirement and to classify respondents accordingly. However, the classification of SHARE respondents into labour status is not a trivial exercise. It calls for attention the fact that there exists old Europeans that self-report themselves as retired, homemaker or disabled but are actually working, or individuals that report themselves as homemakers or disabled but are in fact retired as they are receiving a retirement pension, for example.

Börsch-Supan *et al.* (2005) stresses this issue when saying that in some countries individuals may be allowed to work while collecting pension benefits and report themselves retired even if working. Also an important retirement pathway is through

[56] Individuals are only asked if they are temporarily away from work in the first wave. In the second wave of SHARE that question was omitted.

disable benefits. For that reason we have developed three definitions of retirement that we present below.

3.1.1. In the labour force and not in the labour force

Based on the definition of retirement previously adopted in the literature when analysing the effect of expected longevity on labour supply decisions (Bloom *et al.*, 2006; Hurd *et al.*, 2002; Bloom *et al.*, 2007; and O'Donnell *et al.*, 2008) we classify individuals based on their self-reported labour status. An individual is defined as "not in the labour force" if she reports herself as retiree, disabled or homemaker. On the other hand, an individual is defines as "in the labour force" is she reports herself as employee or unemployed. Clearly, this broad definition is imprecise as we are more interested in the transition from being in the labour force to retirement only, but we have included it for comparison reasons.

3.1.2. In the labour force and retired

It is questionable why we should include disabled and homemakers in our analysis, as we are interested in the transition from being in the labour force to retirement. For that reason, we include an additional definition to be used in the analysis. We define someone as "retired" only if she reports herself as retired. On the other hand, an individual is defines as "in the labour force" if she reports herself as employee or unemployed. Note that the definition of "in the labour force" is the same in the first definition.

This definition has also been used in the literature. O'Donnell *et al.* (2008) and Fischer and Sousa-Poza (2006) analysed labour supply decisions based on this definition of retirement.

3.1.3. Truly in the labour force and truly retired

However, we can go one step further and take into account the fact that some individuals might report themselves as retired, for example, but continue to work. Or, individuals that report themselves as homemaker, but are really retired.

We therefore, make use of further information drawn from the Employment and Pension module of SHARE. In particular, we use the responses to the question regarding if individuals have done any paid work in the last four weeks, if they are temporarily away from work, and if they have ever done a paid work in the past.

In this definition individuals are defined as really "truly retired" if they report themselves as being retired but, additionally, they have not worked in the last month and are not temporarily away from work. We also include disabled individuals and homemakers that have worked in the past but not in the last month. On the other hand, individuals are truly in the labour force if they report themselves as employees or unemployed. But additionally, we include individuals that report themselves as being retired, disabled and homemakers, but have worked in the last month or they are simply temporarily away from work. We believe that this definition is more precise and we use this definition as our base scenario. Surprisingly, to our knowledge, no other study on SSPs uses this definition.

3.2. DESCRIPTIVE ANALYSIS

In this subsection we present a descriptive analysis of the labour status of SHARE respondents aged 45 to 65 years old. Given that we have found significant differences between males and females, when appropriate we provide a differentiated analysis[57].

As mentioned above, SHARE respondents are asked to assess their current labour status into one of these five exclusive options: (i) retired; (ii) employed or self-employed (including working for family business); (iii) unemployed; (iv) permanently sick or disabled; and (v) homemaker. Results for the first wave of SHARE are shown in Figure 1. Panel (a) presents the results for males and panel (b) for females. Results for the second wave are quite similar.

It can be observed outstanding differences between both patterns. More males report themselves as "employed" than females do, especially before age 65. Moreover, more males report themselves as "retired" than females do, especially after 58 years old. Also

[57] We have also considered differences by marital status. However, we have not found significant differences between married and non-married individuals.

remarkable is the fact that the prevalence of "homemakers" is much higher in females than in males. Börsch-Supan *et al.* (2005) remarked that it is highly probable that homemaker women have not had labour market experience during their lifetime.

SHARE respondents are also asked at what age old Europeans decided to stop working and retire. Results are shown in Figure 2. For females, it is observed three clear peaks at ages 56 (6.2%), 60 (9.1%) and at age 65 (5.1%). In the case of males, the pattern is different. Two differentiated peaks are observed at age 60 (11.2%), and 65 (7.5%).

This behaviour contrasts with the observed for the US where there exists clearly two peaks at ages 62 and 65 years old as shown by Rust and Phelan (1997). The reason for this could lie in the fact that in the US individuals could officially retire at 62 years old and enjoy an early retired pension.

Next, we explore the labour supply pattern for SHARE respondents. We use the three definitions presented above and explore their differences. Figure 3 presents, by age, the percentage of individuals classified as in the labour force and not in labour force (panel (a)), in the labour force and retired (panel (b)), and truly in the labour force and truly retired (panel (c)). As expected the percentage of individuals participating in the labour force decreases with age, while the percentage of respondents leaving the labour force increases with age.

The main observed difference is the pattern of "not in the labour force" and that of "retired". The latter is smaller for every age. The reason for this is that individuals reported as disabled and homemakers are included in the first definition, while they are not in the second one. We do not observe significant differences between the different definitions of labour force participation.

Figure 4 presents the same results but we now differentiate by sex of the respondent. As it can be observed, the results are grouped by labour status. It is surprising that for males, the patterns of the different definitions of not participating in the labour force ("not in the labour force", "retired" and "truly retired") as rather similar. This is due to the fact that the prevalence of homemakers is quite limited in the male segment.

As mentioned above, we are also interested in analysing the pattern of the total number of working hours of SHARE respondents. We believe that there might be an alternative retirement pathways such as decreasing the number of hours worked gradually, instead of a one off retirement decision. Figure 5shows the average total number of hours worked per week by SHARE respondents. As expected, we observe a downwards trend as individuals age. The fact that we do not notice any discontinuity supports the idea that respondents may opt for partial retirement. It is also interesting to highlight that, on average, males work more hours per week than females.

3.3. LABOUR SUPPLY TRANSITION

As mentioned in the Introduction, our focus in this Chapter is on the role of SSPs on retirement decisions. The longitudinal aspect of SHARE provides a great opportunity to study these decisions. In particular, we are particularly interested in the transition of individuals in the labour force into retirement for individuals aged 45 to 65 years old.

Table 1 presents the observed transition from labour force into retirement, conditional on the different definitions presented above. Moreover, we present the results for the whole sample and differentiated by sex of the respondent. We find that 19.5%, 14.1% and 9.2% of the respondents in our restricted sample have exited the labour force, depending on the definition of retirement we use. It is not surprising that we find less labour force transition if we use the more restrictive definition of retirement. Overall, we find that more males exit the labour force between wave 1 and wave 2, regardless of the retirement definition used.

Table 2 presents the transition in the number of hours worked, from wave 1 to wave 2. We find that more than half of the respondents in our restricted sample have decreased the hours worked between waves. From those, the majority decreased the number of hours by less than 20 hours per week. We also find that around a fifth of the respondents actually increase the number of hours worked.

4. ECONOMETRIC MODELLING

The main purpose of this Chapter is to determine the role played by SSPs on labour supply decisions of Europeans aged between 45 and 65 years old using SHARE data. We would expect that individuals who expect to die earlier will retire sooner than individuals who expect to live longer, as suggested by several authors.

For this matter, we propose a probit specification in order to model the probability of an individual being retired in wave 2 (who was in the labour force in wave 1) as a function of SSP and a set of independent variables[58].

$$\Pr(Retired_i = 1 \mid SSP_i,\, X_i) = \Phi(\alpha + \beta_1 SSP_i + \beta_2 X_i) \qquad (1)$$

where respondents are indexed i, Φ is the Cumulative Distribution Function of the standard normal distribution, SSP_i are the individual reported expected longevity to age 75, and X_i are covariates. This approach has been used by several authors such as Hurd *et al.* (2004), Delavande *et al.* (2006), Bloom *et al.* (2006, 2007), and O`Donnell *et al.* (2008).

As covariates we have included several variables based on the evidence presented in the academic literature. It should be mentioned that these covariates are based on those used in Chapter 1 and Chapter 2. Therefore, further details on the specification can be found there[59].

In particular, we include standard socio-demographic variables as age, and age squared[60], marital status (dummy variable equals to 1 if married), country of residence (country

[58] We have used a probit approach, instead of a logit one, because it is much more common in the literature we have reviewed. However, results do not change significantly when a logit approach is used.

[59] Table 5 in Chapter 1 presents a brief description of the variables used in the econometric modelling.

[60] Age squared is divided by 100.

dummies), educational level (defined using the seven categories of the ISCED international codes[61]), and number of children of the respondent.

Additionally, we have included two important economic variables. Wealth and income level of the respondent[62]. This variables were also used by Hurd *et al.* (2004), Delavande *et al.* (2006), Bloom *et al.* (2006, 2007), and O`Donnell *et al.* (2008). *A priori*, one would expect that, other things being equal, wealthier individuals would retire earlier. The expected effect of income on the probability of retirement seems to be ambiguous as on the one hand, higher income levels could encourage continuing working to accumulate wealth more rapidly. However, on the other hand, higher income could be related with more wealth already accumulated, and therefore, having the effect of encouraging retirement.

We have also included health related variables in order to model the probability of entering into retirement. This was also used by Hurd *et al.* (2004), Delavande *et al.* (2006), Bloom *et al.* (2006, 2007), and O`Donnell *et al.* (2008). Concretely, we have used the presence of long-term illness[63], two measures of the number of physical limitations: ADL and IADL, the number of severe and non-severe illnesses currently suffering, and finally, self-reported health status.

Lastly, we have included the reported probability of receiving an inheritance based on the evidence presented by Hurd and Smith (2001) and Gan *et al.* (2004). *A priori*, it could be expected that the higher the probability of receiving an inheritance would have the effect of increasing the probability of entering into retirement, because it would decrease the need to accumulate wealth through labour income[64].

---

[61] ISCED CODE: 0 (no education), 1 (primary education or first stage of basic education), 2 (lower secondary or second stage of basic education), 3 (upper secondary education), 4 (post-secondary non-tertiary education), 5 (first stage of tertiary education), and 6 (second stage of tertiary education).

[62] Wealth is divided by 1,000,000 and income is divided by 1,000. Both variables are expressed in 2005 Purchasing Power Parity basis.

[63] Dummy variable 1 if the respondent is suffering from a long-term illness, 0 otherwise.

[64] Ideally, we also wanted to include as a covariate the reported probability of *leaving* an inheritance. However, we found a significant proportion of missing values in our sample. Therefore, we decided not to include this variable.

Table 3 presents a correlation matrix for all of the variables used in the estimation. The lower triangle corresponds to the correlation between variables for males only, while the upper triangle is for females. For example, the correlation between SSP and educational level for males is 0.02, it is 0.10 for females[65]. Interestingly, we find that the correlation between retirement in the second wave and SSP is negative for females but positive for males. We also find that the correlation between ADL and IADL is stronger for males, while the correlation between IADL and self-perceived health status is stronger for females. Finally, it is also observed that the correlation between age and children is more negative for females.

When modelling the probability of entering into retirement, we have conducted separate analysis by sex following the approach used by Bloom *et al.* (2006, 2007), and O`Donnell *et al.* (2008). This is done because the results presented by these authors tend to differ by sex. Other authors conduct a joint analysis and only include a dummy variable to control for the sex of the respondent.

A rather important point that would determine the empirical strategy is the temporal aspect of the estimation. We model the probability that an individual is retired in wave t, conditional on the fact that this individual was in the labour force in wave t-1, as a function of SSP and other covariates. However, the academic literature is not clear on whether SSP and the other covariates should be measured in t or in t-1. Authors such as Hurd *et al.* (2004), Delavande *et al.* (2006) use their values in t. However, O`Donnell *et al.* (2008) presents a convincing argument in favour of using their values in t-1. They state that there is a possibility that retirement itself could have an impact on expected longevity and the other covariates (reverse causality). This could result in an inconsistent estimation of parameters. Therefore, using their values is t-1 would rule out this possibility. Salm (2006) also argues towards this direction. We adopt O`Donnell's approach as we believe it is more suitable.

Equally important is the treatment of SSP when modelling entering into retirement. Most of the authors (Stern, 1989; Delavande *et al.*, 2006; Bloom *et al.*, 2006, 2007; and

[65] The correlation matrix presented in Table 3 corresponds to the subsample of individuals "truly in the labour force" in the first wave. Results do not change significantly if the subsample includes only individuals "in the labour force" in the first wave.

O`Donnell *et al.*, 2008; among others) argued that SSP could suffer from endogeneity. This endogeneity could arise because of two factors: SSP might suffer from classical measurement error, and due to, what in the economic literature is known as "justification bias". This justification bias has also been studied previously. When investigating the effect of disability on labour force participation Stern (1989) claimed that non-working individuals might say they have a severe health problem because they have psychological or economic incentives to do so, for example, they might feel ashamed of not working or afraid of losing some benefits. Similarly, it cannot be rules out a situation where an individual reports very low expected longevity to justify being retired. It should be mentioned that, as mentioned above, following O`Donnell *et al.* (2008)'s approach immediately rules out the possibility of endogeneity caused be reverse causality.

In order to correct this possible endogeneity of SSP, an instrumental variable approach is needed. Delavande *et al.* (2006) and Bloom *et al.* (2006, 2007) instrument SSPs using information on parental mortality such as the age of dead[66]. O`Donnell *et al.* (2008) argues that parental mortality might not be an ideal instrument because it could have a direct impact on the work status of individuals through the care needs of elderly parents. Alternatively, this author proposed to use smoking behaviour as instrument. We believe it is appropriate to adopt a conservative approach and use both, parental mortality and smoking behaviour, as instruments to correct the possible endogeneity of SSP. Both instruments satisfy the relevance condition, as shown in Chapter 1 and 2, and the orthogonal condition.

Following the explanation provided by Rivers and Vuong (1988) and Arellano (2008), assuming a latent variable model ($Retired_i^*$), a probit model with endogenous explanatory variable could be characterized as follows:

$$Retired_i^* = \alpha + \beta_1 SSP_i + \beta_2 X_i + \mu_i \qquad (2)$$

$$SSP_i = \pi_0 + \pi_1 Z_i + v_i \qquad (3)$$

[66] Delavande *et al.* (2006) also corrects endogeneity using basic demographic characteristics, health variables and an optimism index.

$$Retired_i = 1[Retired_i^* > 0] \qquad (4)$$

where respondents are indexed i, SSP_i are the individual reported expected longevity to age 75, X_i are covariates that affect the retirement decisions, and Z_i are covariates that affect expected longevity. It is assumed that the error terms, μ_i and v_i,are jointly normally distributed with variances equals to one. This implies that:

$$\mu = \theta v + e$$

$$\theta = \frac{Cov(\mu,v)}{Var(v)} + e$$

$$E(e|v) = 0$$

$$Var(e) = 1 - \rho^2$$

where ρ is the correlation between μ and v, and e is the error term.

Therefore the Equation 2 can be written as:

$$Retired_i^* = \alpha + \beta_1 SSP_i + \beta_2 X_i + \theta v_i + e_i \qquad (5)$$

Consequently:

$$\Pr(Retired_i = 1 \mid SSP_i, X_i, Z_i, v_i) = \Phi\left(\frac{\alpha+\beta_1 SSP_i+\beta_2 X_i+\theta v_i}{\sqrt{1-\rho^2}}\right) \qquad (6)$$

In order to estimate this equation one need to know v. Rivers and Vuong (1988) proposed a two-step process. Firstly, in order to obtain a consistent estimate of v one should regress Equation 3 and obtain the residuals $\hat{v}$. Next, a standard probit is estimated. In our case the standard probit would consist of regressing $Retired_i$ on SSP_i, X_i, and $\hat{v}_i$.

5. RESULTS

In this section we present the results obtained from the estimation of the determinants of the probability of entering into retirement. We focus on our base case scenario which considers the narrowest definition of labour supply, what we have called *truly in the labour force and truly retired.* Furthermore, we present the results separated by sex and for three different model specifications: (i) probit model without the inclusion of health variables; (ii) probit model including also health variables as independent variables; and (iii) probit model corrected for the possible endogeneity of SSP (i.e. IV probit). Results are shown in Table 4[67].

When health variables are not included (second and third column) we found that the effect of SSP on the probability of entering into retirement is negative and significant for females. This result is coherent with the implications of the LCM, as individuals who expect to live exceptionally longer will retire at a later age than individuals who expect to die early. Surprisingly we obtained a non-significant positive marginal effect of SSP on the probability of entering into retirement for males.

Other studies have found comparable results using standard probit modelling. Hurd *et al.* (2004) found a significant negative effect only for those individuals with very low levels of SSP. Delavande *et al.* (2006) obtained a negative effect, as theory would suggest, but not significant. Contrary to our results, Bloom *et al.* (2007) found a negative and significant effect only for single males and a positive non-significant effect for single females. Bloom *et al.* (2006) obtained a significant negative effect for married females only for SHARE and negative but not significant effect when using ELSA respondents. Finally, O`Donnell *et al.* (2008) obtained negative and significant effects both for males and females.

With respect to age we don't find significant effects when we allow for non-linearity. However, when age squared is not included we obtain a positive and significant effect for age (not reported in Table 4) as it was expected and consistent with previous literature. It should also be noted that because we are using the narrowest definition of retirement,

[67] Country dummies are used in the estimation but not reported to save space.

younger individuals using alternative pathways into retirement, such as partial retirement or disabled claiming, are not considered in this analysis.

Our results also suggest that married and less educated males are more likely to retire. Previous studies point towards this direction. Hurd *et al.* (2004) obtained that married individuals aged 62 or more are more likely to retire. Bloom *et al.* (2006, 2007) concluded that less educated males and females are more likely to retire.

A very important determinant of retirement is wealth accumulated. As mentioned above, *a priori*, it is expected that individuals that have already accumulated higher amounts of wealth would retire sooner. Surprisingly, we find that wealthier males are less likely to retire. Delavande *et al.* (2006) also found similar results. On the other hand, O`Donnell *et al.* (2008) found a positive and significant effect for both males and females of wealth on the probability or retirement.

Finally, our results suggest that the probability that females attach to receiving inheritance has a positive and significant effect on the probability of retirement, as theory would suggest. However, for males we do not find significance.

When health variables are included in the estimation, the picture does not change substantially. We observe that the negative effect of SSP on the probability of retirement is still significant for females. However, the margin effect diminishes[68]. We also found that the presence of a long-term illness for males and a lower self-perceived health status for females increases the probability of entering into retirement.

The sixth and seventh column in Table 4 present the results for the probit model corrected for the possible endogeneity of SSP (i.e. IV probit). Under this specification, our results suggest that SSP has a negative effect (as theory suggests) on the probability of retirement for both males and females. However, we do not find significance. We also find that married and less educated and poorer males are more likely to retire. Furthermore, we find for females that the probability of receiving inheritance increases

[68] Marginal effect of SSP when health variables are not included: -0.0004.

Marginal effect of SSP when health variables are included: -0.0003.

the probability of retirement. The result of the Wald test of exogeneity indicates that one cannot reject the null hypothesis that the standard probit estimate is consistent.

Following Hurd *et al.* (2004) and O`Donnell *et al.* (2008) we have used an alternative specification of our restricted subsample. On the one hand, we have only included individuals aged 52 or more, as oppose to 45 or more. This is justified by the fact that the retirement rate for individuals aged under 52 is very low as shown in Figure 2. Additionally, these individuals might see retirement and death too far away in time. On the other hand, we have excluded observations reporting SSP equals to zero.

Results are shown in Table 5. For females we still obtain a negative and significant effect in the case of the standard probit specification. For males we now obtain a non-significant negative effect when health variables are not included. Similarly to the result obtained with the whole restricted sample, the Wald test of exogeneity indicates that one cannot reject the null hypothesis that the standard probit estimate is consistent.

Overall, our results point towards the finding that for females expectations of longevity are taken into account when taking the decision of retiring, as the LCM would suggest. Those females expecting to live longer are indeed more likely to delay retirement. This suggests that females tend to act as rational economic agents when taking late life decisions such as continue working or not. For males we do find that the sign of the effect is the expected, however we do not find significance. The rationale for this interesting result is discussed in Section 7 after we present a sensitivity analysis in the next section. Finally, we find that education, marital status and wealth affect the retirement decision for males, and the probability of receiving inheritance for females.

6. SENSITIVITY ANALYSIS

In this section we present the results from our sensitivity analysis. As described in Section 3, we have used different definitions of retirement in the analysis. The results presented in Section 5 are based on our base scenario definition of retirement (truly in the labour force and truly). In this definition individuals are defined as retired if they report themselves as being retired but, additionally, they have not worked in the last month and are not temporarily away from work. We also include disabled individuals and homemakers that have worked in the past but not in the last month. On the other hand, individuals in the labour force are those that report themselves as employees or unemployed, but additionally, we include individuals that report themselves as being retired, disabled and homemakers, but have worked in the last month or they are simply temporarily away from work.

In this section we present the results of the effect of SSPs on retirement decision based on the other two definitions of retirement. To start with, we use the first definition (in the labour force and not in the labour force). This definition considers that individuals are retired if they are not in the labour force. Therefore, this definition does not distinguish between retirees, disabled individuals and homemakers. Clearly, this definition is imprecise, but we have included it for comparison reasons, as several authors have also used it (Bloom *et al.*, 2006; Hurd *et al.*, 2002; Bloom *et al.*, 2007; and O'Donnell *et al.*, 2008).

We then present the results of the analysis using the second definition (in the labour force and retired), where individuals are defined retired only if they report themselves as being retired. On the other hand, individuals are assumed to be in the labour force if they report themselves as employed or unemployed. O'Donnell *et al.* (2008), Hospido and Zamarro (2012) and Fischer and Sousa-Poza (2006) also used this definition when analysing labour supply decisions.

Table 6 presents the results for the estimation using the first labour supply definition (in the labour force and not in the labour force). We find that, when health variables are not included, there exists a negative and significant effect of SSP on the probability of entering into retirement. This is true for both males and females. However, when health variables are included, the negative effect of SSP diminishes and it becomes not significant. Similarly, when the possible endogeneity of SSP is addressed, we find that SSP has non-significant negative effect on the probability of entering into retirement.

We have also performed the same analysis to the restricted subsample (individuals aged 52 or more and not reporting SSP equals to zero) as in the case of our base scenario. The results are shown in Table 7. Likewise, we find a significant negative effect for males and females that diminish when health variables are included. Interestingly, when the possible endogeneity of SSP is addressed, we find a negative significant effect for females and the Wald test of exogeneity indicates that one can reject the null hypothesis that the standard probit estimate is consistent.

Table 8 presents the results for the estimation using the second labour supply definition (in the labour force and retired). Results are somewhat different than those of the other labour supply definitions. In particular we do not find any significant effect of SSP on the probability of retirement, neither for males nor for females, irrespectively of the estimation strategy. Alternatively, Table 9 shows the results from the estimation that included the restricted subsample (individuals aged 52 or more and not reporting SSP equals to zero). In this case we find the correct sign of the effect (negative) for males and females including and excluding health variables, though the effects are not significant. In both tables the Wald test of exogeneity suggests that one cannot reject the null hypothesis that the standard probit estimate is consistent.

We believe that the results using the second labour supply definition (in the labour force and retired) are somewhat different because of two reasons. Firstly, individuals defined as retired only include those who report themselves are retired which is too restrictive. As suggested by Börsch-Supan *et al.* (2005) in response to labour market rigidities, many European countries have allowed and sometimes encouraged various forms of early exit from the labour force. The main tools for doing so have been unemployment insurance, an extended access to sickness or disability benefits, or the development of specific pre-retirement schemes. Individuals opting for these forms of retirement pathways would not be considered in the definition. On the other hand the definition of "in the labour force" might not be totally appropriate. Börsch-Supan *et al.* (2005) stresses this issue when saying that in some countries individuals may be allowed to work while collecting pension benefits and report themselves retired even if working.

As already mentioned, it would also be interesting to test the role of SSPs on labour supply decisions from another angle. In particular, it might be the case that Europeans choose alternative retirement pathways such as decreasing the number of hours worked gradually, instead of taking a one-off retirement decision. To our knowledge, no other paper has considered this issue using national databases in the context of the LCM.

We estimate the effect of SSPs on the *change* of number of hours worked per week (wave 2 – wave 1) given that the number of hours worked per week in the first wave was positive using an OLS approach. One would think, *a priori*, that the effect should be positive (on average, an individual reporting lower expected longevity would decrease the number of hours worked).

In particular, we consider the following specification:

$$Change\ hours\ worked_i = \alpha + \beta_1 SSP_i + \beta_2 X_i + \varepsilon_i \qquad (7)$$

where respondents are indexed i, SSP_i are the individual reported expected longevity, X_i are the same covariates used in the analysis of the probability of entering into retirement, and ε_i is the error term.

Similar to the base case of modelling of the probability of retirement, it could be argued that SSP suffers from endogeneity due to classical measurement error and justification bias in all the sensitivity cases. We, therefore, use a two-stage OLS approach where we also instrument SSP using parental mortality experience and smoking behaviour as instruments.

More formally:

$$Change\ hours\ worked_i = \alpha + \beta_1 SSP_i + \beta_2 X_i + \varepsilon_i \qquad (8)$$

such that $Cov(SSP, \varepsilon) \neq 0$

We define a set of instruments Z that have an impact on SSP:

$$SSP_i = \pi_0 + \pi_1 Z_i + v_i \qquad (9)$$

where v_i is the error term and such that $Cov(Z, SSP) \neq 0$ and $Cov(Z, \varepsilon) = 0$

We therefore estimate:

$$Change\ hours\ worked_i = \alpha + \beta_1 \widehat{SSP}_i + \beta_2 X_i + \varepsilon_i \qquad (10)$$

such that $Cov(\widehat{SSP}, \varepsilon) = 0$ and where:

$$\widehat{SSP}_i = \hat{\pi}_0 + \hat{\pi}_1 Z_i \qquad (11)$$

The second and third columns of Table 10 present an OLS estimation excluding health variables. We find that a higher expected longevity is translated into an increase in the number of hours worked, though this effect is not significant both for males and females. Additionally, we find that older individuals reduce in the number of hours worked and more educated people tend to increase the number of hours worked.

Likewise, when health variables are included we find that for males and females SSP has a positive effect on number of hours supplied, but the effect is not significant. As expected, the margin effect of SSP on the dependent variable decreases once health variable. Furthermore, we found that the presence of a long-term illness and increasing number of limitations in ADL have a negative effect on the number of hours worked for females. For males, the number of severe illnesses has a similar effect.

When we control for the possible endogeneity of SSP using an instrumental variable approach we also find that a higher expected longevity is translated into an increase in the number of hours worked, though this effect is not significant both for males and females. The result for the test of endogeneity suggests that one can reject the null hypothesis that SSP is exogeneous in the regression.

Similar to the exercise done when analysing the role of SSP on the probability of entering into retirement, we restrict our subsample to include only individuals aged 52 or more. The reason for this is that it might be the case that younger individuals are less likely to modify their hours supplied because they perceived death as an event too far away in time. Table 11 present the results. We observe that in this case we do find a significant positive effect of SSP on the change in the number of hours worked for females. In particular, an increase in 25% in the expected longevity is associated with an increase of 1 hour of work per week.

As in the case of the estimation for the un-restricted subsample, when we control for the possible endogeneity of SSP, we find a non-significant positive effect for males and females. The result for the test of endogeneity suggests that one can reject the null hypothesis that SSP is exogeneous in the regression.

Overall, we have found that expected longevity affects positively the increase in the number of hours worked per week. However, this effect is not significant. Furthermore, we only find marginal evidence that there is a significant positive effect of SSP on hours supplied for females.

7. CONCLUSIONS

Once we confirmed that SSPs in SHARE do fulfil the three properties proposed by Hurd and McGarry (1997), this Chapter focused on the role of SSPs on models of inter-temporal decision models under uncertainty.

Expectations are important inputs to economists' basic research. In particular, expectations on longevity play a significant role in the LCM. According to the LCM there is a well-defined link between the consumption plans of an individual and her income and income expectations as she passes from childhood, through the work participating years, into retirement and eventual decease. A direct consequence of the LCM is that individuals who expect to live exceptionally longer will retire at a later age than individuals who expect to die early.

The main objective of this Chapter was to examine whether SSPs play a significant role on labour supply decisions of old Europeans using SHARE data. For this matter we have used a probit specification to model the probability of entering into retirement as a function of a set of covariates including SSP. Additionally, we proposed an instrumental variable approach to address the possible endogeneity of SSP in the regression. This endogeneity could be a result of two factors, classical measurement error and the so-called justification bias.

Our results suggest that for females, expectations of longevity do play a significant role when taking retirement decisions. In particular, we find that females who expect to live longer have a lower probability of retiring. This finding is consistent with the LCM. This suggests that females tend to act as rational economic agents when taking late life decisions such as continue working or not. Additionally, we found that this might not be true for males. Our results suggest that males do not take into account longevity expectations when deciding when to retire.

We have developed several sensitivity scenarios in order to test the consistency of this interesting result. We have used different definitions of labour supply and subsamples of individuals and the results tend to point towards the same direction. Moreover, because it might be the case that Europeans choose alternative retirement pathways such as decreasing the number of hours worked gradually, instead of taking a one-off retirement decision, we estimate the role of SSP on the change in the number of hours worked per week. We found that females expecting to live longer tend to increase their number of hours supplied and males do not modify their hours supplied as a consequence of their expected longevity.

These results are similar to those obtained by Hurd *et al.* (2004) and O`Donnell *et al.* (2008). The former found that the effect of SSP on the retirement probability before age of 62 is negligible. However, after age 62 workers with very low survival probability do leave the labour force earlier than those with higher survival expectations. It should be mentioned that the authors did not conduct a separate analysis by sex. O`Donnell *et al.* (2008) found that for both males and females, SSPs play a role on the decision of retiring. In particular the authors found that for males, those that are pessimistic about their survival are least likely to retire. However, as expectations improve, the retirement probability first rises, but then falls over most of the distribution of survival expectations. Contrary, for females, a linear relationship was found.

But why do we find that SSP has an effect on labour supply for females and not for males? Economic literature analysing the role of longevity expectations on retirement does not provide any insight on this issue because they either don't find a significant relationship, they do not conduct separate analysis by sex, or simply they don't discuss the issue. We believe it is important to analyse this interesting result.

One of the reasons why longevity expectations do play a role on female labour supply but not on male labour supply could be related to the fact that males and females have very different retirement length expectations. This is due to two commonly accepted facts. On the one hand, it is a well-known mortality fact that, on average, females live longer than males. In fact, according to the Human Mortality Database (HMD) the difference in life expectancy between females and males in Europe is, on average, 5.2 years. On the other hand, within a married couple wives are, on average, younger than their husbands. This is confirmed in our database as the difference between the wives' age and the husbands' age is, on average, 3.3 years.

These two facts suggest that wives typically will anticipate a longer retirement period than their husbands, and part of that retirement period is likely to be as widow. And this will have important implications when taking late life decisions. This has been explored by some authors. For example, Browning (1995), Lundberg and Ward-Batts (2000), and Lundberg *et al.* (2003) suggested that husbands and wives may have different private interests in savings and wealth accumulation. In particular, because females think they will live longer than males, they need to save more. Moreover, females may prefer lower per period consumption in order to spread resources over their longer life.

Therefore, if females attach a high probability to reaching 75 years old, that might mean that with high probability they will spend some years as widows. If that is the case, females will have the incentive to delay retirement so that the household could accumulate more wealth. Contrary, if males attach a high probability to reaching 75 years old, that does not necessarily means that they will spend some years as widower. Hence, the incentives to save are lower.

In other words, the risk aversion that females fear to fall short of savings and wealth accumulation could have the key to explain why longevity expectations play a role on retirement decisions for females but not for males. Future research could also provide further explanations.

Our analysis is the first approximation to the role of SSP on labour supply using SHARE data. We have found very interesting conclusions that clearly motivate future research. It would be interesting to also consider the characteristics of the different early retirement pathways at a national level and, more importantly, the institutional constraints that an individual may face when retiring before official age. Furthermore, given that SHARE, HRS and ELSA surveys are harmonized, this could be done jointly for the US, the UK, and European countries.

Additionally, it should be taken into account that our analysis is based only on the first two waves of SHARE. Future research could benefit from the fact that once further waves are released in the coming years, the results will be more stable.

As it was mentioned in the Introduction, understanding how expected longevity affects labour supply decisions is a crucial topic for policymakers. Throughout the 20th century the developed world has seen an impressive increase in life expectancy and this trend is also expected to continue during the 21st century. Given the increasing financial pressure on Social Security systems across Europe policymakers should be interested in understanding how to incentive people to work longer so that the sustainability of the system is secured. This Chapter sheds light on this issue and contributes to the public debate.

8. REFERENCES

- Arellano, M. (2008). "Binary Models with Endogenous Explanatory Variables". Class notes. CEMFI. Available at http://www.cemfi.es/~arellano/binary-endogeneity.pdf

- Attanasio, O. (1998). "Consumption Demand". *NBER Working Paper,* 6466. National Bureau of Economic Research, Cambridge, MA, March.

- Benitez-Silva, H. (2000). "Micro Determinants of Labor Force Status among Older Americans". *Working Paper,* 00-07, State University of New York at Stony Brook, Department of Economics.

- Biró, A. (2011). "Subjective Mortality Hazard Shocks and the Adjustment of Consumption Expenditures". Central European University PhD thesis.

- Bloom, D., Canning, D., Graham, B. (2003). "Longevity and Life-Cycle Savings". *Scandinavian Journal of Economics*, 105: 319-38.

- Bloom, D., Canning, D., Moore, M., Song, Y. (2006). "The Effect of Subjective Survival Probabilities on Retirement and Wealth". Workshop on Comparative International Research Based on HRS, ELSA and SHARE, RAND.

- Bloom, D., Canning, D., Moore, M., Song, Y. (2007). "The Effect of Subjective Survival Probabilities on Retirement and Wealth in the United States". In: Clark, R. L., Ogawa, N., Mason, A., *Population Aging, Intergenerational Transfers and the Macroeconomy,* Cheltenham, Edward Elgar Publishing Limited: 67-100.

- Börsch-Supan, A., Brugiavini, A., Jürges, H., Kapteyn, A., Mackenbach, J., Siegrist, J., Weber, G. (2005). *Health, Ageing and Retirement in Europe, First Results from Share*, Mannheim, MEA: 370.

- Browning, M. (1995). "Saving and the Intra-Household Distribution of Income: an Empirical Investigation". *Ricerche Economiche*, 49(3): 277-292.

- Cocco, J., Gomes, F. (2012). "Longevity Risk, Retirement Savings, and Financial Innovation". *Journal of Financial Economics*, 103(3): 507-529.

- Delavande, A., Perry, M., Willis, R. (2006). "Probabilistic Thinking and Early Social Security Claiming". *Working Paper,* 2006-129, Michigan Retirement Research Center, University of Michigan.

- Fang, H., Keane, M., Khwaja, A., Salm, M., Silverman, D. (2006). "Testing the Mechanisms of Structural Models: The Case of the Mickey Mantle Effect". *American Economics Review*, 97(2): 53-59.

- Feinstein, J. (1993). "The Relationship between Socioeconomic Status and Health: A Review of the Literature". *The Milbank Quarterly*, 71: 279-322.

- Fischer, J., Sousa-Poza, A. (2006). "The Institutional Determinants of Early Retirement in Europe". *Discussion Paper*, 2006-08, University of St. Gallen.

- Gan, L., Hurd, M., McFadden, D. (1998). "Subjective Survival Curves and Lifecycle Behaviour". In: Wise, D. (ed.), *Inquiries in the Economics of Aging*, Chicago, University of Chicago Press: 259-309.

- Gan, L., Goug, G., Hurd, M., McFadden, D. (2004). "Subjective Mortality Risk and Bequests". *NBER Working Paper,* 10789, National Bureau of Economic Research, Cambridge, MA, September.

- Gan, L., Hurd, M., McFadden, D. (2005). "Individual Subjective Survival Curves". *Analyses in Economics of Aging:* 377-411.

- Groneck, M., Ludwing, A., Zimper, A. (2011). "A Life-Cycle Consumption Model with Ambiguous Survival Beliefs". *NETSPAR Working Paper,* 044.

- Hospido, L., Zamarro, G. (2012). "Retirement Patterns of Couples in Europe". *Working Paper,* presented at the European Economic Association & Econometric Society in Malaga, Spain.

- Hurd, M., McGarry, K. (1997). "The Predictive Validity of Subjective Probabilities of Survival". *The Economic Journal,* 112(482): 966-985.

- Hurd, M., Mcfadden, D., Merrill, A. (2001). "Predictors of Mortality among the Elderly". In: Wise, D. (ed.), *Themes in the Economics of Aging*, Chicago, University of Chicago Press: 171-198.

- Hurd, M., Smith, P. (2001). "Anticipated and Actual Bequests". In: Wise, D. (ed.), *Themes in the Economics of Aging*, Chicago, University of Chicago Press: 357-392.

- Hurd, M., Rohwedder, S. (2003). "The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending at Retirement". *NBER Working Paper*, 9586, National Bureau of Economic Research, Cambridge, MA, March.

- Hurd, M., Smith, J., Zissimopoulos, J. (2004). "The Effects of Subjective Survival on Retirement and Social Security Claiming". *Journal of Applied Econometrics*, 19: 761-775.

- Kalemli-Ozcan, S., Weil, D. (2010). "Mortality Change, the Uncertain Effect, and Retirement". *Journal of Economic Growth*, 15(1): 65-91.

- Lundberg, S., Ward-Batts, J. (2000). "Saving for Retirement: Household Bargaining and Household Net Worth". *Working Paper,* 2000-004, Michigan Retirement Research Center, University of Michigan.

- Lundberg, S., Startz, R., Stillman, S. (2003). "The Retirement-Consumption Puzzle: A Marital Bargaining Approach". *Journal of Public Economics,* 87(5–6): 1199–1218.

- O'Donnell, O., Teppa, F., van Doorslaer, E. (2008). "Can Subjective Expectations Explain Retirement Behaviour?".*Working Paper,* 188, De Nederlansche Bank.

- Perry, M. (2005). "Estimating Life Cycle Effects of Subjective Survival Probabilities in the Health and Retirement Study". *Working Paper*, 2005-103, Michigan Retirement Research Center, University of Michigan.

- Post, T., Hanewald, K. (2011). "Longevity Risk, Subjective Survival Expectations and Individual Saving Decisions". NETSPAR *Working Paper,* 043.

- Rivers, D., Vuong, H. (1988), "Limited Information Estimators and Exogeneity tests for simultaneous Probit Models", *Journal of Econometrics,* 39: 347–66.

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

- Salm, M. (2006). "Can Subjective Mortality Expectations and Stated Preferences Explain Varying Consumption and Saving Behaviour among the Elderly?",*IZA Discussion Paper,* 2467.

- Steffen, B. (2009), "Formation and Updating of Subjective Life Expectancy: Evidence from Germany", *MEA Studies*, 8.

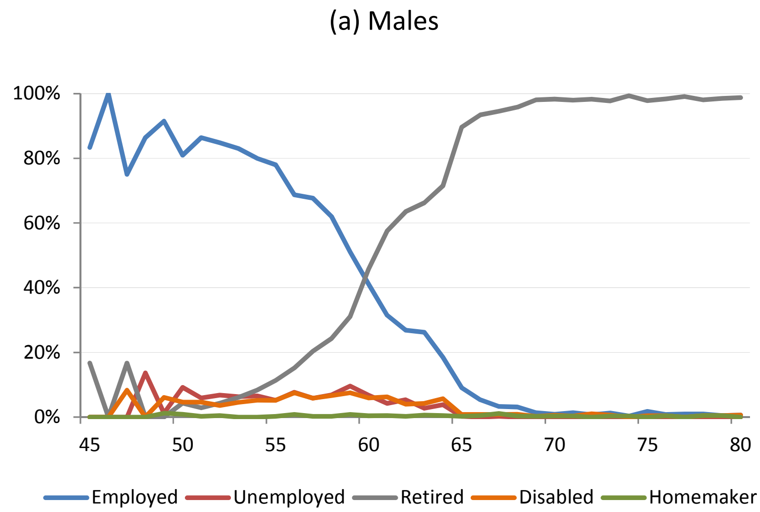
- Stern, S. (1989). "Measuring the Effects of Disability on Labor Force Participation". *Journal of Human Resources*, 24: 361-395.

- Waldron, H. (2001). "Links between Early Retirement and Mortality". *ORES Working Paper*, 93, Division of Economic Research, Social Security Administration.

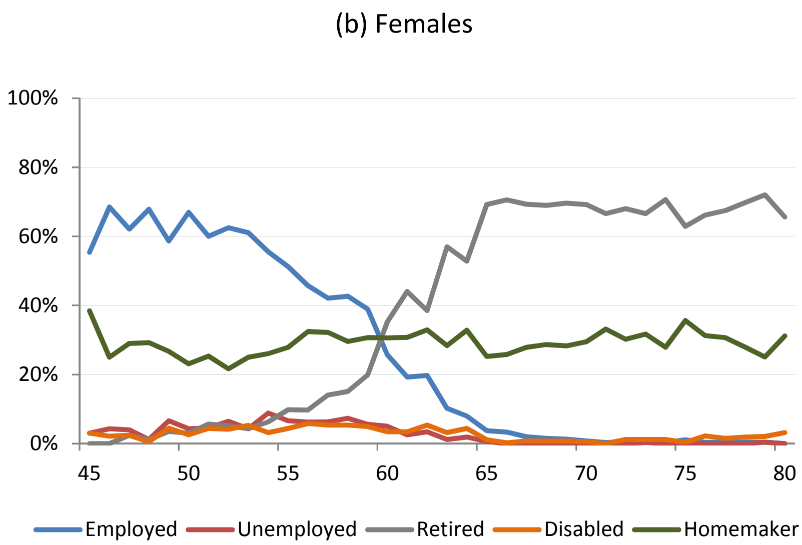
- Wolfe, J. (1983). "Perceived Longevity and Early Retirement". *The Review of Economics and Statistics*, 65(4): 544-551.

9. TABLES AND FIGURES

FIGURE 1: SELF-REPORTED LABOUR STATUS

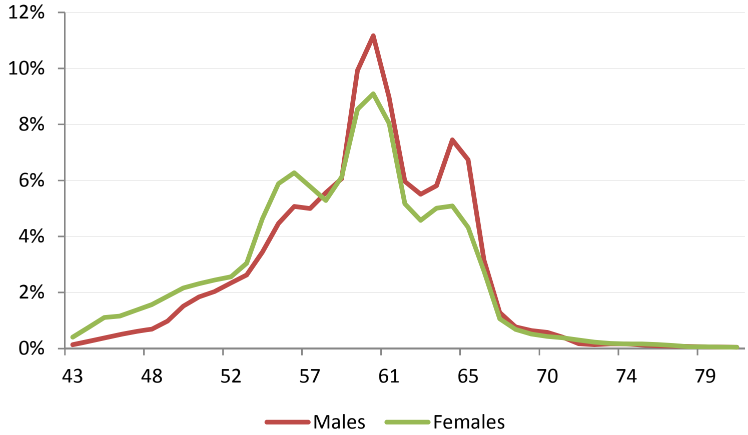


Source: own elaboration



Source: own elaboration

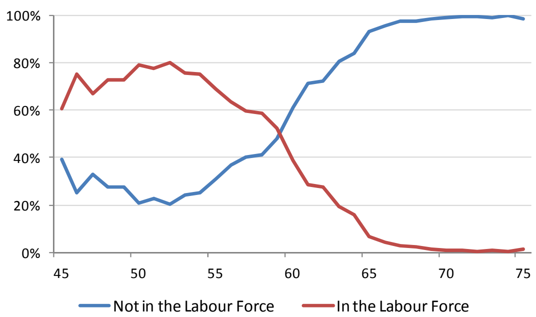
FIGURE 2: AGE WHEN RETIRED



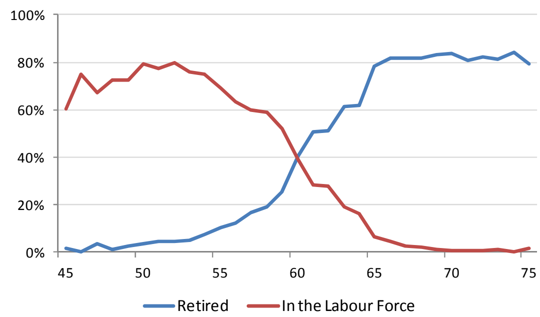
Source: own elaboration

FIGURE 3: LABOUR SUPPLY PATTERN BASED ON DIFFERENT DEFINITIONS

(a) In the labour force and not in the labour force



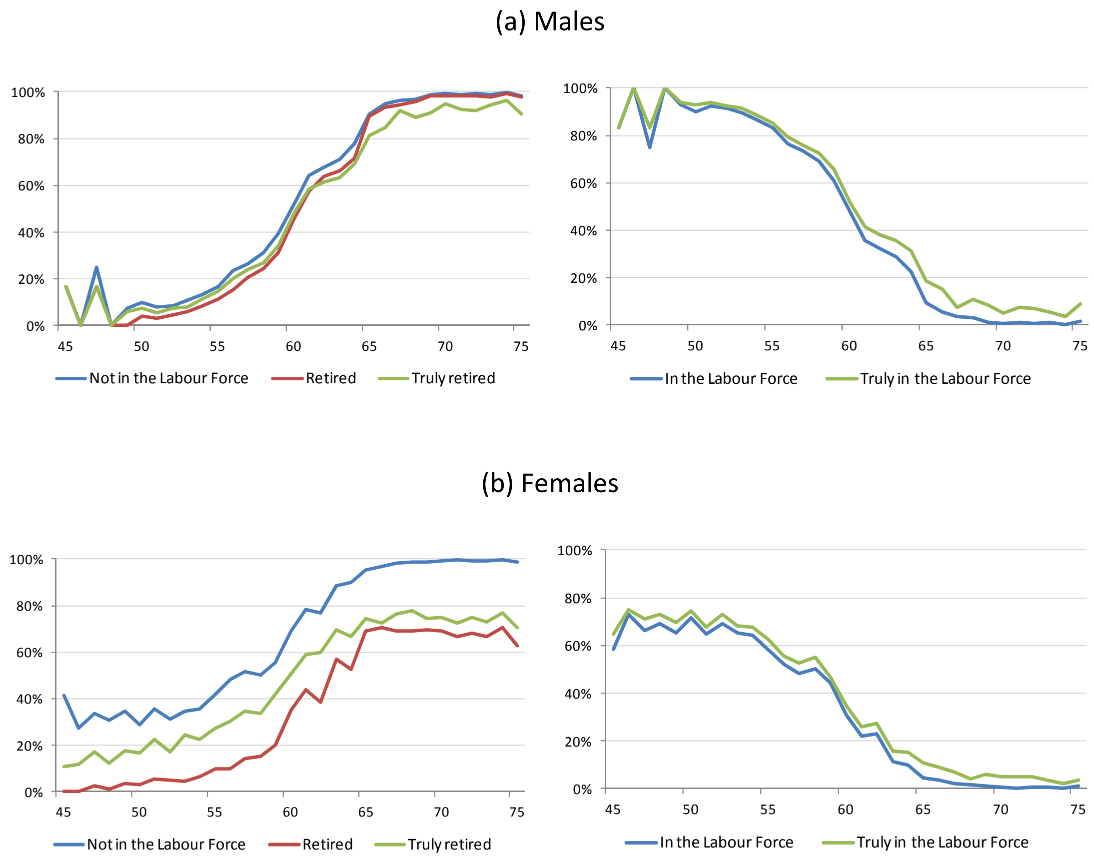
(b) In the labour force and retired



(c)Truly in the labour force and truly retired

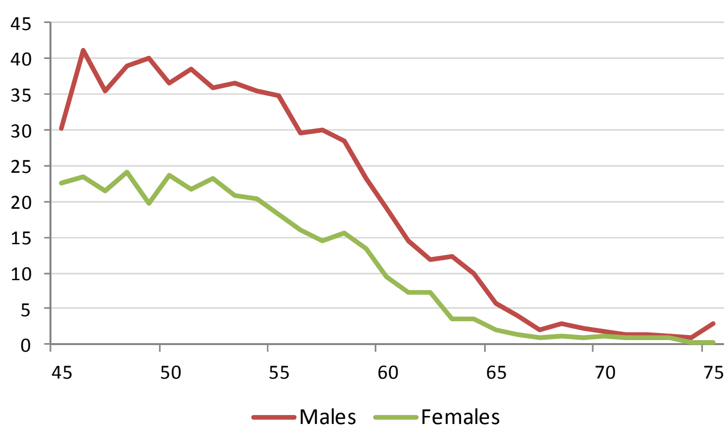
Source: own elaboration

FIGURE 4: LABOUR SUPPLY BY AGE BASED ON DIFFERENT DEFINITIONS



Source: own elaboration

FIGURE 5: TOTAL NUMBER OF HOURS WORKED PER WEEK



Source: own elaboration

TABLE 1: LABOUR SUPPLY TRANSITION

From (wave 1) To (Wave 2)	In the labour force Not in the labour force	In the labour force Retired	Truly in the labour force Truly retired
Full sample	19.5%	14.1%	9.2%
Males	19.7%	16.0%	9.9%
Females	19.3%	12.1%	8.4%

Source: own elaboration

TABLE 2: TRANSITION IN THE NUMBER OF HOURS WORKED PER WEEK

Working hours in the Wave 2 compared with Wave 1	Decreased			Constant	Increased
	> 39 hours	20-39 hours	< 20 hours		
Full sample	10.94%	11.94%	29.36%	27.15%	20.61%
Males	14.56%	10.89%	27.59%	27.70%	19.26%
Females	6.99%	13.09%	31.28%	26.54%	22.10%

Source: own elaboration

TABLE 3: CORRELATION MATRIX

	Retired	SSP	Age	Married	Education	Income	Children	Long-t. illness	Adl	Iadl	Severe illness	Non-severe illness	Self-p. health	Wealth	Receive inherit.
Retired	1	-0.06	0.28	-0.04	-0.05	-0.02	0.00	0.05	0.05	0.04	0.05	0.09	0.12	-0.01	-0.06
SSP	0.02	1	0.06	0.03	0.10	-0.01	0.03	-0.09	-0.13	-0.11	-0.16	-0.08	-0.27	0.05	0.08
Age	0.30	0.07	1	-0.13	-0.10	-0.03	0.01	0.06	0.02	0.01	0.15	0.14	0.05	0.02	-0.15
Married	0.02	0.05	-0.01	1	-0.04	0.08	0.21	-0.03	0.00	0.00	-0.02	-0.07	-0.02	0.08	0.04
Education	-0.08	0.02	-0.05	0.01	1	0.07	-0.08	-0.06	-0.06	-0.10	-0.08	-0.07	-0.22	0.05	0.22
Income	-0.03	0.04	-0.02	0.08	0.14	1	0.00	-0.03	0.03	-0.02	0.00	-0.01	-0.05	0.05	0.01
Children	-0.02	0.02	0.01	0.23	0.00	0.05	1	0.01	0.03	0.02	0.00	0.01	-0.02	0.03	0.01
Long-term illness	0.06	-0.12	0.07	-0.02	-0.04	-0.01	-0.01	1	0.14	0.17	0.26	0.31	0.39	0.00	-0.02
Adl	0.04	-0.10	0.01	-0.03	-0.07	-0.02	0.03	0.13	1	0.32	0.06	0.15	0.20	-0.01	-0.04
Iadl	0.02	-0.10	0.02	-0.02	-0.06	-0.02	0.03	0.10	0.48	1	0.08	0.14	0.23	-0.03	-0.02
Severe illness	0.06	-0.11	0.13	0.00	-0.05	-0.01	0.00	0.30	0.09	0.03	1	0.16	0.28	-0.02	-0.07
Non-severe illness	0.05	-0.08	0.06	-0.01	-0.08	-0.02	0.01	0.28	0.15	0.08	0.14	1	0.29	-0.01	-0.04
Self-p. health	0.09	-0.28	0.03	-0.04	-0.15	-0.06	-0.02	0.38	0.18	0.13	0.30	0.26	1	-0.05	-0.15
Wealth	-0.03	0.05	0.05	0.07	0.07	0.08	0.05	-0.04	0.01	-0.01	-0.01	-0.02	-0.04	1	0.03
Receive inherit.	-0.09	0.04	-0.15	0.04	0.15	0.07	0.04	-0.02	-0.01	0.00	-0.02	-0.04	-0.12	0.04	1

Source: own elaboration

TABLE 4: PROBABILITY OF ENTERING INTO RETIREMENT (TRULY IN THE LABOUR FORCE AND TRULY RETIRED)

Probability of entering into retirement	Probit without health variables				Probit with health variables				IV Probit			
	Males		Females		Males		Females		Males		Females	
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
SSP	**0.0001**	**0.002**	**-0.005 *****	**0.002**	**0.001**	**0.002**	**-0.003 ***	**0.002**	**-0.002**	**0.029**	**-0.012**	**0.021**
Age	0.367	0.281	-0.048	0.254	0.346	0.282	-0.085	0.254	0.348	0.281	-0.053	0.269
Age squared	-0.186	0.241	0.174	0.222	-0.168	0.242	0.207	0.222	-0.169	0.241	0.179	0.238
Married	0.234 **	0.116	0.081	0.099	0.242**	0.117	0.109	0.098	0.248*	0.128	0.114	0.096
Education	-0.086 ***	0.028	-0.005	0.028	-0.079***	0.028	0.015	0.029	-0.080***	0.028	0.019	0.030
Income	-0.001	0.001	-0.000	0.000	-0.001	0.001	-0.000	0.000	-0.001	0.001	-0.000	0.000
Children	-0.039	0.032	0.031	0.035	-0.042	0.032	0.030	0.035	-0.042	0.032	0.032	0.034
Long-term illness					0.140*	0.085	0.019	0.095	0.136	0.093	0.026	0.094
Adl					0.144	0.125	0.114	0.139	0.138	0.142	0.049	0.204
Iadl					-0.060	0.142	-0.022	0.133	-0.073	0.190	-0.040	0.135
Severe illness					0.009	0.044	-0.041	0.053	0.007	0.049	-0.067	0.074
Non-severe illness					-0.028	0.077	0.013	0.068	-0.028	0.077	0.012	0.067
Self-perceived health					0.027	0.045	0.170***	0.055	0.007	0.194	0.119	0.135
Wealth	-0.086 **	0.041	-0.034	0.032	-0.082**	0.040	-0.038	0.034	-0.080*	0.044	-0.033	0.035
Receiving inheritance	-0.001	0.001	0.002*	0.001	-0.001	0.001	0.002**	0.001	-0.001	0.001	0.003**	0.001

Constant	-15.71 *	8.161	-4.013	7.225	-15.30*	8.173	-3.715	7.273	-15.11*	8.540	-3.730	7.116
Observations	2,863		2,648		2,863		2,648		2,863		2,648	
Log-likelihood	-715.43		-580.24		-712.19		-575.42		-13,635.97		-12,465.92	
Wald test of exogeneity (p-value)									0.919		0.666	

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

TABLE 5: PROBABILITY OF ENTERING INTO RETIREMENT (TRULY IN THE LABOUR FORCE AND TRULY RETIRED). RESTRICTED SUBSAMPLE

Probability of entering into retirement	**Probit without health variables**				**Probit with health variables**				**IV Probit**			
	Males		**Females**		**Males**		**Females**		**Males**		**Females**	
	Coeff.	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**
SSP	**-0.0005**	**0.002**	**-0.005 *****	**0.002**	**0.003**	**0.002**	**-0.003 ***	**0.002**	**0.002**	**0.037**	**-0.026**	**0.023**
Observations	2,580		2,115		2,580		2,115		2,580		2,115	
Log-likelihood	-687.95		-520.59		-684.71		-511.16		-12,197.1		-9,923.15	
Wald test of exogeneity (p-value)									0.971		0.402	

Note: Robust standard errors are reported. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

TABLE 6: PROBABILITY OF ENTERING INTO RETIREMENT (IN/NOT IN THE LABOUR FORCE)

Probability of entering into retirement	Probit without health variables				Probit with health variables				IV Probit			
	Males		Females		Males		Females		Males		Females	
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
SSP	**-0.003** **	**0.001**	**-0.003** ***	**0.001**	**-0.001**	**0.001**	**-0.002**	**0.001**	**-0.009**	**0.035**	**-0.020**	**0.016**
Observations	2,689		2,426		2,689		2,426		2,689		2,426	
Log-likelihood	-993.96		-947.08		-984.97		-940.21		-13,102.06		-11,791.32	
Wald test of exogeneity (p-value)									0.816		0.308	

Note: Robust standard errors are shown in parenthesis. *** p<0.01; ** p<0.05; * p<0.1. Source: own elaboration

TABLE 7: PROBABILITY OF ENTERING INTO RETIREMENT (IN/NOT IN THE LABOUR FORCE). RESTRICTED SUBSAMPLE

Probability of entering into retirement	Probit without health variables				Probit with health variables				IV Probit			
	Males		Females		Males		Females		Males		Females	
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
SSP	**-0.003** **	**0.001**	**-0.004** **	**0.002**	**-0.002**	**0.002**	**-0.002**	**0.002**	**-0.024**	**0.029**	**-0.035** ***	**0.010**
Observations	2,416		1,927		2,416		1,927		2,416		1,927	
Log-likelihood	-930.39		-826.46		-923.22		-821.01		-11,686.51		-9,372.82	
Wald test of exogeneity (p-value)									0.468		0.034	

Note: Robust standard errors are shown in parenthesis. *** p<0.01; ** p<0.05; * p<0.1. Source: own elaboration

TABLE 8: PROBABILITY OF ENTERING INTO RETIREMENT (IN THE LABOUR FORCE AND RETIRED)

Probability of entering into retirement	Probit without health variables				Probit with health variables				IV Probit			
	Males		Females		Males		Females		Males		Females	
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
SSP	**-0.0004**	**0.001**	**0.001**	**0.002**	**-0.0001**	**0.001**	**0.001**	**0.002**	**0.011**	**0.027**	**-0.007**	**0.024**
Observations	2,689		2,426		2,689		2,426		2,689		2,426	
Log-likelihood	-821.98		-649.44		-820.36		-645.59		-12,937.39		-11,497.3	
Wald test of exogeneity (p-value)									0.691		0.723	

Note: Robust standard errors are shown in parenthesis. *** p<0.01; ** p<0.05; * p<0.1. Source: own elaboration

TABLE 9: PROBABILITY OF ENTERING INTO RETIREMENT (IN THE LABOUR FORCE AND RETIRED). RESTRICTED SUBSAMPLE

Probability of entering into retirement	Probit without health variables				Probit with health variables				IV Probit			
	Males		Females		Males		Females		Males		Females	
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
SSP	**-0.001**	**0.002**	**-0.001**	**0.002**	**-0.001**	**0.002**	**-0.001**	**0.002**	**-0.016**	**0.027**	**-0.025** ***	**0.024**
Observations	2,416		1,927		2,416		1,927		2,416		1,927	
Log-likelihood	-803.5		-602.2		-801.7		-597.96		-11,565.35		-9,151.57	
Wald test of exogeneity (p-value)									0.566		0.406	

Note: Robust standard errors are shown in parenthesis. *** p<0.01; ** p<0.05; * p<0.1. Source: own elaboration

TABLE 10: CHANGE IN THE NUMBER OF HOURS WORKED PER WEEK

Change in hours worked (Wave 2 – Wave 1)	OLS without health variables				OLS with health variables				IV OLS			
	Males		Females		Males		Females		Males		Females	
	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.	Coeff.	Std. Dev.
SSP	**0.017**	**0.017**	**0.026**	**0.017**	**0.001**	**0.018**	**0.023**	**0.018**	**0.028**	**0.190**	**0.141**	**0.189**
Age	2.941	2.402	3.440*	1.958	3.251	2.391	3.249*	1.969	3.240	2.381	2.982	2.099
Age squared	-3.983 *	2.146	-3.916**	1.806	-4.218**	2.135	-3.745**	1.816	-4.219**	2.126	-3.546*	1.906
Married	-1.532	1.075	0.361	0.962	-1.595	1.083	0.208	0.963	-1.640	1.138	0.227	0.966
Education	1.127 ***	0.297	0.676**	0.262	0.968***	0.299	0.657**	0.265	0.983***	0.310	0.650**	0.264
Income	0.001	0.009	0.006	0.007	-0.001	0.009	0.006	0.007	-0.001	0.009	0.004	0.007
Children	0.086	0.346	-0.171	0.348	0.144	0.347	-0.145	0.348	0.142	0.346	-0.128	0.348
Long-term illness					0.904	0.923	-2.379**	0.971	0.934	0.946	-2.445**	0.990
Adl					-2.795	1.710	-3.381**	1.665	-2.727	1.757	-2.462	2.203
Iadl					-2.889	1.768	-0.821	1.509	-2.799	1.854	-0.861	1.528
Severe illness					-0.909*	0.496	0.735	0.472	-0.877	0.542	1.046	0.708
Non-severe illness					-0.511	0.936	-0.878	0.866	-0.491	0.934	-0.988	0.877
Self-perceived health					-0.738	0.514	0.569	0.480	-0.562	1.325	1.184	1.149
Wealth	0.149	0.271	-0.167	0.216	0.176	0.284	-0.116	0.221	0.165	0.291	-0.183	0.259
Receiving inheritance	-0.004	0.011	0.007	0.010	-0.005	0.011	0.008	0.010	-0.005	0.011	0.005	0.010

Constant	-55.78	66.92	-85.13	53.25	-61.61	66.67	-79.96	53.37	-63.51	67.74	-82.23	52.89
Observations	2,487		2,276		2,487		2,276		2,487		2,276	
P-value	0.127		0.085		0.134		0.093		0.133		0.074	
Tests of endogeneity (p-value)									0.887		0.526	
Test overidentifying restrictions (p-value)									0.532		0.986	

Note: Robust standard errors are shown in parenthesis. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration

TABLE 11: CHANGE IN THE NUMBER OF HOURS WORKED PER WEEK. RESTRICTED SUBSAMPLE

Change in hours worked (Wave 2 – Wave 1)	**OLS without health variables**				**OLS with health variables**				**IV OLS**			
	Males		**Females**		**Males**		**Females**		**Males**		**Females**	
	Coeff.	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**	**Coeff.**	**Std. Dev.**
SSP	**0.017**	**0.021**	**0.041 ***	**0.024**	**-0.005**	**0.022**	**0.042 ***	**0.025**	**0.016**	**0.198**	**0.267**	**0.252**
Observations	1,766		1,411		1,766		1,411		1,766		1,411	
P-value	0.109		0.088		0.119		0.101		0.095		0.048	
Tests of endogeneity (p-value)									0.403		0.352	
Test overidentifying restrictions (p-value)									0.378		0.980	

Note: Robust standard errors are shown in parenthesis. *** p<0.01; ** p<0.05; * p<0.1
Source: own elaboration