Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Vesile Kutlu

Essays on Subjective Well-Being and Retirement Behavior

Research Master Thesis 2011-014

Essays on Subjective Well-Being and Retirement Behavior

by

Vesile Kutlu

ANR: 202820

Research Master Thesis in Economics

Tilburg University

Thesis Supervisor: Prof. Arthur van Soest

Second Reader: Prof. Adriaan Kalwij

JULY, 2011

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my supervisor, Prof. Arthur van Soest, for his excellent guidance and encouragement through the development of this thesis. I gratefully acknowledge research support provided by Netspar via Research Master Scholarship. I would like to thank the participants of Junior Pension Day for their valuable comments and advices.

I am very thankful to my family for their encouragement during my studies. Finally, I would like to give my special thanks to Emre Koç for his invaluable support.

TABLE OF CONTENTS

ABSRTACT

This thesis consists of two chapters. In the first chapter, we are primarily interested in the effect of expectations on psychological well-being of Americans over age 50 and their spouses. In order to circumvent mood effects on both independent and dependent variables, we devise an exogenous proxy for life expectancy of parents. Using the life table survival probabilities, we examine the effect of expectations on psychological well-being and conclude whether psychological well-being is a forward- looking phenomenon in this context.

The sensitivity of psychological well-being to longevity expectations may indicate that older workers anticipate future events and they might incorporate these predictions into their economic decisions. The predictions about life expectancy might particularly affect the choice of retirement timing and leisure preferences of older workers. Therefore, in the second chapter of this thesis we investigate the link between retirement behavior and longevity expectations which are gathered from population life tables.

Overall, the findings of this thesis are as follows: i) there is a causal relationship between longevity expectations and current psychological well-being of individuals, ii) individuals experience less mental distress as their mothers' life expectancy increases, iii) there is a nonlinear relationship between life expectancy and probability of retirement, iv) on average, increased life expectancy leads to delay in retirement.

CHAPTER I: SUBJECTIVE WELL-BEING AND LONGEVITY EXPECTATIONS

1.1 INTRODUCTION

Subjective well-being is a commonly used proxy for the actual utility level of individuals. Happiness research highlights five basic points concerning the determinants of subjective well-being. Firstly, the current well-being of an individual is mainly determined by current factors such as age, marital status, health status, and job status of the individual (Winkelmann and Winkelmann, 1998). Secondly, reported well-being is mostly affected by the relative income level of the individual compared to his target group, rather than his absolute income. (Knight et al, 2009) Thirdly, individuals are strongly influenced by recent jumps in their income. Fourthly, the effects of income changes die out gradually and individuals aspire to new goals after a while (Clark et al, 2008a). Finally, current well-being is an inter-temporal concept in the sense that it is influenced by future income expectations as well as past realizations. (Knight et al, 2009)

The first chapter of this thesis mainly aims to answer the following question: To what extent is self-reported well-being a forward-looking phenomenon? Although researchers have tried to answer this question previously, their findings are limited due to the difficulty of measuring individuals' expectations. In the standard life-cycle models, the current utility of individuals depends on their permanent income which is a function of their expected future income. Therefore, in theory, an effect of future income on current utility is consistent with inter-temporal utility maximization. From empirical perspective, Clark et al (2008b) test the effect of marriage anticipation on current happiness by means of baseline satisfaction analysis in a panel study. The authors observe fluctuations in life satisfaction before and after the marriage of an individual compared to the baseline level. If the individual's life satisfaction increases relative to the baseline satisfaction during the four years prior to marriage then they conclude that the individual anticipates that he/she will be married in the future.

In a cross-section study, Knight et al (2009) find out that, expected changes in income have *ex ante* effects on subjective well-being using subjective expectations of survey respondents. The suggested relationship between subjective well-being and expected income may not be causal if mood effects are present. More explicitly, if income expectations are exposed to the respondent's optimism or pessimism during the survey then they include unobservable mood effects which are not necessarily constant over time.[1] In this case, both the dependent and the independent variables may be influenced by a third factor and the results may simply be spurious. In fact, Schwarz and Clore (1983) find out that survey responses to well-being questions are sensitive to weather conditions. The authors indicate that, the respondents who are interviewed on sunny days report higher life satisfaction and happiness compared to respondents who are interviewed on rainy days. Since reported expectations may also be sensitive to such context effects, using subjective expectations of survey respondents may not imply a causal relationship between expectations and current well-being of individuals.[2]

This thesis proposes an exogenous proxy for expectations, which is definitely not influenced by aforementioned mood effects. This proxy, namely life table survival probabilities of parents, is used to examine whether the respondents' current well-being is affected by their expectations regarding parental longevity. To this end, we use seven waves of Health and Retirement Study (HRS) which represents the population of Americans over age 50 and their spouses and covers the years from 1994 to 2006. To our knowledge, life table survival probabilities have not been previously used as a covariate in a well-being equation.

As opposed to subjective survival probabilities, life table survival probabilities are not influenced by time-variant unobservables. Furthermore, the mortality literature demonstrates that life table survival probabilities predict realizations accurately and individuals' own mortality expectations, namely their answers to expected survival questions, are largely in line with population life tables. (Hamermesh, 1985; Elder, 2007; Post and Hanewald, 2010). Since mortality of the respondent is strongly correlated with his or her parent`s mortality (Hamermesh, 1985; Elder, 2007), this thesis presumes that an individual can predict the expected remaining lifetime of his or her parents without any systematic bias.[3] Therefore, the life table survival probabilities are used to predict the respondent`s subjective expectancy concerning his or her parent`s remaining lifetime.

[1] If mood effects were constant over time, then using fixed effects estimation would result in a causal relationship between subjective expectations and well-being.
[2] To measure a causal relationship between well-being and expectations some studies use natural experiments. For example, the exogenous nature of the 2006 Dutch pension reform is utilized to demonstrate that, pension reforms that have adverse implications for the working population lead to *ex ante* drop in subjective well-being (De Grip et al, 2009), lower job efforts and substantial productivity losses (Montizaan, 2010).
[3] To our knowledge, there are no studies which investigate prediction of parents' remaining life time.

The structure of this chapter is organized as follows: Section 1.2 presents related studies in the literature. Data sources are explained in Section 1.3. The methodology is introduced in Section 1.4. Estimation results are presented in Section 1.5. Finally, Section 1.6 gives concluding remarks.

1.2 RELATED LITERATURE

Subjective well-being is a commonly used proxy for the actual utility level of individuals. Although economists usually rely on revealed preferences in order to measure the utility level, the subjective well-being approach has some advantages. According to Frey and Stutzer (2005), subjective well-being approach does not require a market equilibrium assumption, unlike the conventional revealed preferences approach. As this assumption may not hold due to common rigidities and imperfections, the subjective well-being approach may add another dimension to welfare analysis.

The relationship between current well-being and predictability of future events has been investigated previously; but related studies are limited due to the difficulty of measuring individuals' expectations. In a cross-section study, Knight et al (2009) observe that current well-being is sensitive to expected future changes using the instrumental variable (IV) Probit estimator to correct for the possible endogeneity of the expected income.[4] Clark et al (2008b) test the effect of marriage anticipation on current happiness by means of baseline satisfaction analysis in a panel study. According to their findings, an individual's life satisfaction increases relative to the baseline level during the four years prior to marriage which may imply that the individual anticipates that he or she will be married in the future. The anticipation effect is stronger as the prediction horizon is shorter and it approaches zero for sufficiently long prediction horizons. Although they observe a causal relationship between well-being and being married in the future, their forward-looking analysis relies on strong assumptions such as predictability of future events.

[4] Although using an IV estimator results in a causal relationship, the validity of the suggested instruments is always open to discussion.

As opposed to subjective survival probabilities, life table survival probabilities are not influenced by time-variant unobservables.[5] Empirical studies reveal that individuals are capable of predicting their remaining lifetime with reasonable accuracy and, therefore, these predictions are in line with life-table values. Hamermesh (1985) finds that individuals are fully aware of movements in today's life tables and they incorporate changes in life tables into their subjective survival probabilities. His analysis is based on a non-population representative sample (including PhD economists) as well as a randomly drawn sample (including people chosen from telephone directory). He also shows that on average the respondents' subjective probabilities are roughly coincide with life table probabilities although the distribution of subjective probabilities has greater variance than that of their objective counterparts and this variance increases as individuals have long-lived relatives.

Using the HRS sample, Elder (2007) finds out that, regardless of socioeconomic status, individuals who are older than 50 overestimate whereas the others underestimate their own remaining lifetime. The author points out that, the subjective probabilities contain information which is not contained in the life tables. However, the author also demonstrates that contribution of this additional information is modest. In fact, for those over 64 life tables are stronger predictors of mortality than subjective survival probabilities. Post and Hanewald (2010) investigates whether SHARE respondents are aware of stochastic mortality which refers to the unexpected changes in mortality rates over time. He obtains that respondents are aware of stochastic mortality and their subjective probabilities co-vary systematically with the variation in life table values.

Based on the findings, the life table survival probabilities are used to predict the respondent`s subjective expectancy concerning his or her parent`s remaining lifetime.

1.3 DATA

This thesis utilizes the HRS sample which represents the population of Americans over the age 50 and their spouses and consists of 9 waves covering the years from 1992 to 2008. The data is drawn from the RAND HRS database which is a cleaned and user-friendly version of the HRS raw files. The RAND HRS database is well-suited for the purpose of this thesis since it is a large

[5] These probabilities are based on the mortality rates of the population observed in each year.

sample of elderly population who are interviewed in subsequent waves and it includes detailed information on well-being, health, employment status, and marital status of the respondents and data on age of the respondents' parents.

The RAND HRS data consists of five cohorts: Initial HRS cohort who was born between 1931 and 1941 was first interviewed in 1992. AHEAD cohort born before 1924 was first interviewed in 1993. Children of Depression (CODA) cohort, born 1924 to 1930, War Baby (WB) cohort, born 1942 to 1947, and Early Baby Boomer (EBB) cohort, born 1948 to 1953, were first interviewed in 1998, 1998 and, 2004, respectively. This thesis uses all of the available cohorts in the RAND HRS database except the AHEAD cohort because of two reasons: First, individuals who were born before 1924 are more likely to die in the subsequent waves compared to the members of other cohorts. This may increase the attrition problem which occurs when initial survey respondents drop out of the survey in the subsequent periods. Second, as it is mentioned before, empirical analyses in this thesis rely on the assumption that individuals predict their remaining lifetime with some accuracy and their predictions are in line with the information in the population life tables. The empirical finding that individuals are aware of today's life tables and they form and update their subjective survival probabilities based on the available information in population life tables is less likely to hold in the case of older individuals.[6] Therefore, respondents who were born before 1924 are dropped from the sample.

Subjective well-being is usually measured by survey questions regarding the happiness or life satisfaction level of individuals. As an alternative, mental health index has been commonly used by psychologists as well as economists (Kim and Moen, 2002; Hao, 2008; Finkelstein et al, 2008). This index is based on a score on the Center for Epidemiologic Studies Depression (CES-D) scale. CES-D ranges from 0 to 8 and it is the sum of six negative indicators (depression, everything is an effort, sleep is restless, felt alone, felt sad, could not get going) minus two positive indicators (felt happy and enjoyed life); the higher the score, the more negative the respondent's feelings in the past week.[7] We use the CES-D score as an alternative to happiness and/or life satisfaction measures due to the following reasons. First, life satisfaction question is

[6] Older individuals may have lower cognitive abilities than younger ones.

[7] The CES-D score is defined as the sum of RwDEPRES, RwEFFORT, RwSLEEPR, RwFLONE, RwFSAD, RwGOING, (1-RwWHAPPY) and (1-RwENLIFE) variables of the RAND survey. Each item in the CES-D score stands for a "yes" or "no" question about the feelings of respondents in the past week.

included only in the last three waves of the HRS survey so that the number of observations in our sample drops substantially if we use this measure. Second, a simple survey question about happiness measures only one dimension of individuals' feelings. On the other hand, the CES-D score is more informative about emotional well-being of individuals since it includes both positive and negative indicators. It should be noted that the CES-D score is related to happiness directly but, it is generally an indicator for mental health and depressive symptoms. Following the literature, we will call our dependent variable as 'psychological well-being' or 'mental health' levels of individuals.

Expected remaining life-time of parents is inferred from race, gender, and age specific life-tables of National Vital Statistics Reports which are available for the years from 1992 to 2006. The 1992 wave is excluded from the sample since the allowable responses to the CES-D score are quite different than those in other waves. The 2008 wave is also dropped due to unavailability of life table statistics for that year. In summary, empirical analyses are based on a panel sample which covers the years from 1994 to 2006 and includes Americans born between 1924 and 1953.

Table 1.1: Descriptive Statistics of the Variables Used in the Analysis

	N	*n*	Mean	Std. Dev.	Min	Max
CES-D	75,104	19,426	1.44	1.95	0	8
Mother alive	75,104	19,426	0.27	0.44	0	1
Life expectancy of mother	74,984	19,378	2.26	4.12	0	25.6
Female	75,104	19,426	0.60	0.49	0	1
White	75,104	19,426	0.82	0.39	0	1
Black	75,104	19,426	0.14	0.35	0	1
Education low	75,104	19,426	0.25	0.43	0	1
Education high	75,104	19,426	0.20	0.40	0	1
Age	75,104	19,426	62.34	7.38	40	82
Income in $10,000 per year	75,104	19,426	5.82	17.09	0	2536
Health	75,104	19,426	1.61	1.35	0	8
Net wealth in $10,000	75,104	19,426	31.07	133.23	-473.3	10039
Quartile low (25 %)	75,104	19,426	0.25	0.43	0	1
Quartile high (25 %)	75,104	19,426	0.24	0.43	0	1
Household size	75,104	19,426	2.33	1.17	1	19
Full-time	75,104	19,426	0.32	0.47	0	1
Unemployed	75,104	19,426	0.01	0.10	0	1

Retired	75,104	19,426	0.38	0.48	0	1
Disabled	75,104	19,426	0.03	0.18	0	1
Married	75,104	19,426	0.70	0.46	0	1
Never married	75,104	19,426	0.03	0.17	0	1

Table 1.1 presents descriptive statistics of relevant variables used in the analyses. The corresponding variable definitions and baseline categories for all set of dummies are reported in Table 1.2 The big N and the small n in the Table 1.1 stand for the overall sample size and the number of respondents who are observed over time, respectively. Overall, the sample consists of 75,104 person-year observations. It is an unbalanced sample since the average number of waves in which a respondent was observed is approximately 4 instead of 7.

The reported mean of the dummy variables shows the percentage of respondents who report the outcome 1. Accordingly, 27 % of the respondents have mothers who are still alive. The mean value of the life expectancy of mother is equal to 2.26 suggesting that the expected remaining lifetime of the respondents' mother is quite low on average. However, this is due to the fact that for this variable, the value zero is imputed if the respondent's mother has already died.[8] Health index is based on diagnosis of a variety of diseases by the respondent`s doctor. Therefore, the higher this measure, the less healthy the respondent is. Being an objective measure, this indicator is probably subjected to less reporting bias compared to subjective measures of health status. The mean of health index suggests that, on average, respondents in the sample seem to be healthy. Total income measured in $10,000 has zero values for some respondents.[9] The net wealth denoted in $10,000 has both negative and positive values for the respondents so that it is impossible to present this variable in terms of the natural logarithm. Therefore, three wealth categories are defined with the intention of dividing the distribution of wealth into three parts. These categories are low (Quartile low (25%)), medium (the second and the third quartiles (50%)), and high (Quartile high (25%)).

60 % of the sample consists of females which suggests that women are slightly over-represented compared to men in the sample. About 70 % of respondents are married, 3 % of them are never married and the rest of the sample has other married status. In terms of job status, the sample

[8] See Table 1A in Appendix A for the frequency distribution of this variable.
[9] Although it is not reported, only 0.79 % of the respondents have zero income.

mostly consists of retired individuals. Only 25 % of the respondents have a lower degree than high school diploma which shows that the sample is relatively well-educated.

Table 1.2: Variable definitions

Variable	Description	Baseline
CES-D	Mental Health index (described before)	---
Mother alive	Respondent's mother is alive	Respondent's mother has died
Life expectancy of mother	Expected remaining lifetime of the respondent's mother constructed using the age-gender-race specific life table survival probabilities	---
Female	Respondent is female	Respondent is male
White	Respondent is White/Caucasian	Respondent's race is different from White/Caucasian and Black/African American
Black	Respondent is Black/African American	Respondent's race is different from White/Caucasian and Black/African American
Education low	Respondent's years of education is less than 12 years	Respondent's years of education is between 11 and 16 years
Education high	Respondent's years of education is more than 15 years	Respondent's years of education is between 11 and 16 years
Age	Respondent's age at the end of interview (in years)	---
Income in $10,000	The sum of respondent's annual wage/salary income, bonuses/overtime, pay/commissions/tips, 2nd job or military reserve earnings, professional practice or trade income in nominal dollars	---

Health	Health problem index based on diagnosis of high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric problems, and arthritis by the respondent`s doctor	---
Net wealth in $10,000	The sum of all wealth components except the value of IRAs and Keogh plans less all debt in nominal dollars	---
Quartile low (25 %)	Respondent's net wealth is lower than and equal to the 25th quartile of the net wealth series (<=$40,000)	Respondent's net wealth is between the 25th and 75th quartiles of net wealth series (>$40,000 and <$303,000)
Quartile high (25 %)	Respondent's net wealth is higher than and equal to the 75th quartile of the net wealth series (>=$303,000)	Respondent's net wealth is between the 25th and 75th quartiles of net wealth series (>$40,000 and <$303,000)
Household size	The number of residents in the household including the respondent and the spouse	---
Full-time	Respondent works full-time	---
Unemployed	Respondent is unemployed	Respondent works full-time
Retired	Respondent is retired	Respondent works full-time
Disabled	Respondent is disabled	Respondent works full-time
Other Labor Force Status (OLFS)	OLFS=1-fulltime-unemployed-retired-disabled All labor force categories are working full-time,working part-time, unemployed, partly retired, retired, disabled, or not in the labor force	Respondent works full-time
Married	Respondent is married	Respondent is never married
Never married	Respondent is never married	---
Other Marital Status (OMS)	OMS=1-married-nevermarried All marital status categories include married, married but spouse is absent, partnered, separated/divorced, widowed, and never married	Respondent is never married

Table 1.3: CES-D Frequencies (%)

Waves	1994	1996	1998	2000	2002	2004	2006
No.Obs.	8,704	8,207	12,096	11,087	10,466	12,565	11,979
0	51.23	50.85	41.23	42.41	45.57	46.31	45.54
1	20.49	20.45	23.63	22.98	21.03	20.84	20.9
2	8.61	9.71	12.58	12.07	11.07	11.22	11.18
3	5.61	6.19	7.22	7.57	7.21	6.78	6.91
4	3.96	4.06	5.22	4.78	4.89	4.43	4.48
5	3.09	2.98	3.57	4.03	3.62	3.51	3.87
6	2.93	2.71	3.03	2.96	3	2.98	3.19
7	2.37	1.96	2.36	2.17	2.32	2.42	2.48
8	1.7	1.09	1.17	1.03	1.3	1.5	1.45

Table 1.3 shows the distribution of the CES-D score over nine possible outcomes for each subsequent wave. From 1994 to 2006, the number of observations increases by approximately 38 percent. In each wave, more than 40 percent of the respondents report the lowest CES-D category. The number of respondents who has the highest CES-D score is quite low across waves.

1.4 METHODOLOGY

The "psychological well-being'' or "mental health" equation is estimated using both linear (OLS) and nonlinear (Logit) models with fixed effects. The former assumes the cardinality of happiness scores whereas the latter is designed for the ordinal concept of happiness answers.[10] According to Ferrer-i-Carbonell and Frijters (2004), the assumption of both ordinality and cardinality of happiness scores leads to similar results if one estimates the model with fixed

[10] Usually, happiness/life satisfaction questions range from 0 to 10 where 0 means totally unhappy/dissatisfied and 10 means totally happy/satisfied. The cardinality assumption implies that the relative difference between happiness/satisfaction scores has a quantitative meaning which makes scores comparable across individuals. The ordinality assumption says that the relative difference in those scores is unknown, what matters is the ordering of the scores and individuals share the same interpretation about the ordering. In our study, the CES-D score includes one item about happiness which is answered 'yes' or 'no'. Although we do not directly estimate the happiness equation, we follow the methodology in happiness literature since one of the underlying variables in the CES-D score is happiness.

effects. When fixed effects are dropped from the well-being equation, coefficient estimates as well as standard errors differ substantially across linear and non-linear models.[11]

Under the linear FE specification, the mental health equation is estimated as following:

$$S_{it} = \alpha_i + x_{it}^{'}\beta + \varepsilon_{it} \quad i = 1,.....,N \text{ and } t = 1,......,T \tag{1.1}$$

where S_{it} denotes the CES-D score, x_{it} is a vector of explanatory variables which include life expectancy of mothers and a set of controls such as age, job status, marital status, years of education, health status of individuals, etc., β is a vector of homogenous coefficient across cross-sections, α_i are fixed effects, and ε_{it} are the error terms. The fixed effects term captures the effect of unobserved time-invariant characteristics such as personality traits or genes on the CES-D score which consists of eight subjective questions. In order to use the standard FE estimator, it is assumed that the explanatory variables in all periods are unrelated to the error terms in all periods (strict exogeneity).

For non-linear model estimation, an ordered logit model is appropriate since the CES-D score is an ordinal measure and represented by a categorical variable. In fact, FE ordered logit models such as Das and Van Soest (1999) and Ferrer-i-Carbonell and Frijters (2004) estimators are the most suitable ones to estimate the non-linear mental health equation with fixed effects. On the other hand, Winkelmann and Winkelmann (1998) proposed to transform the psychological well-being variable into a binary variable in order to incorporate fixed effects to the estimation. Although the resulting binary estimator is consistent, it is not efficient as some information is lost at the transformation stage. In order to use the estimator suggested by Winkelmann and Winkelmann (1998), the original CES-D score, which takes values from 0 to 8, is transformed into a binary variable. The mean of the CES-D score is equal to 1.44 as it is reported in the Table 1.1. The CES-D score is coded as 1 if it is below 2 and as 0 otherwise. That is, individuals who report below-average are treated as mentally healthy and individuals who report above-average

[11] For example, Ferrer-i-Carbonell and Frijters (2004) show that when fixed effects are allowed, the influence of income on well-being is reduced by 2/3 and the effect of having children is found to be insignificant, although the corresponding coefficient is negative and significant in the pooled estimations. The authors conclude that the effect of other variables used in the literature will turn out to be different once fixed effects are accounted for. They deduce that this will be the case in both individual-level and national-level studies.

are treated as mentally unhealthy. Note that higher value of the binary variable implies higher mental well-being, whereas higher CES-D score implies lower mental health level.

Under the non-linear FE specification, the following latent model is adopted:

$$S_{it}^{*} = \alpha_i + x_{it}^{'}\beta + \varepsilon_{it} \quad i = 1,\ldots\ldots, N \text{ and } t = 1,\ldots\ldots, T \tag{2.1}$$

where S_{it}^{*} is an unobserved continuous index of mental health, x_{it} is a vector of explanatory variables which include life expectancy of mothers and a set of controls such as age, job status, marital status, years of education, health status of individuals, etc., β is a vector of homogenous coefficient across cross-sections, α_i are fixed effects, and ε_{it} are the error terms.

The observed binary variable is defined as follows:

$$S_{it} = \begin{cases} 1 \text{ if } S_{it}^{*} > 0 \\ \\ 0 \text{ if } S_{it}^{*} \leq 0 \end{cases} \tag{3.1}$$

The ε_{it} 's are assumed to be independently identically distributed (i.i.d) with a logistic distribution and they are independent of $x_{i1},\ldots\ldots, x_{iT}$. In a panel data model with fixed effects, the number of parameters to be estimated increases with the number of individuals in the panel. This leads to inconsistent maximum likelihood estimates for α_i and β's if the time-span of the sample is finite. In order to solve the so-called incidental parameter problem, and to find consistent estimates for β's, Chamberlain (1980) proposes a conditional maximum likelihood estimator. The probability of observing $S_{i1},\ldots. S_{iT}$ conditional on $\sum_{t=1}^{T} S_{it}$ can be written as:

$$P\left(S_{i1},\ldots. S_{iT} \middle| x_{i1},\ldots.. x_{iT}, \alpha_i, \sum_{t=1}^{T} S_{it} \right) = \frac{\exp\left[\sum_{t=1}^{T} \left(x_{it}^{'}\beta \right) S_{it} \right]}{\sum_{d \in D_i} \exp\left[\sum_{t=1}^{T} \left(x_{it}^{'}\beta \right) d_t \right]} \tag{3.1}$$

where D_i denotes the set of all the possible combinations of $S_{i1},......,S_{iT}$ which sum up to $\sum_{t=1}^{T} S_{it}$.

In the case of T=2, the likelihood equals to $\frac{\exp(S_{i1}x_{i1} + S_{i2}x_{i2})\beta}{\exp(x_{i1}\beta) + \exp(x_{i2}\beta)}$ which is independent of α_i 's and only includes individuals for whom $S_{i1} + S_{i2} = 1$. This means that the individuals whose responses are invariant over time are excluded in the first step and the model is estimated using the individuals whose responses are changing over time.

In both linear and non-linear panel data models, if the individual effects are independent of $x_{i1},......,x_{iT}$, $\varepsilon_{i1},......,\varepsilon_{iT}$ and $\alpha_i \sim N(0,\sigma_\alpha^2)$ a random effects estimator can be utilized since it leads to consistent and asymptotically efficient estimates under these assumptions. In practice, these assumptions are likely to be violated most of the time so that random effects models might lead to inconsistent estimates. In order to provide a full and robust analysis, the linear model in (1.1) is also estimated with random effects. Moreover, the null hypothesis that individual effects are independent of explanatory variables is tested using the so-called Hausman test.[12]

The presence of individual effects in a binary Logit model is tested by Winkelmann and Winkelmann (1998) by means of the Hausman test. The FE Logit estimation is a conditional maximum likelihood method whereas pooled Logit estimation is obtained using is a maximum likelihood estimator. Following the latter study, the Hausman test is used to check the significance of fixed effects in the mental health equation in (2.1). Under the null hypothesis, the test statistic is:

$$H = (\hat{\beta}_{CML} - \hat{\beta}_{ML})' (\hat{V}_{CML} - \hat{V}_{ML})^{-1} (\hat{\beta}_{CML} - \hat{\beta}_{ML}) \sim \chi_k^2$$

where k is the number of explanatory variables. The Hausman test statistic is used to compare the pooled logit and FE logit estimators since the latter is consistent under the null and alternative hypotheses whereas the former is consistent and asymptotically efficient under the null but it is inconsistent under the alternative hypothesis.[13]

[12] Results are available in Appendix A (Table 2A and Table 4A).
[13] Results are given in Appendix A (Table 3A and Table 5A)

1.5 EMPIRICAL RESULTS

This section presents estimation results for linear and non-linear models described in equation (1.1) and in equation (2.1), respectively. As shown in Appendix A (see footnote 12 and 13) the Hausman test statistic suggests that both random effects and pooled Logit estimators give inconsistent estimates. Therefore, the reported coefficients as well as the marginal effects are based on linear and non-linear models with fixed effects.[14]

Table 1.4: Fixed Effects Estimation

	(1) FE CES-D	(2) FE logit CES-D binary	(3) FE CES-D	(4) FE logit CES-D binary
Age	-0.0444***	0.0635**	-0.0575***	0.0752**
	(0.0148)	(0.0303)	(0.0156)	(0.0320)
Age squared	0.000297**	-0.000502**	0.000389***	-0.000587**
	(0.000116)	(0.000237)	(0.000121)	(0.000247)
Mother alive	**-0.0515***	**0.102***		
	(0.0265)	**(0.0552)**		
Ln(income)	-0.0383***	0.0534***	-0.0373***	0.0509***
	(0.00954)	(0.0194)	(0.00954)	(0.0194)
Health	0.144***	-0.218***	0.144***	-0.217***
	(0.0111)	(0.0226)	(0.0111)	(0.0226)
Quartile low (25%)	0.0350	0.00735	0.0355	0.00536
	(0.0225)	(0.0438)	(0.0225)	(0.0439)
Quartile high (25%)	-0.0113	0.0298	-0.0122	0.0314

[14] In the fixed effect models the variables such as female, white, black, years of education and the constant term have been dropped since they are time invariant. These variables were included to the random effects and pooled Logit estimations while performing the Hausman test statistic.

	(0.0213)	(0.0457)	(0.0213)	(0.0457)
Household size	0.00941	-0.0211	0.00948	-0.0203
	(0.00831)	(0.0165)	(0.00831)	(0.0165)
OLFS	0.0258	0.0123	0.0265	0.0114
	(0.0213)	(0.0439)	(0.0213)	(0.0440)
Unemployed	0.419***	-0.505***	0.421***	-0.504***
	(0.0605)	(0.115)	(0.0605)	(0.115)
Retired	0.123***	-0.132***	0.125***	-0.135***
	(0.0224)	(0.0457)	(0.0224)	(0.0457)
Disabled	0.482***	-0.503***	0.483***	-0.503***
	(0.0434)	(0.0864)	(0.0434)	(0.0865)
Married	-0.736***	0.981***	-0.739***	0.984***
	(0.125)	(0.233)	(0.125)	(0.234)
OMS	-0.101	0.159	-0.101	0.159
	(0.121)	(0.226)	(0.121)	(0.226)
Life expectancy of Mother			**-0.0105*** **	**0.0128* **
			(0.00326)	**(0.00665)**
Constant	3.658***		4.118***	
	(0.499)		(0.530)	
N	75104	37607	74984	37547

Table 1.4 gives coefficient estimates and the corresponding standard errors for four different models. The dependent variable in models (1) and (3) is the CES-D score whereas the dependent variable in models (2) and (4) is the binary indicator of the CES-D score. As noted before, higher value of the binary variable implies higher mental well-being, whereas higher CES-D score implies lower mental health level. As a result, the same variables in models (1) and (3) have opposite signs in models (2) and (4), respectively.

Before looking at the relation between parental longevity and psychological well-being, it would be informative to check if respondents whose parents are still alive report lower CES-D scores (less depressive symptoms) compared to those who have lost their parents. Therefore, we first focus on the results reported in columns (1) and (2) in the Table 1.4.[15] According to Ballas and Dorling (2007), the death of a parent has a negative impact on happiness level of individuals. Accordingly, one may expect that the presence of parents is positively associated with mental well-being of their children. Therefore, the expected sign of the variable "Mother alive" is negative and positive in models (1) and (2), respectively. The estimation results suggest that the coefficient of "Mother alive" has an expected sign and it is statistically significant at 10 percent significance level. The reported standard errors are higher in column (2) than those reported in column (1) since the FE Logit estimator is inefficient as discussed before. According to these results, on average, respondents whose mother is alive experience better mental health than those who have lost their mother in the past.

Table 1.5: Marginal Effects after the Fixed Effects Logit Estimation

	(1) FE logit CES-D binary		(2) FE logit CES-D binary
Age	0.0001	Age	0.0003
Mother alive	**0.0224**	**Life expcectancy of mother**	0.0028
Ln(income)	0.0117	Ln(income)	0.0112
Health	-0.0479	Health	-0.0477
Unemployed	-0.1110	Unemployed	-0.1108
Retired	-0.0290	Retired	-0.0297
Disabled	-0.1106	Disabled	-0.1106
Married	0.2156	Married	0.2163

(1): Model 2 in the Table 1.4, (2): Model 4 in the Table 1.4

In order to demonstrate how much the probability of being mentally healthy increases with the presence of mothers, we calculate the marginal effect for an average respondent after the FE

[15] Instead of using the variable 'Mother alive', we created a dummy variable which takes 1 if the respondent's father is alive, and re-estimated the models (1) and (2) in the Table 1.4. Results showed that the respondents whose father is alive do not significantly report less depressive symptoms than those who have lost their fathers.

Logit estimation. The probability of being mentally healthy is equal to 0.674 for the average respondent in our sample, that is, $P(S_{it} = 1|x_{it}, \alpha_i) = 0.674$. The marginal effect formula is:

$$\frac{\partial P(S_{it} = 1|x_{it}, \alpha_i)}{\partial x_{it,3}} = \beta_3 \{P(S_{it} = 1|x_{it}, \alpha_i)(1 - P(S_{it} = 1|x_{it}, \alpha_i))\} = 0.0224$$

An average respondent whose mother is alive is 2.24 % - points more likely to be mentally healthy compared to an average respondent whose mother has died. Based on this finding, it is reasonable to analyze whether a respondent becomes less depressed (or mentally healthier) as his mother's remaining lifetime gets longer.

Now we return to the relation between parental longevity and psychological well-being. As explained before, the main variable of interest throughout the analysis is remaining lifetime of respondents' mothers which is denoted by the variable 'Life expectancy of Mother' in the Table 1.4. Columns (3) and (4) give the FE and FE Logit estimation results for the effect of parental longevity on mental health. Since the presence of mothers is positively associated with mental health levels of their children, one may expect that respondents become less depressed as they live longer with their mothers. Therefore, the expected sign of the variable "Life expectancy of Mother" is negative and positive in models (3) and (4), respectively.

Table 1.4 shows that the coefficient of "Life expectancy of Mother" has an expected sign and it is statistically significant at 1 percent significance level in the linear model with FE (column 3). As expected, the precision of this variable becomes smaller if one uses inefficient FE logit estimator (column 4). On average, respondents experience less mental distress as they expect to live longer with their mothers. The calculated marginal effect in the Table 1.5 also suggests that the probability of being mentally healthy increases by 0.28 % - points for each additional year the respondent expects to live with his/her mother.

The marginal effects of other significant control variables are calculated in a similar way and summarized in Table 1.5. According to the Table 1.5, being married has the highest positive effect whereas being unemployed has the highest negative effect on probability of being mentally healthy. On average, life expectancy of mother variable has a relatively small effect compared to the impacts of other statistically significant determinants.

Overall, respondents' mental health levels are positively associated with expected remaining lifetime of their mothers. Considering the exogenous nature of life expectancy variable, this finding suggests that there is a causal relationship between longevity expectations and current psychological well-being of individuals.

1.6 CONCLUSION

The first chapter of this thesis mainly investigates whether there is a link between psychological well-being and survival expectations. Using an exogenous proxy for life expectancy, we observe a causal relationship between current psychological well-being of HRS respondents and their expectations regarding parental longevity. This finding supports the claim that individuals estimate the number of years that they will live with the people in their social environment. It is also consistent with previous research which shows that older workers anticipate their far future and they report lower happiness scores due to adverse implications of future pension reforms (De Grip et al, 2009).

Since the similarity of the predictions concerning one`s own life expectancy and life-table information is well-documented, one`s prediction concerning his/her mother remaining lifetime might also be in line with the life-table probabilities. The analysis throughout this thesis relies on the assumption that respondents' predictions concerning their own survival probabilities and their mother`s survival probabilities are accurate and, therefore, they coincide with the information in the life-tables. However, to our knowledge, there are no studies which explore the accuracy of the later prediction and this is in part due to lack of relevant survey data.

The sensitivity of psychological well-being to longevity expectations may indicate that older workers anticipate future events and they might take their predictions in consideration when they make economic decisions. The predictions about life expectancy might particularly affect the choice of retirement timing and leisure preferences of older workers. Therefore, in the second part of this thesis we investigate the link between longevity expectations and retirement behavior.

CHAPTER II: RETIREMENT BEHAVIOR AND LONGEVITY EXPECTATIONS

2.1 INTRODUCTION

The social security budgets of many industrialized countries are under strain due to the demographic shift towards older population. Since the 1980s, life expectancy in those countries has increased significantly due to the improved health care facilities. For example, over this period, the gain in life expectancy at birth in the USA is 6 years for black males, 4.2 years for black females, 5 years for white males, and 2.5 years for white females.[16] Since 1990, every OECD country has adopted social security reforms with the aim of alleviating the problem of aging population. Retirement age and life expectancy are closely interrelated. The legal retirement age in USA gradually has increased to from age 65 to age 67 and, currently, those who were born in 1960 and later can receive full Social Security benefits at age 67.

The second chapter of this thesis investigates whether individuals' retirement decision is sensitive to their longevity expectations. Theoretically, life cycle models (LCM), in which individuals make consumption and leisure choices over their lifetimes, predict that individuals defer retirement as their expected lifespan extends. The reason behind this prediction is that with a longer life, individuals need to work longer in order to maintain their consumption in old age. Empirically, several researchers have investigated the relationship between retirement decision and longevity using subjective survival probabilities as a proxy for individuals' life expectancy. With the availability of survey questions about survival probabilities, researchers often prefer subjective probabilities to their objective counterparts (life-table probabilities) in order to explain individual decision making properly. In cross- sections the subjective survival probabilities aggregate well to life table levels although they show higher variation than average population probabilities at the individual level.

The studies which test the validity of subjective survival probabilities suggest that these probabilities may contain serious measurement error possibly due to unfamiliarity of respondents with probabilities. In fact, Hurd and McGarry (1995) show that 2.5 percent of the respondents reported larger values for survival probability to age 85 than for the survival probability to age 75. Another study, Hurd, et al (1998), demonstrates the tendency of respondents to provide focal-

[16] See National Vital Statistics Reports for years 1980 and 2006.

point answers (0, 0.5, or 1) to survival probability questions. In order to correct for the potential measurement error in these measures, researchers employ instrumental variable (IV) techniques with the aim of finding valid instruments for subjective survival probabilities (Bloom et al 2006, and O'Donnell et al 2008).

Although the relation between survival expectations and retirement decision is investigated before, to our knowledge; no study uses life table survival probabilities as a covariate in the retirement equation. Therefore, we aim to extend the findings of previous literature by using an exogeneous proxy for respondents' life expectancy. To this end, we again use eight waves of RAND HRS database covering the years from 1992 to 2006.

The structure of this chapter is organized as follows: Section 2.2 presents related studies in the literature. Data sources are explained in Section 2.3. The methodology is introduced in Section 2.4. Estimation results are presented in Section 2.5. Finally, Section 2.6 gives concluding remarks.

2.2 RELATED LITERATURE

The relationship between retirement behavior and longevity has been previously investigated both theoretical and empirical perspectives. According to theory, in a simple LCM with one-period labor-leisure choice, the consumer's maximization problem is to choose lifetime consumption and leisure over his/her constraint so that his/her lifetime utility is maximized (Chang, 1991). In such a model, an increase in the consumer's lifetime results in a parallel shift in consumer's budget set provided that he/she earns the same amount of money in those extra years. The income effect increases both consumption and leisure under the assumption that consumption and leisure are both normal goods. Increased consumption means increased labor supply since extra consumption is financed through working. As a result, individuals delay retirement.

Kalemli-Ozcan and Weil (2007) obtain the same result in a richer specification. In their model, individuals maximize their utility subject to their budget constraint by choosing a path of lifetime consumption and endogenous retirement age with two different assumptions about uncertainty. In the absence of the uncertainty effect, that is to say, when individuals are certain about their date of death, increased life expectancy implies higher retirement age since extra years of consumption is financed by working. Kalemli-Ozcan and Weil (2007) also report that, according to life-tables of US males for 1900-2000, conditional of having reached age 20, expected life expectancy increases, whereas the variance of life expectancy decreases. Decreased variance of life expectancy has an opposite effect on retirement age. In life cycle models, in the case of increased uncertainty and under the assumption of incomplete markets, the agent may have higher wealth when he passes away. As the effective return on savings decreases due to increased uncertainty, individuals tend to delay their retirement in order to finance their old age consumption. Therefore, higher expected life expectancy leads to delayed retirement, whereas lower uncertainty that accompany higher expected life expectancy leads to earlier retirement. In order to capture the possible non-linear relationship between life expectancy and retirement age we add higher order life expectancy terms to our set of covariates.

Empirically, two previous studies for the USA are much related to our research question. Firstly, Hurd et al. (2004) investigate the effect of subjective survival on retirement and Social Security claiming using four waves of HRS. They find that there is no relationship between life expectancy and retirement before age 62. Beginning at age 62, individuals with zero probability of survival go into retirement with a higher rate compared to those with moderate or high survival probabilities and this trend continues after age 62 as well. The study of Bloom et al (2006) utilizes subjective survival expectations from HRS to explain retirement and saving rates separately. To correct for the measurement error, focal points in responses and possible reverse causality regarding subjective survival expectations, they report the IV Probit estimates with an instrument of parental longevity. Their findings suggest no negative relationship between higher survival probabilities and the probability of retirement. Another study by O'Donnell et al (2008) investigates the theoretically suggested relation between retirement and longevity using three waves of the English Longitudinal Study Ageing (ELSA). Using the IV Probit estimator, they obtain a concave relationship between mortality expectations and retirement. According to their results, individuals with very low survival probabilities are least likely to retire; as individuals become more optimistic, the probability of retirement increases and then it falls for most of the distribution of survival expectations.

2.3 DATA

As it is in the first part of this thesis, the data is drawn from nine waves of the RAND HRS database which represents the population of Americans over the age 50. The database is well-suited for our purpose, since it includes information about retirement status, health status, income, and wealth of the respondents and it collects data over time.

The measure for the dependent variable is a binary retirement variable which is based on the survey respondent's description of his or her job status. The value of dependent variable takes 1 if the respondent reports that he/she is completely or partly retired and it takes 0 if he is not retired yet (neither completely nor partly). This thesis analyzes the retirement behavior of Americans over the age 50 as a function of their life expectancy and a set of control variables. For this purpose, we model the probability of retirement between subsequent waves of the HRS panel. To be included in the sample, respondents must report that they are non-retired (either completely or partly) in the first year they enter the RAND HRS survey. Therefore, we exclude individuals who are already retired in the first year they attend the survey. And then, we follow the remaining individuals until they exit from non-retired state into retired state (until they report that they are either completely or partly retired). Once they are retired, they are dropped from the sample. As a result, the final sample has up to 8 observations per person (an individual is used once if he/she reports retirement in the second wave, an individual is used twice if he/she is retired by the third wave, etc.).

The life expectancy of respondents is inferred from race, gender, and age specific life-tables of National Vital Statistics Reports which cover the years from 1992 to 2006. For the reasons discussed in section 1.2, respondents who were born before 1924 (AHEAD cohort) are dropped from the sample. The 2008 wave is also dropped due to unavailability of life table statistics for that year. In summary, empirical analyses are based on a panel sample which covers the years from 1992 to 2006 and includes Americans born between 1924 and 1953.

Table 2.1 : Descriptive Statistics of the Variables Used in the Analysis

	N	*n*	Mean	Std. Dev.	Min	Max
Retired	31,964	11,118	0.05	0.21	0	1
Life expectancy	31,964	11,118	24.07	4.93	7.3	43.2
Education low	31,964	11,118	0.18	0.39	0	1
Education high	31,964	11,118	0.24	0.43	0	1
Health	31,964	11,118	1.05	1.05	0	7
Tenure	31,964	11,118	13.45	11.30	0	66.8
Wage in $1,000 per week	31,964	11,118	0.74	1.76	0	124.45
Black	31,964	11,118	0.14	0.35	0	1
White	31,964	11,118	0.82	0.39	0	1
Male	31,964	11,118	0.44	0.50	0	1
Wealth in $10,000	31,964	11,118	26.60	117.15	-473.3	10039
Quartile low (25 %)	31,964	11,118	0.23	0.42	0	1
Quartile high (25 %)	31,964	11,118	0.23	0.42	0	1
Couple	31,964	11,118	0.77	0.42	0	1
Age	31,964	11,118	57.99	6.01	38	81

Table 2.2: Variable definitions

Variable	Description	Baseline
Retired	Respondent is completely or partly retired	Respondent is not retired (neither completely nor partly retired)
Life expectancy	Expected remaining lifetime of the respondent constructed using the age-gender-race specific life table survival probabilities	---
Male	Respondent is male	Respondent is female
White	Respondent is White/Caucasian	Respondent's race is different from White/Caucasian and Black/African American
Black	Respondent is Black/African American	Respondent's race is different from White/Caucasian and Black/African American

Education low	Respondent's years of education is less than 12 years	Respondent's years of education is between 11 and 16 years
Education high	Respondent's years of education is more than 15 years	Respondent's years of education is between 11 and 16 years
Age	Respondent's age at the end of interview (in years)	---
Wage in $1000	Respondent's weekly wage rate in nominal dollars	---
Health	Health problem index based on diagnosis of high blood pressure, diabetes, cancer, lung disease, heart disease, stroke, psychiatric problems, and arthritis by the respondent`s doctor	---
Net wealth in $10,000	The sum of all wealth components except the value of IRAs and Keogh plans less all debt in nominal dollars	---
Quartile low (25 %)	Respondent's net wealth is lower than and equal to the 25th quartile of the net wealth series (<=$36,000)	Respondent's net wealth is between the 25th and 75th quartiles of net wealth series (>$36,000 and <$262,000)
Quartile high (25 %)	Respondent's net wealth is higher than and equal to the 75th quartile of the net wealth series (>=$262,000)	Respondent's net wealth is between the 25th and 75th quartiles of net wealth series (>$36,000 and <$262,000)
Tenure	Respondent's years of tenure on the current job	---
Couple	Respondent lives with a partner/spouse	Respondent is single
Age51,...,Age75	A set of dummies for age	Respondent's age is 51
Other age	1-age51-age52-age53-......-age75	Respondent's age is 51

Table 2.1 presents descriptive statistics of relevant variables used in this part of the thesis. The corresponding variable definitions and baseline categories for all set of dummies are reported in Table 2.2. The big N and the small n in the Table 2.1 stand for the overall sample size and the number of respondents who are observed over time, respectively. Overall, the sample consists of 31,964 person-year observations. It is an unbalanced sample since the average number of waves in which a respondent was observed is approximately 3 instead of 8.

Table 2.1 indicates that 5 % of the respondents exited from non-retired state into retired state during the whole sample period. The expected remaining lifetime (life expectancy) of respondents ranges from 7 to 43 years with a mean of 24 years. Respondents' tenure on the current job is equal to 13 years on average.[17] The mean of health index suggests that, on average, respondents in the sample are healthy since the number of diseases they have is approximately 1 out of 7 diseases. Weekly wage rate measured in $1000 has zero values for some respondents.[18] The net wealth denoted in $10,000 has both negative and positive values for the respondents so that it is impossible to present this variable in terms of the natural logarithm. Therefore, we again create three wealth categories as in the first section. These categories are low (quartile low (25%)), medium (the second and the third quartiles (50%)), and high (quartile high (25%)).

44 % of the sample consists of males which suggests that men are slightly under-represented compared to women in the sample. About 77 % of respondents are couples. In terms of race, the sample consists of white individuals mostly. Only 18 % of the respondents have a lower degree than high school degree which shows that the sample is relatively well-educated.

2.4 METHODOLOGY

Following the literature, the retirement equation is estimated using a pooled Probit estimator (Hurd, et al 2004, Bloom et al 2006, and O'Donnell et al 2008). This estimator is a standard Probit estimator which ignores unobserved individual effects in a panel data estimation. As it is described in the section 2.3, in order to model the probability of retirement between subsequent waves, the sample is arranged in such a way that individuals have up to 8 observations. That is,

[17] Although it is not reported, only 0.43 % of the respondents do not have tenure at all.
[18] Only 0.67 % of the respondents have zero weekly wages.

an individual is used once if he/she reports retirement in the second wave, an individual is used twice if he/she is retired by the third wave, etc. Therefore, number of years that an individual is observed is limited in the sample. In this case, the pooled Probit estimator can be more appropriate to model retirement transitions compared to FE logit estimator which uses within-individual variation.[19]

Under the pooled Probit specification, the following latent model is estimated:

$$R_{it}^{*} = \alpha + x_{it}^{'}\beta + \varepsilon_{it} \quad i = 1,.....,N \text{ and } t = 1,......,T \tag{2.1}$$

where R_{it}^{*} is an unobserved continuous index of retirement, x_{it} is a vector of explanatory variables which include life expectancy of respondents and a set of controls such as age, tenure, income, wealth, years of education, health status of respondents, etc., β is a vector of homogenous coefficients across cross-sections, and ε_{it} are the error terms.

The observed binary retirement variable is defined as follows:

$$R_{it} = \begin{cases} 1 & \text{if } R_{it}^{*} > 0 \\ \\ 0 & \text{if } R_{it}^{*} \leq 0 \end{cases} \tag{2.2}$$

The ε_{it} 's are assumed to be independent of $x_{i1},......,x_{iT}$. They are also standard normally distributed, i.e, $\varepsilon_{it} \sim N(0,1)$ [20]

The marginal effect for the pooled Probit model is given in the following formula:

[19] In order to investigate whether unobserved fixed effects are important in the retirement equation, we use the whole sample (without dropping the individuals who are already retired in the first year they attend the survey). The dependent variable is again the self-reported retirement (binary variable).We estimate FE logit and RE logit models and compare both models by means of the Hausman test. Estimation results are reported in Appendix B, in Table 1B.

[20] Alternatively, one might use pooled Logit estimator which is obtained when error terms follow a logistic distribution. The variance of the error terms is 1 and $\pi^2/3 \approx 3.29$ in Probit and Logit models, respectively. Logistic distribution has fatter tails compared to standard normal distribution. Since the variance of the Logit model is higher than that of Probit, slope coefficients of former are $\sqrt{\pi^2/3} = 1.81$ times higher than those of latter.

$$\frac{\partial \Pr(R_{it} = 1|x_{it})}{\partial x_{it,j}} = f_{\varepsilon_{it}}(x_{it}^{'}\beta)\beta_j$$

where $f_{\varepsilon_{it}}(x_{it}^{'}\beta)$ is the density of ε_{it}. The marginal effect depends on the individual characteristics x_{it} because $f_{\varepsilon_{it}}(x_{it}^{'}\beta)$ varies across individuals. Therefore, either marginal effects are calculated for each individual in the sample and then the mean of these marginal effects are reported or they are reported for an individual with average characteristics (at the mean values of explanatory variables).

In order to incorporate unobserved heterogeneity into the retirement equation in (2.1), we re-estimate the model using RE Probit estimator. For the latter, the assumption that the individual effects are independent of $x_{i1},......,x_{iT}$, $\varepsilon_{i1},......,\varepsilon_{iT}$ and $\alpha_i \sim N(0,\sigma_\alpha^2)$ is required. We also test if unobserved heterogeneity is important. For this purpose, we test whether the proportion of the total error variance contributed by the variance of individual effects, $\hat{\rho} = \frac{\hat{\sigma}_\alpha^2}{\hat{\sigma}_\alpha^2 + \hat{\sigma}_\varepsilon^2}$, is equal to zero in the RE Probit estimation. If we cannot reject that $\hat{\rho}$ is equal to zero then RE Probit coincides with Pooled Probit estimator.

2.5 EMPIRICAL RESULTS

This section presents estimation results for the pooled Probit and the RE Probit estimators for the retirement equation described in (2.1). The main variable of interest throughout the analysis is expected remaining lifetime of respondents which is denoted by the variable 'Life expectancy' in the Table 2.3. We add higher order-polynomials of life expectancy variable to the estimation in order to test the possible nonlinear relation between life expectancy and retirement decision, as it is suggested by the theory. All higher-order polynomial terms turn out be insignificant except for the square term. Results in the Table 2.3 shows that there is a statistically significant nonlinear relationship between life expectancy and the probability of retirement although the squared term is not highly significant. As Kalemli-Ozcan and Weil (2007) show in their theoretical model, increased life expectancy has two possible effects. First, increased horizon effect leads to delay in retirement. Second, decreased uncertainty about the length of life increases the effective return on savings so that it becomes attractive to retire early. Our estimation results suggest that there may be two different mechanisms through which life expectancy affects retirement behavior but the first effect seems to be more pronounced.

Table 2.3. Pooled Probit versus RE Probit Estimation, Coefficient Estimates

	(1) Pooled Probit	(2) RE Probit
Education high	0.203***	0.216***
	(0.0378)	(0.0399)
Education low	-0.114***	-0.123***
	(0.0382)	(0.0417)
Health	0.0832***	0.0882***
	(0.0126)	(0.0142)
Tenure	-0.0189***	-0.0201***
	(0.00159)	(0.00185)
Ln(wage)	-0.427***	-0.445***
	(0.0223)	(0.0232)
Black	-0.0141	-0.0180
	(0.0894)	(0.0914)
White	0.0730	0.0759
	(0.0736)	(0.0771)
Life expectancy	**-0.115***	**-0.120***
	(0.0415)	**(0.0449)**
Life expectancy^2	**0.00152***	**0.00157***
	(0.000776)	**(0.000898)**
Male	0.110*	0.111*
	(0.0620)	(0.0579)
Quartile low	-0.258***	-0.269***
(25 %)	(0.0396)	(0.0435)
Quartile high	0.313***	0.328***
(25 %)	(0.0350)	(0.0387)
Couple	-0.0262	-0.0273
	(0.0359)	(0.0381)
Constant	2.128***	2.262***
	(0.656)	(0.641)
$\hat{\rho}$		0.072
		(0.065)
LR test (p-value)		1.24
		(0.133)
N	31964	31964

Clustered standard errors are in parentheses, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, Age dummies are included but they are not reported. LR: Likelihood Ratio test with the null hypothesis that rho=0

In order to compare pooled Probit and RE Probit estimators, we test the null hypothesis that $\hat{\rho}$ is equal to zero using Likelihood Ratio test. The p-value of the LR test shows that the null hypothesis cannot be rejected at the 5 percent significant level which suggests unobserved heterogeneity in the our panel sample is unimportant. Therefore, both the RE Probit and the pooled Probit estimators give similar results. The dependent variable in both models above is the binary indicator for retirement as defined before. The probability of retirement is explained by life expectancy and a set of controls. Since the variable 'Life expectancy' is calculated from age-gender-race specific population life tables, it is possible this variable to be correlated with variables 'male, white and black, the set of age dummies' in the baseline specification. The age, gender, and race measures may have a direct impact on retirement behavior. If one of these variables is excluded from the baseline specification, the magnitude of life expectancy may change since the estimated effect of life expectancy contains the effect of the omitted variable on retirement. The Table 2.3 suggests that the coefficient of life expectancy is still significant after controlling for age, gender and race measures which implies a direct effect of longevity on retirement.

Table 2.4: Marginal Effects at the Mean after Pooled Probit Estimation

Education high	0.018***
Education low	-0.008***
Health	0.006***
Tenure	-0.001***
Ln(wage)	-0.033***
Life expectancy	**-0.0032*****
Male	0.008*
Quartile low (25 %)	-0.016***

Quartile high (25 %)	0.029***

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, age dummies are included but they are not reported.

Table 2.4 gives estimated marginal effects at the mean of the explanatory variables which are shown to be significant in the Table 2.3 based on the pooled Probit estimation. The estimated marginal effect of "Life expectancy" for an average respondent is negative and it is statistically significant at 1 percent significance level. If an average respondent expects to live one more year, the probability of retirement decreases by 0.32 % - points.

The marginal effects of control variables are also summarized in the Table 2.4. The amount of wage has the highest impact on the probability of retirement. Respondents who earn higher wages are more likely to delay retirement compared to those with lower wages. According to the LCM, when wage increases, the income effect leads to the purchase of more of all normal goods, including leisure, which makes the retirement more likely. The substitution effect, on the other hand, makes the probability of retirement less likely since higher wage increases the opportunity cost of leisure. The estimation results suggest that the substitution effect dominates income effect so that individuals prefer working as their wages increase.

Wealth dummies suggest that respondents whose wealth is less than and equal to $36,000 are less likely to retire compared to those whose wealth is more than $36,000 and less than $262,000. Similar comparison indicates that highest wealth owners are more likely to retire compared to the same reference group. Financial wealth may increase retirement rates, since richer retirees can maintain their pre-retirement consumption even though their income decreases substantially after retirement.

Health has a positive impact on transition to retirement. The probability of retirement is higher for males than females. Respondents with more years of tenure are more likely to delay retirement, on average. Interestingly, more educated individuals are more likely to retire compared to those with medium level of education and low educated respondents are less likely to retire compared to the same reference group. This result is against the previous findings of the literature, that is, as educated individuals tend to hold jobs that are not physically demanding, years of education may have a negative coefficient in the retirement regressions.

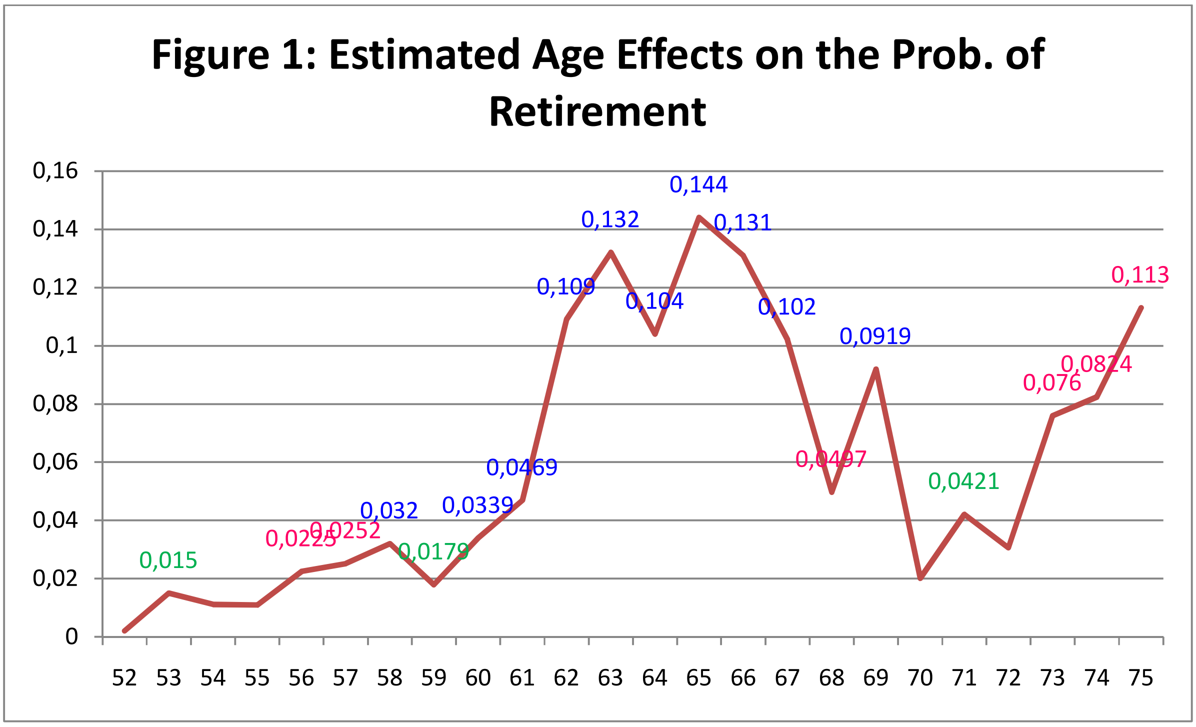

Figure 1: Estimated Age Effects on the Prob. of Retirement

Notes: Blue, pink and green mean significance at 1%, 5%, and 10 %, respectively.

Figure 1 shows the estimated marginal effects of age dummies on the probability of retirement. Starting from age 53, the set of dummies becomes significant which shows the direct effect of age on retirement. After the age 62, retirement seems to be more likely because respondents become eligible to receive Social Security benefits once they reach age 62.[21] The highest age effect on retirement occurs at age 65 which is the legal retirement age in the United States.

2.6 CONCLUSION

In the second chapter of this thesis, we examine whether retirement behavior is sensitive to longevity expectations using a panel survey of Americans over age 50 and their spouses. Similar to the first chapter, life expectancy is measured by life table survival probabilities which are not exposed to possible measurement errors and focal-point responses. Based on the findings of mortality literature we assume that respondents are aware of life table probabilities and their own expectations coincide with life table values.

After controlling for age, gender, and, race, we still observe a significant nonlinear effect of life expectancy on probability of retirement. Although this nonlinear relationship is not very strong,

[21] Reference age is 51.

it cannot be rejected. This estimated relation is consistent with implications of a LCM in which individuals are uncertain about their expected life spans. A linear relationship between life expectancy and retirement timing is also a theoretical possibility if we assume that there is no uncertainty about length of life. Overall, the estimated marginal effects suggest that if an average respondent expects to live one more year, the probability of retirement decreases by 0.32 % -points which is probably because more wealth is required to maintain consumption over an extended lifespan. This empirical finding may suggest that reforms which increase the legal retirement age are not ill-advised and future regulations about current retirement age can be expected in advance. Even in the absence of policy announcements, the working population may anticipate the upcoming policy reforms and revise their economic decisions in the light of these projections.

OVERALL CONCLUSION

Using a panel survey of elderly Americans and their spouses, we tried to answer two questions. First; to what extent is self-reported psychological well-being a forward-looking phenomenon? In order to estimate a causal relationship, we use life table survival probabilities of parents which are independent of optimism or pessimism of survey respondents when they answer survey questions. Based on the findings of mortality literature, we assume that respondents are aware of movements in today's life tables and they incorporate changes in life tables into their own survival probabilities and their mother`s survival probabilities. Our results for the first question suggest that, on average, respondents experience less mental distress as they expect to live longer with their mothers. In other words, the probability of being mentally healthy increases by 0.28 %-points for each additional year the respondent expects to live with his/her mother. This finding supports the claim that individuals estimate the number of years that they will live with the people in their social environment. It is also consistent with previous research which shows that older workers anticipate their far future and they report lower happiness scores due to adverse implications of future pension reforms.

The sensitivity of psychological well-being to longevity expectations may indicate that older workers anticipate future events and their predictions may play a role in their economic decisions. For example, increased life expectancy is closely related to retirement decision of older workers. Therefore, the second question; is retirement timing of respondents influenced by changes in life expectancy based on population life-tables? Estimation results show that there is a nonlinear relationship between life expectancy and probability of retirement. Higher expected life expectancy leads to delayed retirement, whereas lower uncertainty that accompany higher expected life expectancy leads to earlier retirement. On average, when respondents expect to live one more year, the probability of retirement decreases by 0.32 % - points which is probably because more wealth is required to maintain consumption over an extended lifespan.

References

Ballas, D., Dorling, D. 2007. “Measuring the Impact of Major Life Events upon Happiness” *International Journal of Epidemiology* 36(6): 1244-1252.

Bloom, D. E., Canning D., Moore, M. and Song, Y. 2006. “The Effect of Subjective Survival Probabilities on Retirement and Wealth in the United States’’*NBER Working Papers*, published: Clark, Robert, Naohiro Ogawa, and Andrew Mason (eds.) *Population Aging, Intergenerational Transfers and the Macroeconomy*. Cheltenham, U.K. and Northampton, MA: Elgar, 2007.

Charles, K. K. 2004. “Is Retirement Depressing?: Labor Force Inactivity and Psychological Well Being in Later Life." *Research in Labor Economics,* 23: 269-299.

Chamberlain, Gary. 1980. “Analysis of Covariance with Qualitative Data” *Review of Economic Studies*, 47: 225-238.

Chang, F-R. 1991. “Uncertain lifetimes, retirement and economic welfare”, *Economica*, 58: 215-232.

Clark, A. E., Frijters P. and Shields M.A. 2008a. “Relative Income, Happiness, and Utility: An Explanation for the Easterlin Paradox and Other Puzzles,” *Journal of Economic Literature*, 46(1): 95–144.

Clark, A. E., Diener, E., Georgellis, Y. and Lucas, R. E. 2008b. “Lags and Leads in life Satisfaction: A Test of the Baseline Hypothesis” *Economic Journal*, 118(529): F222-F243.

Das, M., and van Soest, A. 1999. “A panel data model for subjective information on household income growth,” *Journal of Economic Behavior and Organization*, 40: 409-426.

De Grip, A., Lindeboom, M. and Montizaan, R., 2009. “Shattered Dreams: The Effects of Changing the Pension System Late in the Game,” *IZA Discussion Papers* 4034, Institute for the Study of Labor (IZA).

Elder, T. E. 2007. “Subjective Survival Probabilities in the Health and Retirement Study: Systematic Biases and Predictive Validity,” *Working Paper WP 2007-159*, Michigan.

Ferrer-i-Carbonell, A. and Frijters, P. 2004. “How Important is Methodology for the Estimates of the Determinants of Happiness,” *Economic Journal*, 114(497): 641-659.

Finkelstein, A., Luttmer, E.F.P, Notowidigdo, M.J., 2008. “What good is wealth without health? The effect of health on the marginal utility of consumption,” *NBER Working Paper 14089*, Cambridge MA.

Frey, B.S. and Stutzer, A. 2005. “Happiness Research: State and Prospects,” *Review of Social Economy*, 63(2): 207-208.

Hamermesh, D. S. 1985, “Expectations, Life Expectancy, and Economic Behavior,” *Quarterly Journal of Economics*, 100: 389–408.

Hao, Y. 2008. “Productive Activities and Psychological Well-being Among Older Adults”. *Journal of Gerontology: Social Sciences*, 63B: S64– S72.

Hurd, M.D., Smith, J.P. and Zissimopoulos, J.M. 2004. “The Effects of Subjective Survival on Retirement and Social Security Claiming," *Journal of Applied Econometrics*, 19, 761-75.

Hurd, M.D. and K. McGarry. 1995. “Evaluating subjective probabilities of survival in the Health and Retirement Study”, *Journal of Human Resources*, 30 (Supplement), S268-S292.

Hurd, M.D., McFadden D. and Gan, L. 1998. “Subjective survival curves and life cycle behavior”, in David Wise (ed.), *Inquries in the Economics of Aging*, Chicago: University of Chicago Press.

Kalemli-Ozcan, S. and D.N. Weil 2007. “Mortality chage, the uncertainty effect and retirement”, *NBER Working Paper 8742*, Cambridge MA.

Kim, J. E. and Moen, P. 2002. “Retirement Transitions, Gender, and Psychological Well-being: A Life-course, Ecological Model.” *Journal of Gerontology: Social Sciences*, 57B: S212–S222.

Knight, J., Song, L., Gunatilaka, R. 2009. “Subjective Well-being and its Determinants in Rural China” *China Economic Review,* 20(4): 635-649.

Luttmer, E.F.P. 2005. “Neighbors as Negatives: Relative Earnings and Well-Being,” *The Quarterly Journal of Economics*, 120(3): 963-1002.

Montizaan, R. 2010. “De-motivating Workers: The Role of Negative Reciprocity” In R. Montizaan, eds., *Pension Rights, Human Capital Development and Well-Being.* Netspar Thesis, 73-90.

O’Donnell O, Teppa F, Van Doorslaer E. “Can subjective survival expectations explain retirement behavior?” Report No.: 188. Amsterdam: DNB, 2008.

OECD. Directorate for Employment, Labour and Social Affairs. 2009. *Pensions at a Glance 2009: Retirement-Income Systems in OECD Countries*

Post, T. and Hanewald, K. 2010. "Stochastic Mortality, Subjective Survival Expectations and Individual Saving Behaviour, " Available at SSRN: http://ssrn.com/abstract=1651167.

Schwarz, N. and Clore, G.L. 1983. "Mood, Misattribution, and Judgments of Well-being: Informative and Directive Functions of Affective States" *Journal of Personality and Social Psychology*, 45: 513 - 523.

Winkelmann, L. , Winkelmann R. 1998. "Why are the unemployed so unhappy? Evidence from panel data" *Economica*, 257, 1-15.

APPENDIX A

Table 1A: Expected remaining lifetime of the respondents' mothers, Frequencies

Total Number of Obs: 74,984

Values	Freq.	Values	Freq.
0	49,885	11.5	343
2.2	18	11.6	8
2.4	40	11.7	222
2.5	34	11.8	41
2.6	24	11.9	20
2.7	54	12	265
2.8	78	12.1	192
2.9	40	12.2	198
3	91	12.3	144
3.1	48	12.4	20
3.2	62	12.6	103
3.3	148	12.7	268
3.4	42	12.8	286
3.5	127	12.9	52
3.6	137	13	134
3.7	59	13.2	19
3.8	158	13.3	150
3.9	129	13.4	159
4	62	13.5	141
4.1	339	13.7	63
4.2	91	13.8	14
4.3	24	13.9	20
4.4	397	14	182
4.5	100	14.1	175

4.6	17	14.2	99
4.7	437	14.4	64
4.8	129	14.5	4
4.9	120	14.6	51
5	147	14.7	89
5.1	410	14.8	106
5.2	153	14.9	48
5.3	34	15.1	37
5.4	335	15.2	6
5.5	475	15.3	36
5.6	65	15.4	122
5.7	17	15.5	76
5.8	349	15.6	66
5.9	559	15.8	20
6	12	15.9	34
6.1	162	16	11
6.2	332	16.1	70
6.3	1,707	16.2	66
6.4	233	16.3	28
6.5	23	16.5	13
6.6	419	16.6	16
6.7	362	16.7	5
6.8	367	16.8	54
6.9	203	16.9	5
7	217	17	52
7.1	405	17.2	21
7.2	331	17.4	9
7.3	234	17.5	17
7.4	217	17.6	26
7.5	213	17.7	31
7.6	437	17.8	26
7.7	282	17.9	12
7.8	245	18	6
7.9	278	18.1	2
8	198	18.2	2
8.1	423	18.3	20
8.2	384	18.4	1
8.3	33	18.5	18
8.4	395	18.6	2
8.5	166	18.8	4
8.6	445	19	2
8.7	328	19.1	28
8.8	25	19.2	8

8.9	215	19.3	18
9	254	19.7	3
9.1	627	19.8	5
9.2	92	19.9	13
9.3	293	20	9
9.4	8	20.1	3
9.5	187	20.3	1
9.6	427	20.5	1
9.7	371	20.6	16
9.8	121	20.7	1
9.9	123	20.8	7
10.1	305	21.4	10
10.2	379	21.5	1
10.3	408	21.6	7
10.4	109	21.9	1
10.5	109	22.1	1
10.6	86	22.2	7
10.7	150	22.9	1
10.8	410	23	4
10.9	358	23.1	3
11	94	23.2	2
11.1	78	23.7	1
11.2	121	23.9	2
11.3	197	25	1
11.4	201	25.6	1

Table 2A: Fixed Effects versus Random Effects Estimation

Dependent variable: CES-D

Independent variable: Mother alive

	(1) FE	(2) RE
Age	-0.0444***	-0.0676***
	(0.0148)	(0.0127)
Age squared	0.000297**	0.000340***
	(0.000116)	(0.0000996)
Female	-	0.218***
		(0.0212)

Mother alive	**-0.0515***	**-0.0390****
	(0.0265)	**(0.0190)**
Ln(income)	-0.0383***	-0.0895***
	(0.00954)	(0.00823)
Health	0.144***	0.302***
	(0.0111)	(0.00663)
Quartile low (25%)	0.0350	0.203***
	(0.0225)	(0.0185)
Quartile high (25%)	-0.0113	-0.0393**
	(0.0213)	(0.0178)
Household size	0.00941	0.0269***
	(0.00831)	(0.00658)
OLFS	0.0258	0.0854***
	(0.0213)	(0.0185)
Unemployed	0.419***	0.511***
	(0.0605)	(0.0555)
Retired	0.123***	0.191***
	(0.0224)	(0.0192)
Disabled	0.482***	0.848***
	(0.0434)	(0.0376)
Married	-0.736***	-0.382***
	(0.125)	(0.0553)
OMS	-0.101	0.171***
	(0.121)	(0.0550)
White	-	-0.377***
		(0.0504)
Black	-	-0.400***
		(0.0553)
Education low	-	0.540***
		(0.0257)

Education high	-	-0.220***
		(0.0272)
Constant	3.658***	4.942***
	(0.499)	(0.421)
Hausman test (d.f.)		820.48 (14)
$\chi^2_{0.95,k}$		23.69
N	75104	75104

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3A: Fixed Effects Logit versus Pooled Logit Estimation

Dependent variable: CES-D Binary

Independent variable: Mother alive

	(1) FE logit	(2) Pooled logit
Age	0.0635**	0.115***
	(0.0303)	(0.0160)
Age squared	-0.000502**	-0.000731***
	(0.000237)	(0.000126)
Female	-	-0.241***
		(0.0183)
Mother alive	**0.102***	**-0.0114**
	(0.0552)	**(0.0210)**
Ln(income)	0.0534***	0.146***
	(0.0194)	(0.0110)
Health	-0.218***	-0.351***
	(0.0226)	(0.00680)
Quartile low (25%)	0.00735	-0.276***
	(0.0438)	(0.0218)

Quartile high (25%)	0.0298	0.102***
	(0.0457)	(0.0236)
Household size	-0.0211	-0.0402***
	(0.0165)	(0.00759)
OLFS	0.0123	-0.111***
	(0.0439)	(0.0247)
Unemployed	-0.505***	-0.612***
	(0.115)	(0.0777)
Retired	-0.132***	-0.258***
	(0.0457)	(0.0250)
Disabled	-0.503***	-1.147***
	(0.0864)	(0.0525)
Married	0.981***	0.238***
	(0.233)	(0.0509)
OMS	0.159	-0.241***
	(0.226)	(0.0504)
White	-	0.279***
		(0.0427)
Black	-	0.318***
		(0.0465)
Education low	-	-0.426***
		(0.0207)
Education high	-	0.314***
		(0.0250)
Constant	-	-4.289***
		(0.529)
Hausman test (d.f.)		284.44 (14)
$\chi^2_{0.95,k}$		23.69
N	37607	75104

Notes Dependent variable: 1 if the happiness level of individuals is above the average, 0 if their happiness level is below the average. Standard errors are in the second rows. ( )*, ( )** , ( )*** denote significance at 10 %, 5% and 1 %, respectively.

Table 4A: Fixed Effects versus Random Effects Estimation

Dependent variable: CES-D

Independent variable: Life expectancy of mother

	(1) FE	(2) RE
Age	-0.0575***	-0.0743***
	(0.0156)	(0.0132)
Age squared	0.000389***	0.000388***
	(0.000121)	(0.000102)
Female	-	0.218***
		(0.0212)
Life expectancy of mother	**-0.0105*** **	**-0.00585*** **
	(0.00326)	**(0.00220)**
Ln(income)	-0.0373***	-0.0885***
	(0.00954)	(0.00824)
Health	0.144***	0.301***
	(0.0111)	(0.00664)
Quartile low (25%)	0.0355	0.202***
	(0.0225)	(0.0185)
Quartile high (25%)	-0.0122	-0.0394**
	(0.0213)	(0.0178)
Household size	0.00948	0.0271***
	(0.00831)	(0.00658)
OLFS	0.0265	0.0859***
	(0.0213)	(0.0185)
Unemployed	0.421***	0.514***
	(0.0605)	(0.0555)

Retired	0.125***	0.192***
	(0.0224)	(0.0192)
Disabled	0.483***	0.843***
	(0.0434)	(0.0377)
Married	-0.739***	-0.381***
	(0.125)	(0.0554)
OMS	-0.101	0.173***
	(0.121)	(0.0550)
White	-	-0.379***
		(0.0505)
Black	-	-0.400***
		(0.0554)
Education low	-	0.543***
		(0.0257)
Education high	-	-0.220***
		(0.0272)
Constant	4.118***	5.159***
	(0.530)	(0.439)
Hausman test (d.f.)		819.52 (14)
$\chi^2_{0.95,k}$		23.69
N	74984	74984

Notes; Standard errors are in the second rows. ( )*, ( )** , ( )*** denote significance at 10 %, 5% and 1 %, respectively.

Table 5A: Fixed Effects Logit versus Pooled Logit Estimation

Dependent variable: CES-D Binary

Independent variable: Life expectancy of mother

	(1)	(2)
	FE Logit	Pooled Logit

Age	0.0752** (0.0320)	0.113*** (0.0165)
Age squared	-0.000587** (0.000247)	-0.000718*** (0.000129)
Female	-	-0.240*** (0.0184)
Life expectancy of mother	**0.0128*** **(0.00665)**	**-0.00127** **(0.00237)**
Ln(income)	0.0509*** (0.0194)	0.145*** (0.0110)
Health	-0.217*** (0.0226)	-0.351*** (0.00680)
Quartile low (25%)	0.00536 (0.0439)	-0.277*** (0.0218)
Quartile high (25%)	0.0314 (0.0457)	0.102*** (0.0236)
Household size	-0.0203 (0.0165)	-0.0399*** (0.00759)
OLFS	0.0114 (0.0440)	-0.113*** (0.0247)
Unemployed	-0.504*** (0.115)	-0.618*** (0.0778)
Retired	-0.135*** (0.0457)	-0.260*** (0.0250)
Disabled	-0.503*** (0.0865)	-1.143*** (0.0525)
Married	0.984*** (0.234)	0.239*** (0.0509)
OMS	0.159 (0.226)	-0.241*** (0.0504)

White	-	0.279***
		(0.0427)
Black	-	0.317***
		(0.0465)
Education low	-	-0.427***
		(0.0207)
Education high	-	0.313***
		(0.0250)
Constant	-	-4.220***
		(0.547)
Hausman test (d.f.)		285.25 (14)
$\chi^2_{0.95,k}$		23.69
N	37547	74984

Notes Dependent variable: 1 if the happiness level of individuals is above the average, 0 if their happiness level is below the average. Standard errors are in the second rows. ( )*, ( )** , ( )*** denote significance at 10 %, 5% and 1 %, respectively.

APPENDIX B

Table 1B: FE Logit versus RE Logit Estimations (Full Sample)

	(1) FE logit	(2) RE logit
Education high	-	0.703***
		(0.0937)
Education low	-	-0.716***
		(0.0993)
Health	0.138**	0.263***
	(0.0641)	(0.0289)
Tenure	-0.0784***	-0.0827***

	(0.00648)	(0.00340)
Ln(wage)	-0.975***	-1.471***
	(0.0501)	(0.0367)
Black	-	0.116
		(0.221)
White	-	0.649***
		(0.201)
Life expectancy	**-0.504*****	**-0.483*****
	(0.117)	**(0.0734)**
Life expectancy^2	**0.00200**	**0.00304***
	(0.00261)	**(0.00164)**
Male	-	0.0996
		(0.0955)
Quartile low (25 %)	-0.109	-0.626***
	(0.118)	(0.0809)
Quartile high (25 %)	0.249**	0.930***
	(0.108)	(0.0734)
Couple	0.0111	0.138*
	(0.159)	(0.0824)
Constant		14.43***
		(0.888)
Hausman test (d.f)		685.71 (33)
$\chi^2_{0.95,k}$		47.40
N	12801	43328

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$, age dummies are also included but not reported.