June 2022

Two Essays in Applied Microeconomics: Retirement, Income Inequalities, and other Economic Indicators of Health and Life Satisfaction

Zahra Akbari
University of South Florida

Two Essays in Applied Microeconomics:

Retirement, Income Inequalities, and other Economic Indicators of Health and Life Satisfaction

by

Zahra Akbari

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Economics
College Of Arts and Sciences
University of South Florida

Major Professor: Gabriel Picone, Ph.D.
Padmaja Ayyagari, Ph.D.
Andrei Barbos, Ph.D.
Abolfazl Saghafi, Ph.D.

Date of Approval:
June 24, 2022

Keywords: Health Disparities, Socioeconomic Status, Self-Reported Health, Happiness

Acknowledgements

I would like to acknowledge and give my thanks to my main advisor, Dr. Gabriel Picone, who made this work possible. His guidance and consistent support carried me through all the stages of my Ph.D. program.

I would also like to express my appreciation to my committee members, Dr. Padmaja Ayyagari, Dr. Andresi Barbos, and Dr. Abolfazl Saghafi, for being there for me and helping me through every moment of this path.

Finally, I would like to give special thanks to the wonderful people in the Economics Department at the University of South Florida, especially Dr. Bradley Kamp for his insights and kind support.

Dedication

To the dearest people in my life: my parents, who dedicated every second of their lives to giving me the love and support I needed and helping me take the most courageous steps even when they disagreed with my decisions. I would like to dedicate this work to the rest of my family: my brother, sister-in-law, and my little nephews whose love motivated me and encouraged me to keep going.

To my amazing friends who helped me through the difficult path of graduate school. I would love to give my warmest thanks to my friends who were my rock during times of hardship and loneliness: Raha, Zohre, Riccardo, Javad, Hadi, Hamidreza, Maysam, Fedra, Mohsen, Alireza, Abolfazl, and Skye.

Table of Contents

List of Figures

List of Tables

Abstract

The following proposal includes two essays in applied microeconomics. The first essay studied the relationship between income differences among siblings and the health outcomes of the individuals. Health inequalities and the factors associated with them have been a significant interest of health economists. Among those factors that can lead to health differences in adulthood, many studies have studied financial status. There are still many questions about these factors which should be answered in this area, especially about the adulthood income relative to a reference group and how it can be related to differences in adulthood health. The main goal of this paper was to estimate the relationship between health outcomes and income differences among siblings. We considered the siblings as one of the reference groups that could have a meaningful impact on people's health. This paper also examined whether the income of siblings can have a causal relationship with the good health of the individuals. The causal association between income differences and health indicators, self-reported general health, was examined using the Wisconsin Longitudinal Study data. The results showed a significant positive relationship between the income of a randomly selected sibling and the health outcome but controlling for the endogeneity of the differences in income made the coefficient of the income differences less significant. The results of the fixed effect model showed that the relationship disappeared when we controlled for individual fixed effects. We also discussed the issues of estimating the relationship with controlling for individual fixed effects and suggested a way to solve the issue.

The second essay was on the causal effect of retirement on life satisfaction. The Health and Retirement Study data was used to estimate the impact of retirement on life satisfaction. Additionally, also two-stage

process was used to find the potential mechanism through which retirement impacts life satisfaction. Regression discontinuity design was applied to deal with the reverse causality between retirement and life satisfaction. The eligibility age for pension was used as a rule for treatment assignment. The initial result showed that retiring leads to an increased probability of being satisfied with life. Physical activity, sleep quality, and social contacts are how retirement changes life satisfaction.

Chapter 1: Income Differences among Siblings and Health Indicators

1.1 Abstract

Health inequalities and the factors leading to them have been a significant interest of health economists. As a factor causing health inequalities, financial status has been studied in many papers. However, many questions have not been answered about financial/economic indicators of health inequalities, especially how income relative to a reference group can be related to differences in adulthood health. The main goal of this paper was to estimate the relationship between health outcomes and income differences among siblings. We consider the siblings as one of the reference groups that can impact people's health. This paper also examined whether the income of siblings had a causal relationship with the health of the individuals. The causal association between income differences and health indicators, self-reported general health, was examined using the Wisconsin Longitudinal Study data. The results identified a significant positive association between the income of a randomly selected sibling and the individual's health, but the association got smaller and less significant after controlling for the endogeneity of the differences in income. The results of the fixed effect model showed that the relationship disappeared when we controlled for individual fixed effects.

1.2 Introduction

How can the income of others affect me? This question has always been a major interest of social scientists. Economists have studied this issue for a long time since they started to realize also, the consumption of the others (Alpizar et al., 2005). The issue is even more critical when it gets into the specific areas such as well-being, health behavior and eventually, health outcomes of the individuals, considering that the health status of the individuals might be affected by other people's income and socioeconomic status.

Studying the association between the income of others and the health outcomes will help to answer the some of the most frequently asked questions in the study of health inequality: “why some people are healthier than others” and “which socioeconomic factors are the most important indicators of the health status.” Identifying the more essential factors in determining current and future health can help policymakers build new plans and policies to improve society’s health outcomes. These policies will decrease the future cost of health and increase welfare.

In this paper, we aimed to find the relationship between the health outcomes of the individuals and their income relative to their randomly selected siblings. In theory, it has been discussed that relative income can impact well-being in two directions, which are through comparison effect and the altruistic effect (Senik, 2005). Our goal was to contribute to a collection of the numerous peer studies that have tried to disentangle these two significant paths. In fact, by considering the siblings as the reference group, this paper tests for two hypotheses of the opposite direction that siblings’ income can play a role in the health status of the individuals. The first hypothesis states that if the sibling of an individual is doing better financially, the comparison effect will cause the individual to feel stressed, eventually leading to a worse health outcome for the individual. Despite the first hypothesis, the second one affirms that the higher income the siblings have compared to the individual is considered a signal by the individuals that assures them they will have a better situation or financial support in the future if they need it.

The other contributions of this paper were related to the other side of the story, which is the effect of socioeconomic status on health outcomes. Researchers have tried to work with theory and real-world data to find the sufficient indicators that lead us to find the causes of health inequality. Health inequalities and the associated factors have been of considerable interest to health economists. Various issues can be studied in health inequalities and indicators of it. One of these

issues is to compare the childhood experiences and status of the individuals and see how and to what extent each factor can play a role in determining adulthood health outcomes. A major problem in studying the effect of adulthood earning on health is that childhood and adulthood socioeconomic status are correlated (Cohen et al., 2010). Additionally, the unobserved factors may make it difficult to find a causal relationship between socioeconomic factors and adulthood health, even in rich data sets. The unobserved factors include many variables such as early childhood experiences and genetic factors (Case et al., 2005). That is why finding a new method to control for childhood experiences, parents' socioeconomic status, and genetic effects can be very useful in estimating the extent to which adulthood socioeconomic status affects adults' health.

Although it is daunting to separate the childhood and current socioeconomic factors, a comprehensive data set that has detailed information on both the individual and the peers can provide an opportunity to control for some unobserved childhood and peer fixed effects. We used Wisconsin Longitudinal Study that collects the data on the high school graduates in 1957 and their randomly selected siblings. This data enabled us to control for siblings' fixed effects and family fixed effects to estimate better the correlation between relative income and individual health outcomes (We assumed that kids with the same parents went through similar family environment and experiences).

To summarize this section: the primary goal of this paper was to estimate the relationship between health outcomes and income differences among siblings in adulthood. The method used in this study was to control for childhood status (including neighborhood environment, household environment and other unobserved socioeconomic and environmental factors in childhood) and genetics by looking at siblings' data. By considering the data for siblings, we were automatically controlling for any shared childhood socioeconomic status. Since it could be complicated to

separate the childhood and adulthood status in evaluating the effect of income inequality on health inequalities in adulthood, we looked at the current socioeconomic indicators, explicitly the impact of income differences among siblings on health. The differences in financial situation among siblings were considered one of the economic indicators of socioeconomic status in adulthood. There were some financial indicators whose impact on health outcomes has been estimated before, including household income, wealth, financial difficulties, and financial status. But the effects of differences in income of siblings on health outcomes have not been studied.

Therefore, another contribution of this paper is to estimate how people's financial situation relative to close ones affect their health. Since there are few studies that could estimate the relative financial situation with controlling for main childhood status (ex. home environment in childhood).

The next parts of the paper will be as follows: 1) a brief review of theory and literature on two arguments, how the relative income can impact the well-being and how socioeconomic status can change the health outcomes, 2) data, dependent and independent variables, and methodology, 3) results, 4) conclusion.

1.3 Theoretical Background and Literature Review

This section includes two sub-sections which discuss the theory and literature review on both arguments of this paper. First, the study discusses the important theories and papers which are published about the effect of relative income on well-being and health, and second part is about the background and literature of the impact of the socioeconomic status on adulthood health outcomes.

1.3.1 How Does Other's Income Change Our Well-Being?

Veblen (1909) talks about the problem of the classic utility function in his book "limitation of the marginal utility". He mentions an assumption that the individual's utility is a function of other's

consumption which leads to the importance of the relevance of income in utility function (Ferrer-i-Carbonell, 2005). The role of others in the individual's utility function has led to the introducing a new concept considered as "comparison income effect" which states that individual's relative position to other "people like you" in the society determines the degree of their well-being or happiness (Carporale et al, 2009). To test the income comparison effect, there have been a lot of papers which studied the relationship between relative income and well-being. One of the first studies that has shown that family members compete with each other is the study that Neumark and Postlewaite (1998) do to estimate the effect of family member's financial situation on the women's behavior. They use the national longitudinal survey of youth and they find that the probability of working for the women increases with higher income of their sister's husband. They also show that the probability of young women being in the labor force has a positive and significant correlation with their sister-in-law's employment. These findings proves that there might be a direct preference interdependence (Senik, 2005). Ferrer (2005) studies a study of a large German survey, German Socioeconomic Panel, to check the income comparison effect for a sub-sample over 1992-1997. The findings show that the individuals are happier if their income is larger than their reference group- people in the same age, education and region- as well as the comparison effect works mostly up-ward which means that individuals compare themselves with people with higher income (Ferrer-i-Carbonell, 2005). There are other studies that have considered different reference group for their studies to find that how people react to the increase of their peers' income. Using the data on 5,000 British households, Clark and Oswald find some evidence to prove that people's satisfaction negatively depends on the wage of their reference group. The effect of the comparison income effect in the study gets estimated by considering the employees with the same age and the same level of education as the reference group (Clark and Oswald, 1996). In a more

recent study by Card and his co-authors, the results show that employees' work satisfaction is negatively correlated with their peers' salary. When the employees know that the salary of their peers is higher than their own salary, they are more likely to look for another job (Card et al., 2012).

On the other hand, studies have also looked at the relationship between income of the others and happiness. One of these studies compares the happiness of UK and US's household members finds out that in both countries, social comparison in income would change the happiness of household members (Blanchflower & Oswald, 2004). Headed and Wooden (2004) look for a relationship between happiness and relative income by using the waves of 2001 and 2002 from the data of the Household, Income and Labor Dynamic of Australia. The results show that it is the relative income that cause the difference in people feeling happy not the net income itself and it can change preferences (Headey and Wooden, 2004). They are also some studies which reject the comparison income effect. For instance, Di Tella and MacCulloch find that for over 380,000 observations, in estimating the life satisfaction the coefficient of relative income is positive (Di Tella and MacCulloch, 2003).

In spite of the comparison income effect, there is another side of the argument which has been discussed by the other studies which are based on the "cognitive category ". These studies include is a large number of projects which introduce the peer's income as a cognitive category in examination of the satisfaction in life or in the work environment. They consider the income relative to a reference group as a signal to the future situation of the salary, well-being or health. This is a closer concept to the "tunnel effect" that implies that people see the progression of their reference group as a sign of their own future improvement (Hirschman and Rothschild, 1973). Senik (2004) studies the 1994-2000 waves of Russian Longitudinal Monitoring Survey and apply

a two-stage analysis to test the effect of the reference group- similar profession, diploma, branch, and region- on the individual's satisfaction. She finds a positive relationship between these two variables and provides the evidence to the consideration of income of others as a signal for the future self-improvement (Senik, 2004). Using a Canadian cross-sectional survey, Lévy-Garboua and Montmarquette, show that the previous wage gap has a positive relationship with job and wage satisfaction. The effect of wage gap on the satisfaction declines by higher years of experience (Lévy-Garboua and Montmarquette, 2003).

There has always been a struggle in the effect of economic indicators of socioeconomic status on health outcomes in compared to the effect of more conventional indicators of education and occupation (Duncan & others, 2002). In his paper, Duncan discusses the need for a comprehensive indicator of SES and looks for a relationship between SES and mortality. By assessing a sample of 3734 individuals with age 45 and older, he finds out that wealth and recent family income have the strongest association with mortality, even when other SES indicators are controlled for.
Although there is little support for the direct relationship of income inequality and health, decreasing the income inequality can improve some health indicators (Lynch et al., 2004). Trying to find this relationship in some states around United States, Lynch and her coauthors find out that income inequality can affect some health indicators like homicide rate.

There are many other works that have proved that people compare their level of living with the close people around them and it effects their levels of satisfaction in the life (Usui et al., 1985). There are some studies about the effect of relative income on people's mental health. There is a positive relationship between the relative income and positive feelings (Yu & Chen, 2016). In the study which is conducted by Yu and Chen, the association of absolute and relative income with well-being has been estimated in some areas of China. They show that subjective well-being has

a strong negative relationship with people feeling negatively or depressed. It also has been shown that it is the relative income that is associated with the differences in people feeling happy, not the net income itself, and it can change preferences (Heady and Wooden, 2004).

1.3.2 How Socioeconomic Status Can Change Our Health Status

Researchers have shown that childhood and current experiences can be linked to adulthood health status. This link can be through several channels such as adulthood psychological, health behavior and physiological mechanisms which affect the health outcomes (Cohen et al., 2010). The correlation between health outcomes and socioeconomic status in both childhood and adulthood have been explained by theoretical support based on environmental and physiological factors (Gruenewald et al., 2012). The childhood status along with adulthood mechanisms gives a cumulative measure of socioeconomic status which can lead to health differences, this issue is shown in the figure 1 which is extracted from the paper written by Gruenewald and her colleagues. The figure shows the potential pathways through which social and economic status is linked to health status in adulthood.

There has been a wide literature of studies which have studied the association of socio-economic situations and health status through different channels. Some of these studies have used the life history approaches (Hagemaster, 1992). These studies mix the recent biological findings from examining the natural history of the disease with the adult life path of socio-economic status considered by social sciences (Wadsworth, 1997). Wadsworth's paper is one of the papers that has used the life history approach to introduce methods for estimating the relationship between socio-economic situations and health in adulthood. His paper can be categorized among papers who have studied the factors that can be related to health inequalities among individuals. He mentions that life history studies were the beginning to focus on importance of factors other than genetics - such

as social and economic situations – in determining the health path of adulthood. While Wadsworth doesn't use any data to estimate the association of socio-economic status and health outcome, some studies use the same approach as they work on data to estimate the extant that socio-economic factors can associate with health. Van De Mheen and others try to answer the question of how childhood environment contributes to the explanation of inequalities in adulthood health. They gather the retrospective data from an interview which was a section of longitudinal study on socio-economic health differences in south-east of Netherlands (Van De Mheen and others, 1997). The variables which have been used are family structure and financial situation for socio-economic status, and general health and self-reported of chronic conditions for health indicators respectively. They find that the health differences can be assigned to the differences in the characteristics of childhood environment, the most important characteristics are the father's job, mother's education, and financial situation.

In the recent papers, more expanded issues of childhood status' effect on adulthood health have been worked on. Richie Poulton and his co-authors mention the measurement errors of childhood socioeconomic status and use a new way to control for these errors. The child's initial infant health status and also adulthood socioeconomic status should be controlled for to make a better estimation of the association of childhood socioeconomic status with current health (Poulton & others, 2002). In their paper, Poulton and his colleagues do a longitudinal study of an unselected cohort of 1000 children from Dunedin, New Zealand. They report that there are significant differences in health of children grew up in low socioeconomic status and children who were from families of high socioeconomic status. The important determinants of childhood status can change the adulthood health through their role in forming the initial adult health and through a lasting direct effect of childhood health in middle age (Case et al., 2004). Studying the lasting effect of

childhood status on adulthood health has shown that children with poor health have less educational achievements, poorer adulthood health and lower socioeconomic status (Case et al., 2004). The association of early health and adulthood health has been studied in another research. It has been shown that differences in early health status can be effective in reproducing the socioeconomic inequalities in adulthood (Palloni et al., 2009). In this paper, the findings reveal that childhood health status have "small but non-trivial" effect on socioeconomic gradient of health in adulthood. After controlling for adulthood socioeconomic status, childhood environment can be a powerful factor in determination of adulthood cardiovascular mortality and other cause mortality (Cohen et al., 2010). Cohen and others also mention in their paper that childhood status can change adult health through adulthood socioeconomic status. That is why they go through a long list of evidence which show how childhood socioeconomic status can affect adulthood health status, after controlling for adulthood socioeconomic status.

As it is mentioned before, there are many other papers which have evaluated the importance of childhood environment on health in adulthood. It has been tried to control for childhood socioeconomic status to see how effective adulthood status can be and there has been a big challenge of answering the question of which one- childhood or adulthood status- is the most important. That's why the comparison on past and current status's role in health will rise to be in our attention. While the early experiences have been introduced as a very important determinant of adulthood health, some studies have challenged their value in compared to adulthood situation's role in health (Rahkonen & others, 1997). Rahkonen and his coauthors used a survey on Living Conditions data from Finland. They looked for past and current socio-economic situation of individuals who could have an impact on adulthood health for Finn men and women. For childhood and current living condition they used the factors such as one concerning economic problem, three

concerning family-related social problem and degree of urbanization of living area. They found out that economic problem in both childhood and current situation have stronger association with adulthood health. They also ended up with a result of current socioeconomic status being the strongest factor determining the health. There are more recent papers who have compared the childhood and current socio-economic situations and their relationship with health in adulthood. One of them is the work of Louise and Zhao who looked for the effects of family structure and adulthood experiences on life satisfaction. They collected their data from the General Social Survey to measure the variables of financial situation which was the respondent's assessment of family income at age 16 in compared to the average American family (Louise and Zhao, 2002). They acquired that family structure can affect adult well-being, but after controlling for adulthood experiences, their importance was less significant, which shows the importance of experiences in adulthood. Hambleton and other co-authors use the collected data from "SABE project" in Barbados to estimate the relative contribution of historical and current status in elderly. Their findings show that the 5.2% contribution of individual's past experiences to health is reduced to 2.0% when they consider the current experiences. Measuring the current status gets more accurate when three current indicator –current socioeconomic status, lifestyle risk factors and disease indicators- is used (Hambleton & others, 2005). As they find, disease indicators are the strongest indicators for elderly people. Current paper accounts for childhood and current socioeconomic status by using individuals and sibling's current economic situation and their parent's socioeconomic status when they were kids.

The study contributes to the literature in two ways: first, it controls for unobserved childhood environment and genetics using the randomly selected sibling as a reference group; second, it adds to the literature of the association between health and the relative income of close one. In other

words, the study teste for the “comparison income effect” and “cognitive category” with the reference group of siblings, which has not been studied before.

1.4 Data

The data was collected from the Wisconsin Longitudinal Study (WLS), which includes the information from a longitudinal study of a random sample of 10,317 Wisconsin high school graduates of 1957. This data set also includes the information for the sample’s randomly selected siblings by following their life path through repeated surveys. The surveys have been conducted in years 1957, 1964, 1975, 2004-2006 and 2010-11. WLS is a comprehensive and detailed collection of educational, social, economic, mental, and physical health information in a relatively homogeneous population (Sewell et al., 2004; Hauser, 2009). The sample size was 9,318 (90.1% of surviving participants), 8,493 (87% of surviving participants), 7,265 (80% of surviving participants), and 5,986 (74.9% of surviving participants) for the year 1975, 1992, 2004-2006 and 2011 respectively (Taylor & Shivers, 2013).

WLS started gathering data for graduate respondents with an in-person questionnaire in 1957, followed by parents’ data collection seven years after the graduation of students in 1964, 1975 survey, 1993 and 2004 telephone and mail surveys, and 2011 in-person questionnaire. The sibling was randomly selected from a list of all siblings. If the graduate’s sibling was twins, the twin was picked. For the siblings' data set, 2000 siblings were empaneled in 1977, and the total sibling sample was implemented in 1994. The siblings' data expanded the WLS survey in 1992-93 to include a randomly selected sibling of every respondent with at least one brother or sister. The content was extended to obtain detailed occupational histories and job characteristics; incomes, assets, and inter-household transfers; social and economic information of parents, siblings, and children and descriptions of the respondents’ relationships with them; and extensive information

about mental and physical health and well-being (WLS website). Once empaneled, a sibling survey has been fielded either subsequently or concurrently to the graduate survey in each round (Pudrovska, 2014). The parallel siblings' data is collected for about 56.4% of graduates (n = 6,897). The sample size for siblings was 4,804 (70% of surviving participants) for the years 1993-1994, 4,270 (74% of surviving participants) for the year 2003-2005, and 3,397 (78% of surviving participants) for the year 2011 (table 1).

This study used the data from the graduates with available data for the randomly selected siblings in two steps. First, the graduates were chosen from surveys from 1992-1993, 2003-2005, and 2011. Then graduates with no data on siblings were removed from the data set, leaving us with a sample size of 4804, 4270, and 3397 for 1992-1993, 2003-2005, and 2011 respectively. Table 1 shows the number of surviving participants for each wave of surveys (siblings and graduates). We used the sample with all information for siblings and individuals' income and health status (N=2718 per year).

1.5 Measures

Six categories of indicators were used to find the association between health status and income differences between siblings: graduate's health indicators, graduate's financial indicators including income and income differences[1], graduate's current socioeconomic indicators other than income, graduate's health behavior indicators, family's characteristics, and sibling's socioeconomic characteristics.

Health indicators:

The original health indicator was a self-reported measure, which asked graduates to rate their general health. The indicator ranged from 1 to 5 (poor, fair, good, very good, and excellent), which

[1] Income relative to the sibling.

was converted to a binary variable that was 1 if the health status was good or above and 0 if otherwise.

1.5.1 Financial Indicators

The total household income was used to measure the financial status of the graduates and siblings. The log of graduate's income was subtracted from the and the log of sibling's income to calculate the income differences.

1.5.2 Current Socioeconomic Indicators Other Than Income

Education: The highest degree level was used to build a binary education variable, which was 1 if the individuals had at least a 4-year degree or 0 if otherwise. Social contact was measured by adding the frequencies of "getting together" and the hours of "talking on the phone" with friends or relatives during the last week. Employment status was a binary indicator that was 1 if the individual was currently employed and 0 otherwise. The marital status was a yes/no question of if the graduates were currently married. Family size measured the number of households that the individual lived with in each wave.

1.5.3 Health Behavior Indicators

Smoking status was assigned as 0, non-smoker, and 1, current smoker. Alcohol consumption was categorized as heavy-drinker, 1, or non-heavy drinker as 0. Heavy consumption of alcohol was 15 or more drinks in a week, as defined by the "National Institute of Alcohol Abuse and Alcoholism."

1.5.4 Family Characteristics

These variables measured the different characteristics of the family that individual lived with in 1957. The education of parents identified that if the mother and father of the individual had the

BA or higher degree. Father's job indicated the occupation code of the father which was assigned from 1-5, 1 for being in the category of unskilled and 5 for being in the top managers and CEO job category. The degree of urbanization measured the population of the region individual was living in 1957. We also controlled for the financial situation of the individual relative of the other parts of the society and the parent's income in 1957.

1.5.5 Sibling's Socioeconomic Status

We control for sibling's characteristics including their gender, marital status, social contacts, family size, education, and employment.

We did not control for race, because less than 2% of graduates from 1957 were non-white and asking about ethnic background was illegal at the initial time of collecting data. Even after it was legal, the number of non-whites in the sample was too small to justify any analysis of among non-white members of the sample (Sewell et al., 2003). Table 2 shows the summary statistics for all of the indicators which are used in our model.

1.5.6 Empirical Specification

For analyzing the data, a fixed-effect model with IV was used and compared to the pooled OLS regression with period dummies to a fixed-effect model. A two-stage analysis was applied:

$$\overline{Inco\ \ me}\overline{Diff}_t = \alpha_1.AgeS_{it} + \alpha_2.AgeS_{it}{}^2 + \alpha_3.Unemp18 + \alpha_4.X_{it} + \alpha_5.S_{it} + v_i + z_{it}+ \quad \text{A.t}$$

$$h_{it} = \beta.\overline{Inco\ \ me}\overline{Diff}_t + \alpha.Income_{it} + \theta.X_{it} + \gamma.S_{it} + v_i + u_{it}+ \quad \text{B.t}$$

Which h_{it} is the health indicator which collects the information on general health of the individual in period t, $IncomeDiff_t$ is the difference between the log of household income for sibling and the individual in period t, $Income_{it}$ is the log of individual's household income in period t and X_{it} is other characteristics of the individual in period t. X_{it} includes the social contacts,

employment, marital status, and health behaviors of the individual. As mentioned before, t is the period of survey which includes 1992-1993, 2003-2004 and 2011. S_{it} is the socioeconomic status of the sibling (including education, employment status, marital status, family size, social contact, and gender), v_i is the individual fixed effects that are not time-variant. These variables include the individual's childhood socioeconomic status which can be a collection of various characteristics like parents' income, their education or occupation, suffering from some financial disadvantage in childhood and the rest of childhood environment. As it has been shown in previous studies childhood socioeconomic status can be associated with health, but with controlling them by using fixed effect model we can get a better result. v_i, if it includes childhood environment cannot be associated with income differences since it is the same for individual and her sibling, but if it's related to other individual characteristics like performance in the school or other personal characteristics like gender, it can be correlated with income differences. Therefore, controlling for all time invariant variables makes the results less biased. u_{it} measures other factors which are effective on health, and we have not considered yet, these are the characteristics that are time variant. A.t are dummy variables for the waves. The analysis was done using STATA 17.

1.5.7 Instrumental Variables

The IVs used are three variables which indicate the income of siblings which include unemployment rate in the state when the sibling was 18, age and $age^2$of the sibling. The reason we used the age and age^2 is that income changes by age in U shaped function. The age and $age^2$of the sibling will change the income of sibling and not the income of the individual, therefore we use those three together as an IV for the differences in income. Unemployment rate when sibiling was 18, and Age-squared do not directly correlated with health of the individual, they can be related to the health of the individual through the effect they have on the sibiling's income. The

results for F test (F= 67.46) show that the age and age^2 and the unemployment rate at 18 are good instruments for the income differences. Table 6 shows the Kleibergen-Paap statistic which tests for weak instrument. The results of test rejected the null hypothesis that three used IVs were weak instruments.

1.6 Results

The first pooled OLS regression of good general health indicator on independent variables shows that there is a significant positive relationship between the individual and her sibling's different income and individual's general health. In 95 % confidence interval, the coefficient of difference in income is 0.006 which is significant by t equals to 2.123. This shows that if the sibling's income increases by 10 percent relative to individual's income, the probability of the individual being in the group of people with good health increases by 0.06 percent point. The results show a positive significant relationship between general health and household income, social contacts, being employed, family size and education. It is shown that having a good general health has a significant positive association with being a heavy drinker, which is not expected, the reason for this result might be a reverse causality between drinking and health status. Another reason can be because of many missing data for health behavior. Including IV decreases the significance of the differences in income but the magnitude of the coefficient gets bigger. The coefficient of the employment gets bigger and more significant which is expected in respect to the nature of the IV. The summary of results for pooled OLS model with and without IV are shown in table 3.

To check the robustness of the pooled OLS regression, we first remove the sibling's characteristics and family's socioeconomic indicators from the main regression. As table 4 shows by comparing the pooled OLS and pooled OLS with IV, the magnitude of the difference in income

gets bigger by using the IV and the t value gets bigger. Overall, the magnitude and the direction stay the same with excluding the family and sibling's characteristics. We also exclude just sibling's indicators for another robustness check, as it is shown in table 5, the coefficient which is associated differences in income stays the same when we use the IV in the regression. Table 6 compares the three model together, the magnitude of the coefficients for differences in income in pooled OLS are very similar and they are also significant. But for pooled OLS with IV, the pattern of differs between three models. The significance of the coefficient for income differences gets larger by including IV in both models with excluded siblings and excluded siblings and the family's characteristics. The magnitude for the coefficient decreases to 0.13 and 019 which are so close, and it is interesting considering the larger sample.

Table 7 shows the results for the fixed effect model, including with and without the IV. The first column shows the results for the model without IVs. The results for income, although in line with the previous findings of our study, are insignificant. But we still see a positive relationship between income differences and health status. The results for income and income differences became insignificant when we included IVs in our model with individual fixed effects. The only variables that still stay significant are employment, marital status, family size, and heavy drinking.

1.7 Discussion and Future Research

As the results show there is a positive and significant relationship between the income of siblings relative to the individuals' income in pooled OLS model, with and without IV. The results prove that the income of reference group, sibling's income in our study, can be considered as a signal to the future improvement in socioeconomic status of the individual and eventually it can cause to a better health status. The results of fixed effect model show that by controlling for individual fixed effects, the positive relationship of income of reference group disappears. There

are some issues with using a fixed effect model which drives us toward using pooled OLS instead. The biggest issue is the unclear nature of controlling for individual fixed effect when estimating the health changes. This issue mostly is caused by differences in individuals' ratio of health depreciation over the time. Since these differences depend on the individuals' characteristics including time-variant and time-invariant, controlling for individual fixed effects might not be an efficient tool to estimate the effect of relative income on the health over the time. We suggest two methods to solve this problem. First, we will interact the time trend with individual fixed characteristics. The other way is to separate the sibling's fixed effects from individual fixed effects.

There are some aspects in which we should consider about the data set, since the data set is homogeneous. The main question is that how cultural differences can change the results. To answer this question, we should have a less homogeneous sample. The future study should include other cultures either a sample from other countries or a sample from other states in the United States. Because the differences in income can affect individuals' health in a different extent and different ways based on the cultural differences among individuals. Our sample is from a very specific region and with almost completely homogeneous racial culture. So, it is necessary to look at the other cultures and compare the results with our findings. We have another issue of having non-consistent data set during the time. There is a 10-year gap between surveys, and it must be a reason of non-clear results. So, the next future work will be to find a data set with information on all of the years. The next aspect of the study will be to compare the results with other people close to individuals other than siblings because of possible correlations among financial situations of siblings, so the comparison should be done on different close people to the sample like friends and the people from the same cohort graduated from the same high school. Another issue can be in defining the correct variables. Defining the write well-being indicators have been always a big

challenge. To decide whether the indicators are the dependent variables we should look at or they are the indicators which affect other well-being indicators and should be considered as independent variables can be crucial and effective in the results.

Therefore, for the future study and improving the current results, a sample including various cultures and ethnicities should be considered. Also, relative income should be studied in different levels and the comparison among levels should be examined carefully. Of course, all of these modifications should accompany the big correction in defining the new variables. Also, the possibility of other instrumental variables should be considered.

1.8 References

Alpizar, Francisco, Fredrik Carlsson, and Olof Johansson-Stenman. "How much do we care about absolute versus relative income and consumption?" *Journal of Economic Behavior & Organization* 56.3 (2005): 405-421.

Blanchflower, D. G., & Oswald, A. J. (2004). Well-being over time in Britain and the USA. *Journal of public economics*, *88*(7), 1359-1386.

Card, David, et al. "Inequality at work: The effect of peer salaries on job satisfaction." *The American Economic Review* 102.6 (2012): 2981-3003

Card, David, Alexandre Mas, Enrico Moretti, and Emmanuel Saez. "Inequality at work: The effect of peer salaries on job satisfaction." *The American Economic Review* 102, no. 6 (2012): 2981-3003.

Caporale, Guglielmo Maria, et al. "Income and happiness across Europe: Do reference values matter?" *Journal of Economic Psychology* 30.1 (2009): 42-51.

Case, A., Fertig, A., & Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of health economics*, *24*(2), 365-389.

Clark, Andrew E., and Andrew J. Oswald. "Satisfaction and comparison income." *Journal of public economics* 61.3 (1996): 359-381.

Cohen, S., Janicki-Deverts, D., Chen, E., & Matthews, K. A. (2010). Childhood socioeconomic status and adult health. *Annals of the New York Academy of Sciences*, *1186*(1), 37-55.

Di Tella, R., and R. MacCulloch. Income, happiness and inequality as measures of well-being. Mimeo: *Harvard Business School*, 2003.

Duncan, G. J., Daly, M. C., McDonough, P., & Williams, D. R. (2002). Optimal indicators of socioeconomic status for health research. *American journal of public health*, *92*(7), 1151-1157.

Ferrer-i-Carbonell, Ada. "Income and well-being: an empirical analysis of the comparison income effect." *Journal of Public Economics* 89.5 (2005): 997-1019.

Gruenewald, T. L., Karlamangla, A. S., Hu, P., Stein-Merkin, S., Crandall, C., Koretz, B., & Seeman, T. E. (2012). History of socioeconomic disadvantage and allostatic load in later life. *Social science & medicine*, *74*(1), 75-83.

Hagemaster, J. N. (1992). Life history: a qualitative method of research. *Journal of advanced nursing*, *17*(9), 1122-1128.

Hambleton, I. R., Clarke, K., Broome, H. L., Fraser, H. S., Brathwaite, F., & Hennis, A. J. (2005). Historical and current predictors of self-reported health status among elderly persons in Barbados. *Revista Panamericana de Salud Pública*, *17*(5-6), 342-352.

Headey, B., & Wooden, M. (2004). The Effects of Wealth and Income on Subjective Well-Being and Ill-Being. *Economic record*, *80*(s1).

Hirschman, Albert O., and Michael Rothschild. "The Changing Tolerance for Income Inequality in the Course of Economic Development* With A Mathematical Appendix." *The Quarterly Journal of Economics* 87.4 (1973): 544-566.

Lévy-Garboua, Louis, and Claude Montmarquette. "Reported job satisfaction: what does it mean?" *The Journal of Socio-Economics* 33.2 (2004): 135-151.

Louis, V. V., & Zhao, S. (2002). Effects of family structure, family SES, and adulthood experiences on life satisfaction. *Journal of Family Issues*, *23*(8), 986-1005.

Lynch, J., Smith, G. D., Harper, S. A., Hillemeier, M., Ross, N., Kaplan, G. A., & Wolfson, M. (2004). Is income inequality a determinant of population health? Part 1. A systematic review. *Milbank Quarterly*, *82*(1), 5-99.

Palloni, A., Milesi, C., White, R. G., & Turner, A. (2009). Early childhood health, reproduction of economic inequalities and the persistence of health and mortality differentials. *Social science & medicine*, *68*(9), 1574-1582.

Poulton, R., Caspi, A., Milne, B. J., Thomson, W. M., Taylor, A., Sears, M. R., & Moffitt, T. E. (2002). Association between children's experience of socioeconomic disadvantage and adult health: a life-course study. *The lancet*, *360*(9346), 1640-1645.

Senik, Claudia. "Income distribution and well-being: what can we learn from subjective data?" *Journal of Economic Surveys* 19.1 (2005): 43-63.

Senik, Claudia. "When information dominates comparison: Learning from Russian subjective panel data." *Journal of Public Economics* 88.9 (2004): 2099-2123

Sewell, W. H., Hauser, R. M., Springer, K. W., & Hauser, T. S. (2003). As we age: A review of the Wisconsin Longitudinal Study, 1957–2001. *Research in social stratification and mobility*, *20*, 3-111.

Siegrist, J., Wahrendorf, M., Von Dem Knesebeck, O., Jürges, H., & Börsch-Supan, A. (2007). Quality of work, well-being, and intended early retirement of older employees—baseline results from the SHARE Study. *The European Journal of Public Health*, *17*(1), 62-68.

Pudrovska, T. (2014). Early-life socioeconomic status and mortality at three life course stages: an increasing within-cohort inequality. *Journal of health and social behavior*, *55*(2), 181-195.

Rahkonen, O., Lahelma, E., & Huuhka, M. (1997). Past or present? Childhood living conditions and current socioeconomic status as determinants of adult health. *Social science & medicine*, *44*(3), 327-336.

Taylor, J. L., & Shivers, C. M. (2011). Predictors of helping profession choice and volunteerism among siblings of adults with mild intellectual deficits. *American journal on intellectual and developmental disabilities*, *116*(4), 263-277.

Thoits, P. A. (2010). Stress and health major findings and policy implications. *Journal of health and social behavior*, *51*(1 suppl), S41-S53.

Van De Mheen, H., Stronks, K., Van Den Bos, J., & Mackenbach, J. P. (1997). The contribution of childhood environment to the explanation of socio-economic inequalities in health in adult life: a retrospective study. *Social Science & Medicine*, *44*(1), 13-24.

Usui, W. M., Keil, T. J., & Durig, K. R. (1985). Socioeconomic comparisons and life satisfaction of elderly adults. *Journal of Gerontology*, *40*(1), 110-114.

Veblen, Thorstein. "The limitations of marginal utility." *Journal of political Economy* 17.9 (1909): 620-636.

Wadsworth, M. E. J. (1997). Health inequalities in the life course perspective. *Social science & medicine*, *44*(6), 859-869.

Yu, Z., & Chen, L. (2016). Income and Well-Being: Relative Income and Absolute Income Weaken Negative Emotion, but Only Relative Income Improves Positive Emotion. *Frontiers in Psychology*, *7*.

1.9 Tables and Figures[1]:

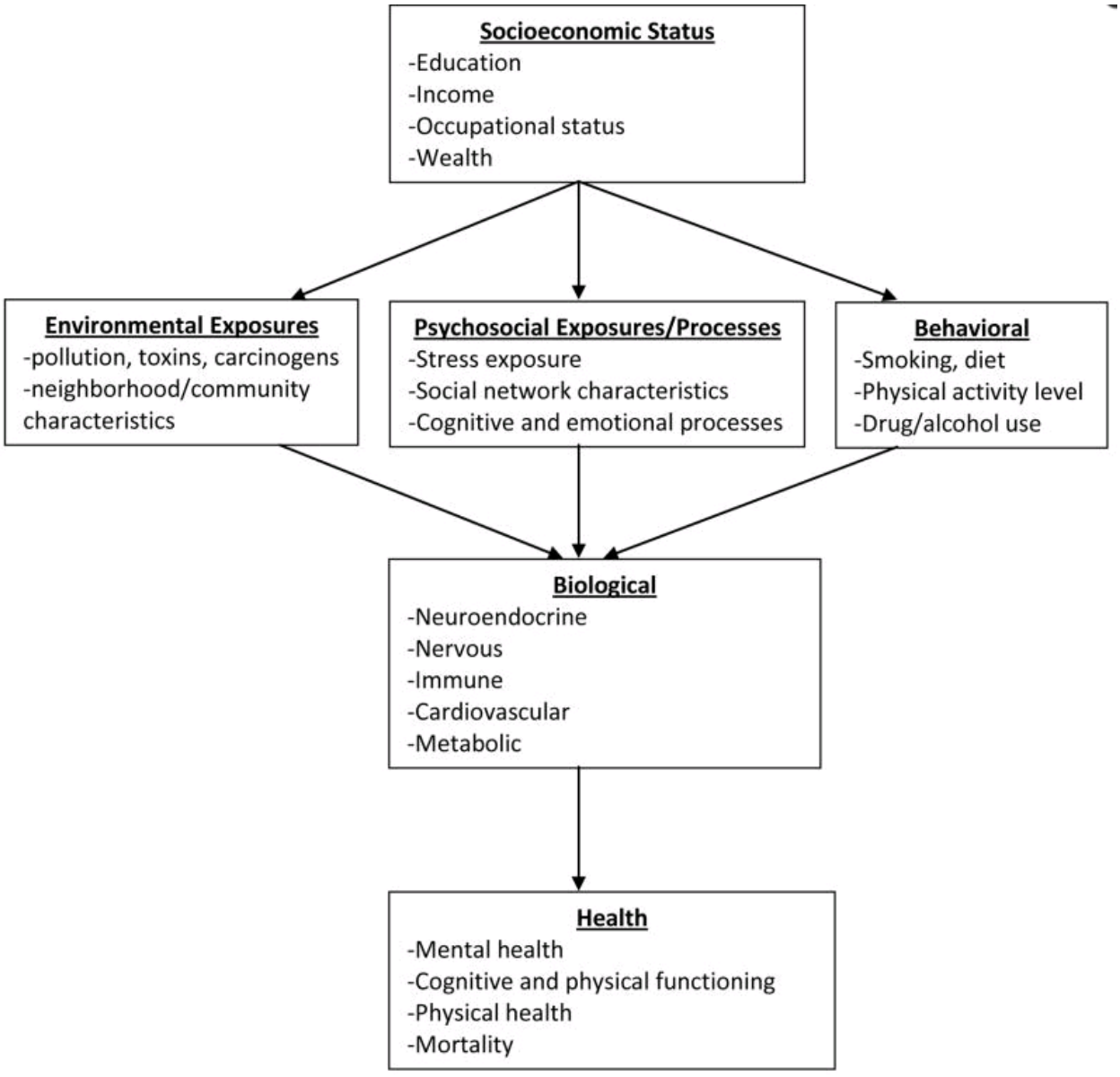

Figure 11-1: Potential channels which socioeconomic status can associate with health in adulthood (Gruenewald et al., 2012)

[1] The tables and figures are listed in the order they appear in the chapter.

Table 1-1: Sample Size
Total Size of existing data and subsamples for graduates and siblings by year of survey-randomly picked siblings

		Graduate			Sibling	
	Survey Year	total	Women	Men	total	Women	Men
Existing Data							
	1992-1993	8493	4513	3980	4804	2480	2324
			(53%)	(47%)		(52%)	(48%)
	2003-2005	7265	3895	3370	4270	2240	2030
			(54%)	(46%)		(52%)	(48%)
	2011	5968	3191	2777	3397	1768	1692
			(53%)	(47%)		(52%)	(48%)
Sample Size for randomly picked siblings							
	1992-1993	4804	2537	2267		2480	2324
			(53%)	(47%)		(52%)	(48%)
	2003-2005	4270	2220	2050		2240	2030
			(52%)	(48%)		(52%)	(48%)
	2011	3397	1789	1608		1768	1692
			(53%)	(47%)		(52%)	(48%)
Sample Size for one sibling sample							
	1992-1993	1205	642	563		659	546
			(53%)	(47%)		(55%)	(45%)
	2003-2005	1070	551	519		589	481
			(52%)	(48%)		(55%)	(45%)
	2011	868	444	424		478	390
			(51%)	(49%)		(55%)	(45%)

Table 1-2: Descriptive Statistics for the variables

VARIABLES	(1) N/%	(2) Mean	(3) SD	(4) Min	(5) Max
Health		0.695	0.460	0	1
Good	69.5				
Below good	30.5				
Income	8,154	10.70	0.980	1.099	13.47
Income Difference	8,154	-0.0971	1.283	-8.319	10.62
Age	8,154	61.917	7.481	51	74
Sex		1.516	0.500	0	1
Female	52				
Male	48				
Employment		0.548	0.497	0	1
Currently employed	54.86				
Currently unemployed	45.14				
College		0.273	0.445	0	1
BA or higher	27.3				
No college	72.7				
Smoking		0.155	0.362	0	1
Current smoker	15.6				
Non-current smoker	84.4				
Heavy drinker		0.067	0.270	0	1
Heavy drinker	6.7				
Non-heavy drinker	93.3				
Physical activity		0.735	0.440	0	1
Yes	73.75				
No	26.35				
Family financial status	8,154	3.149	0.577	1	5
Urbanization	8,154	4.772	2.348	1	8
Father's job	8,154	2.438	1.453	1	5
Father's education		0.0913	0.288	0	1
BA or higher	9.2				
No college	90.8				
Mother's education		0.0955	0.294	0	1
BA or higher	9.5				
No college	90.5				
Parents' income	8,154	9.008	0.578	5.704	12.43
Sibling's age	8,154	61.253	9.769	29	92
Sibling's sex		1.495	0.500	0	1
Female	49.58				
Male	50.42				
Sibling's employment		0.418	0.493	0	1
Currently employed	41.8				
Currently unemployed	58.2				
Sibling's marital status		0.763	0.425	0	1
Currently married	76.3				
Currently non-married	23.7				
Sibling's family size	8,154	2.237	0.994	1	11
Social gathering	8,154	7.224	6.593	0	128
Sibling's education		0.31	0.460	0	1
	30.6				
	69.4				

Table 1-3: Results for the main regression, pooled OLS and pooled IV

VARIABLES	(1) pooled OLS	(2) Pooled IV
Income	0.011***	0.071**
	(2.645)	(2.082)
Income Difference	0.006**	0.08*
	(2.123)	(1.835)
Employment	0.105***	0.111***
	(4.209)	(3.356)
College	0.0317	-0.0123
	(1.217)	(-0.295)
Marital Status	0.0201	0.0376
	(0.525)	(0.769)
Social Gathering	0.00314**	0.00281
	(2.195)	(1.460)
Family Size	0.0534***	0.0862***
	(3.383)	(3.693)
Female	0.0349	-0.00554
	(1.414)	(-0.146)
Female (S)	0.0113	0.0601
	(0.489)	(1.215)
Family Size (S)	0.0149	0.0128
	(1.150)	(0.813)
Employment (S)	-0.0268	-0.145**
	(-1.106)	(-1.984)
Social Gathering (S)	0.00438**	0.00545**
	(2.199)	(2.142)
Marital Status (S)	-0.0121	-0.163
	(-0.368)	(-1.523)
Education (S)	-0.00715	-0.121
	(-0.282)	(-1.456)
Father's education	0.0728**	0.0327
	(2.038)	(0.643)
Mother's education	-0.0164	0.0195
	(-0.479)	(0.419)
Parents' income	0.00895	-0.0242
	(0.369)	(-0.597)
Family financial status	0.0140	0.0254
	(0.674)	(0.825)
Father's job	0.00468	-0.000112
	(0.529)	(-0.00952)
Urbanization	-0.00229	-0.00190
	(-0.467)	(-0.289)
Smoking	3.12e-05	0.0164
	(0.00102)	(0.398)
Drinking	0.0851**	0.0968*
	(2.089)	(1.666)
Physical activity	0.0504*	0.0595*
	(1.892)	(1.714)
Constant	-0.434	-3.429**
	(-1.608)	(-2.006)
Observations	8,154	8,154
R-squared	0.064	0.064

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1-4: Results of the regression excluding family's characteristic and sibling's SES

VARIABLES	(1) pooled OLS	(2) Pooled IV
Income	0.004*	0.002
	(2.001)	(1.647)
Income Difference	0.0001	0.006
	(1.502)	(0.558)
Employment	0.0941***	0.0735***
	(5.916)	(3.978)
College	0.0806***	0.0623***
	(4.654)	(3.029)
Marital Status	0.0509**	0.0415
	(2.060)	(1.538)
Social Gathering	0.00162	0.00106
	(1.183)	(0.852)
Family Size	0.0261**	0.0379***
	(2.329)	(3.024)
Female	0.0527***	0.0395**
	(3.173)	(2.132)
Smoking	0.0354*	0.0446*
	(1.691)	(1.886)
Drinking	0.0595**	0.0357
	(2.053)	(1.063)
Physical activity	0.0200	0.0242
	(1.113)	(1.212)
Constant	-0.180	-1.820***
	(-1.391)	(-4.434)
Observations	8,154	8,154
R-squared	0.047	0.045

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1-5: Results for main regression excluding the sibling's characteristics

VARIABLES	(1) pooled OLS excluding S's variables	(2) Pooled IV excluding S's variables
Income	0.008***	0.023**
	(2.450)	(2.036)
Income Difference	0.004*	0.020**
	(1.799)	(2.301)
Employment	0.0831***	0.0686***
	(4.156)	(3.127)
College	0.0551**	0.0422*
	(2.574)	(1.745)
Marital Status	0.0221	0.0182
	(0.738)	(0.577)
Social Gathering	0.000956	0.000824
	(0.676)	(0.601)
Family Size	0.0471***	0.0599***
	(3.922)	(4.585)
Female	0.0505**	0.0402*
	(2.458)	(1.790)
Father's education	0.0336	0.0174
	(1.039)	(0.503)
Mother's education	0.0234	0.0403
	(0.751)	(1.205)
Parents' income	0.00107	-0.0157
	(0.0579)	(-0.730)
Family financial status	0.0126	0.0129
	(0.751)	(0.706)
Father's job	-0.00164	-0.00254
	(-0.221)	(-0.316)
Urbanization	-0.00485	-0.00565
	(-1.174)	(-1.273)
Smoking	0.0335	0.0504*
	(1.296)	(1.770)
Drinking	0.0588*	0.0518
	(1.673)	(1.352)
Physical activity	0.0303	0.0432*
	(1.399)	(1.838)
Constant	-0.283	-1.312***
	(-1.378)	(-2.978)
Observations	8,154	8,154
R-squared	0.052	0.050

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1-6: Comparing the main regression with models with no controls for sibling's and family's characteristics

VARIABLES	(1) Pooled OLS, main	(2) Pooled IV, main	(3) Pooled OLS no control for siblings	(3) Pooled IV, no control for siblings	(1) Pooled OLS no control for siblings & family	(2) Pooled IV, no control for siblings & family
Income	0.011***	0.071**	0.008***	0.023**	0.004*	0.002
	(2.645)	(2.082)	(2.450)	(2.036)	(2.001)	(1.647)
Income Difference	0.006**	0.08*	0.004*	0.020**	0.0001	0.006
	(2.123)	(1.835)	(1.799)	(2.301)	(1.502)	(0.558)
Kleibergen-Paap statistic	-	19.02	-	21.8	-	22.9
		0.009		0.003		0.003
Observations	8,154	8,154	8,154	8,154	8,154	8,154
R-squared	0.064	0.064	0.052	0.050	0.047	0.045

Robust t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 1-7: Results of FE model

VARIABLES	(1) Fixed Effects	(2) Fixed Effects IV
Income	0.006	0.002
	(0.86)	(0.03)
Income Difference	0.009*	0.001
	(1.82)	(0.01)
Employment	0.112**	0.191*
	(2.247)	(1.84)
Education	0.017**	0.0181
	(2.13)	(-0.0983)
Marital Status	-0.297	-0.008*
	(-0.810)	(-1.756)
Social gathering	-0.00122	-0.00128
	(-0.305)	(-0.157)
Family size	0.106***	0.0898
	(2.979)	(1.398)
Drinking	0.261***	0.271
	(2.746)	(1.496)
Physical activity	0.0472	0.0830
	(0.935)	(0.795)
Constant	0.528	5.627
	(0.850)	(1.552)
Observations	8,154	8,154
R-squared	0.129	

Robust t-statistics (z-statistics) in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Chapter 2: Causal Effects of Retirement on Life Satisfaction, Possible Mechanisms, And A Regression Discontinuity Approach

2.1 Abstract

The health outcomes of retirement have been of significant interest to researchers and policymakers working in social security policy. This study sought to examine the causal relationship between retirement and life satisfaction. A regression discontinuity design was used to address the endogeneity problem. Data were collected from the Health and Retirement Study to estimate the retirement effects on two measures of life satisfaction. The paper also looked for mechanisms through which getting retired affects the individuals' life satisfaction. The results show that retirement increases the probability of being satisfied with life. Sleep behavior, the time spent on physical activity, and participation in religious gatherings are the potential mechanisms through which being retired impacts well-being and life satisfaction.

2.2 Introduction and Literature Review

A growing share of the aging population brings up new concerns about the retirement age policies in developed countries. On the one hand, policymakers have been encouraged to raise the normal retirement age because of an increase in life expectancy, the need for the older workforce in the future, and the high ratio of benefits paid by Social Security to what it collects in the payroll tax (Dave et al., 2006). On the other hand, the indirect financial effects of these policies through the health and well-being of the individuals have added a new aspect to the issue, which requires more careful considerations (Gorry et al., 2018). Because, if retirement harms health and well-being, the policies which extend the working time before retirement may be desirable, but if these

improve the health, we should account for the health impacts in the evaluation of the retirement policies which delay retirement.

There is a growing literature in estimating the effects of retirement on health outcomes to address the concerns about pension policies (Kuhn, 2018). The previous studies have found an ambiguous health effect of retirement because of two reasons. The first reason is the complex nature of these effects, which can be negative because of the consequences of job loss and positive due to more time invested in health improvement. The second reason is that different studies have evaluated the effects in a different context, in different countries, and by different methods (Kuhn, 2018). The difference in context is a result of choosing various health outcomes, objective or subjective, and a different way of dealing with the reverse causality problem, which is an inherent methodological issue in the study of health effects of retirement. This issue originates from the fact that the individuals who have experienced health shocks and lower life satisfaction scores tend to retire earlier, confirmed by other studies (Gorry et al., 2018). To this end, this paper looks for a causal relationship between retirement and life satisfaction as an essential measure for health.

As mentioned before, there is a significant body of literature evaluating the retirement effects on many health outcomes. In one of the most recent papers, Shai uses an exogenous increase in the retirement age to prove that employment in older age worsens the health (Shai, 2018). In another research, Coe and Zamarro use the Survey of Health, Aging, and Retirement in Europe[1] to find the effects of retirement on self-reported health, depression, and cognitive ability. They use the country-specific early and full retirement age as instruments to solve the endogeneity problem and find that retirement has a positive impact on overall health (Coe and Zamarro, 2011). Similar results have been found by Atalay and Barrett, again, using the Instrumental Variable method and

[1] SHARE

studying the Australian pension reform (Atalay and Barrett, 2014). On the other side, using a regression discontinuity design, Johnston and Lee find that retirement decreases the tension of work stress in retirees, but it does not necessarily improve their physical health (Johnston and Lee, 2009).

Despite the extensive literature on the health impacts of retirement, only a few papers have studied these effects on life satisfaction. Latif (2011) looks for the relationship between retirement and psychological well-being in Canada while addressing the endogeneity problem using the fixed-effect instrumental variable method. His results show a positive impact of getting retired on psychological well-being. In a more recent study, Zhu and He (2015) try to answer how women's life satisfaction responds to retirement with the help of a two-stage analysis. Their findings illustrate an immediate improvement in women's life satisfaction, but life satisfaction diminishes during retirement (Zhu and He, 2015). In another relevant paper, Gorry and others study the effects of leaving a job on general health and life satisfaction using the Health and Retirement Study data by IV method. They find that retirement can improve life satisfaction in the short term, but it does not change well-being in the long term. (Gorry et al., 2018).

This paper aimed to study the effects of retirement on life satisfaction while addressing the problem of endogenous retirement using a fuzzy Regression Discontinuity Design[1]. This method mimics a random assignment in a non-experimental setting which leads to more reliable results. RDD uses the discontinuous change in the probability of retiring at age 62 and age 71 to deal with the endogeneity problem. The rationale behind this strategy is that the probability of being retired after a specific age increases due to financial incentives, and this discontinuous increase was used as the source of exogenous variation in retirement. The method is similar to the studies that have

[1] RDD

used the eligibility age for social security benefits as an instrumental variable. In order to measure the retirement effects on health outcomes, these studies specify a quadratic age trend on retirement over the entire age range.

In contrast, the current study's method enables the model to allow the age trend to differ on both sides of the threshold since seeing some nonlinearities at old age that might not be captured using a quadratic age trend was expected (Eibich, 2015).

An important contribution of this paper is to find the mechanisms through which retirement affects life satisfaction. There are very few papers that have looked for the mechanism through which retirement impacts health outcomes (Zhang et. all 2018, and Eibich 2015), and there is none that have studied life satisfaction specifically. Eibich (2015) examines the effect of retirement on subjective health status and mental health and provides evidence that the health behaviors, time use, and effect heterogeneity can be the potential mechanisms. Insler (2014) estimates the effects of retirement on subjective health and finds out that health behavior is a mechanism through which retirement can affect health. He shows that with more leisure time after retirement, individuals have more time to spend on healthier behavior (Insler, 2014).

2.3 Theoretical Background

The effect of retirement on health and well-being can be explained theoretically by the human capital model of the demand for health (Grossman, 1972). In this model, Grossman introduces the demand for an individual's health capital by combining the household production model of consumer behavior with the theory of human capital investment. Health is assumed as a consumptive and an investment good, which means it can increase utility directly and raises the earning through more healthy time or lower work loss due to illness (Dave et al., 2006). Therefore, withdrawing from work can decline the motivation to invest in health, which causes more income.

So, we expect health to be worse after retirement. On the other side, the individual might spend more time to improve the health after retirement since the health is directly adding to the individual's utility. In this frame, the health is subject to improve after retiring from the job. The utility function defined by Grossman contains health and other goods consumed by individuals:

(1) $U = U(\phi_t H_t, Z_t)$

Where H_t is health at time t, ϕ_t is the service flow per unit stock of health, $\phi_t H_t$ is the total consumption of health services, and Z_t is the total consumption of other goods. Individuals maximize the utility subject to two constraints of time and income. The time constraint requires the total amount of time available in any period to be exhausted by all uses. Income constraint equates to the present value of money spent on goods to the initial assets plus the discounted value of the earnings (full wealth) as an individual spent all her time at work (Grossman, 1999). This optimization problem leads us to the following first order condition for period t:

(2): $G_t[w_t + (Uh_t/\lambda)(1+r)^t] = \pi_{t-1}(r - \bar{\pi}_{t-1} + \delta_t)$

Where G_t is the marginal product of health capital[1], w_t measures the hourly wage rate, (Uh_t/λ) is the discounted monetary value of the increase in utility caused by a one unit increase in healthy time[2], π_{t-1} represents the marginal cost of gross investment in health capital in time t-1, and $\bar{\pi}_{t-1}$ is the percentage change in the marginal cost between period t-1 and period t. Equation 2 implies that the undiscounted value of marginal product[3] of health capital at period t, left-hand side, should be equal to the supply price of health capital. This equation enables us to find the optimal amount of health capital for any individuals at period t.

[1] Increase in the amount of healthy time due to a one unit increase in the stock of health
[2] λ is the marginal utility of wealth and $Uh_t$is the marginal utility of healthy time?
[3] Or marginal benefit

Investment in health changes health capital. This investment, in turn, depends on time, medical care, etc. Since individuals have more time after withdrawal from their jobs, the effects of retirement depend on the changes in the marginal value of time. If the marginal value of time increases after retirement, the marginal benefit of improved health rises as well. At the same time, the marginal cost of investing in health, such as exercising or visiting a doctor, gets higher when time is more valuable after retiring (Behncke, 2012). Therefore, the health effects of retirement are ambiguous.

There are other ways we expect retirement impact well-being and health. One of them is the relief from the stress caused by work (Bound and Waidmann, 2007). It is because stressful and highly physically demanding jobs could worsen life satisfaction. On the other hand, retirement itself can be a stressful event that decreases physical activity, social interactions, and satisfaction from a rewarding job (Gorry et al., 2018). It seems that we find either a positive or a negative effect of retirement on well-being. Both the direction and magnitude of either of the impacts can be different across individuals. In the end, we can estimate an average effect.

2.4 Data

Data was acquired from the Health and Retirement Study (HRS), a longitudinal panel data that biennially surveys a sample of 20,000 people over the age of 50 in the United States. HRS collects information on the economic, health, marital, and family status, also public and private health support systems for older people. The period that the HRS provides data for is the waves from the year 1999 to the year 2018. Because of the limited availability of the life satisfaction data, not all the waves of HRS could be used. The used waves include the cohort of the Early Baby Boomer, which entered the survey in 2004; the cohort of Middle Baby Boomer, which came first in the survey in 2010, the cohort of the Late Baby Boomer, which entered the survey in 2016, and

finally the last cohort, which entered the survey in 2018. Most of the data used was from the RAND version of the HRS data set, a clean and user-friendly data set, including a subset of data from HRS. However, the life satisfaction variables were merged from the Leave-Behind Questionnaires and Demographic Questionnaires from the HRS raw data.

For this study, a retired individual is defined as a person who has reported to be completely retired. It is because it is assumed that withdrawal from work impacts life satisfaction through behavior adjustment and stress relief from work. Therefore, being partially employed is not considered as retired. The data was restricted to individuals between the ages of 55 and 80. This restriction leaves us with a maximum sample size of 38,509 person-wave observations.

2.4.1 Two Sets of Observations

There are two sets of observations in this study. The first data set is shaped based on a measure of life satisfaction that has been collected from all participants in HRS. This measure asks a question from the individuals on how satisfied they are with life as a whole. This question was converted to a binomial variable being 1 when the individual is completely or very satisfied with life and 0 if otherwise. The observation set includes data for six waves, 2008-2018, which shape the main data set with 72% of individuals in the sample who report they are completely or very satisfied with their lives.

The second set of observations includes the people who have been asked to respond to the life satisfaction questions from Leave-Behind Questionnaires. These questions are given to a rotating, random, 50% of the core panel participants who participated in the enhanced face to face interview (HRS, 2021). The questions on life satisfaction ask the individual how much they agree or disagree with the following five statements:

- "In most ways, my life is close to ideal."

- “The conditions of my life are excellent.”
- “I am satisfied with my life.”
- “So far, I have gotten the important things I want in life.”
- “If I could live my life again, I would change almost nothing.”

A binomial measure was made for each statement, setting it equal to “1” if the answer is “strongly agree” and “0” otherwise. Then, the scores for all the statements were added and built a measure for a degree of life satisfaction, ranging from 5 being the highest level of satisfaction to 1, which is the lowest level of satisfaction. This information set contains eight waves of data, 2004-2018, which forms a data set with about 19,688 person-wave observations.

The other dependent variables in this study included the measures for health behaviors and time use. HRS offers various variables to measure health behavior in different waves. Data on smoking, alcohol consumption, physical activity, BMI, and sleep were used. In this study, smoking status was captured with a dummy variable coded as “1” if the individual smokes now and “0” if not. Alcohol consumption was measured with two variables. The first one asks if the individual does not drink alcoholic beverages. The second one asks about the number of days per week the person drank in the last three months. The second variable was redefined a dummy variable that measures if the individual consumes alcohol regularly. This measure was defined equal to “1” if the respondent drinks three or more days a week. One variable was used to determine the amount of physical activity for each observation. This variable measures the frequency of moderate physical activity per week. The information was converted to a dummy variable, which is “1” if individual exercises every day or more than once a week and “0” if otherwise. The data on an individual's BMI was collected directly from RAND version of HRS. To estimate sleep behavior, one variable was used. The measure asks about how often individuals feel rested when they wake

up in the morning. The binary variable was “1” if they answer “most of the time” and “0” otherwise.

Various questionnaires were used from HRS core data to evaluate the time individuals spend on different activities. This allowed using the time respondents spent with other people as a proxy for social interactions. To capture this, the frequency of individuals attending religious services was observed and was coded as a dummy variable, with “1” indicating attendance of religious services at least once a week and “0” if otherwise. The last variable asks individuals if they spend 100 hours or more taking care of their grandkids, which was coded “1” for yes and “0” otherwise. The sample sizes for each outcome in each data set are shown in Table 1 and Table 2.

Other variables were included to estimate heterogeneous effects like gender, race, marital status, education, and income. The gender was coded as "female" was “1” if an individual was female and “0” otherwise. There was a dummy variable for being white, which was “0” when the race was not white. Marital status was measured by a binomial variable, which was “1” if an individual was married and “0” otherwise. For education, a dummy variable was used, which was equal to “1” if a person had a BA or higher degree and “0” if she did not.

This study used the age of the eligibility for Social Security and private defined-benefit plan as an assignment for getting retired. The considered age thresholds in this study are 62, when the individuals are eligible for early but reduced social security benefits, 65[1] which is the official age of eligibility for a complete benefits package, and 71[2] which is the age that retirement after will lead to receive more than 30% higher monthly payment.

[1] It can be 65 or 66 based on the year individuals are born, but since I am using the probability of getting retired, it won’t hurt my general results.
[2] This also can be 70 to 71.

2.5 Econometric Model

2.5.1 Endogeneity

Two important issues must be considered while estimating the causal effects of retirement on life satisfaction and health. The first one is that the individuals' decision to retire and health status depends on some unobserved variables, omitted variable bias. That is why the level of life satisfaction for individuals with different life histories and time preferences can be different (Dave et al., 2006). To address this problem, the unobservable, time-invariant individual characteristics were controlled by using a fixed-effect panel data model.

The second issue is that deciding to get retired can be because of well-being or health status before getting retired. This kind of endogeneity of retirement has been studied and proved by many papers (Mandal & Roe, 2007). People who have a lower degree of satisfaction in their lives tend to retire as soon as possible (Celidoni et al., 2017). Also, negative shocks to mental health may force an individual to decide to withdraw from her job earlier than others (Dwyer & Mitchell, 1999). This endogeneity of retirement means that we cannot find a causal relationship only just by comparing life satisfaction before and after retirement. To eliminate this problem, a Regression Discontinuity Design was used.

2.5.2 Regression Discontinuity Design

RDD uses the rules which determine if an individual belongs to a treatment group, retired, or not. To use this design, we need an assignment variable that establishes whether an individual belongs to the treatment group or not. Observations above the threshold are in the treatment group, while the individuals below the threshold are not. Then, if there is a discontinuity in outcome under some minor assumptions, we can explain it as a causal impact of treatment (Eibich, 2015).

In this paper, age was used as an assignment for being retired. Because of financial motives, the eligibility age for pension benefits can be an exogenous source of variation in deciding to get retired. Three age thresholds were considered as potential threshold in this study; 62, 65, and 71. Since the age is 65 (62 if an individual decides to retire early) to be eligible to benefit from social security pays, individuals cannot get paid by reduced or full amount of pension before the threshold of 62. Before age 65, an individual can get a partial amount as long as she retires after age 62. Age 71 is also considered as another threshold since individuals who delay their retirement till 70 gets delay credit monthly on their paycheck. 70 is the age they get their maximum credit. Therefore, holding the retirement till age 71 seems financially desirable. This implies that retirement is not completely determined by the threshold of age at 62, 65, or 71. Alternatively, the probability of withdrawal from work rises discontinuously at these age thresholds for full retirement. Therefore, [1]fuzzy Regression Discontinuity Design was used in the analysis of this study. The estimated effect would be a local average effect (LATE). The estimated parameter shows the average impact of retirement on the life satisfaction of the individuals who are near age thresholds who would retire if their age moved from just below 62, 65, or 71 to just above it.

The fuzzy RDD method that was used has some advantages in comparison to alternative approaches. Some papers have used simple fixed-effects models to solve the selection problem (Bonsang & Klein, 2011; Dave et al., 2006). The issue with this method is that it addresses just the selection on time-invariant observable factors. In contrast, some factors, such as shocks to well-being and health, which motivate individuals to get retired, are time-variant unobservable. A few

[1] RDD has two styles, sharp and fuzzy. Researchers use sharp RDD when treatment status is a deterministic and discontinuous function of a predictor. In other words, if individuals are over an exact threshold on the covariate, they are treated. On the other hand, fuzzy RDD is suitable to use when the predictor does not deterministically show the treatment status. Instead, the probability of being treated is a function of the variable and jumps at the threshold (Angrist & Pischke, 2008).

of these papers exploit a solution of restricting observations to people with no severe illness or no issues with well-being. This approach can resolve the issue while sacrificing the unbiased results, which can happen due to the misreporting of the criteria to exclude the individuals. On the contrary, the RDD approach requires two assumptions, which will be mentioned under the assumption part later, that guarantee the unbiased results. Another way to address the endogeneity problem used by a few papers is to consider policy reforms as an exogenous source of variation for retirement (Shai, 2018; Hallberg et al., 2014; Blake & Garrouste, 2012). The problem with these studies is that most of the time, the reforms are applied to a subgroup of individuals and not on everyone, e.g., the army employees in Halbert et al., 2014, or the male workers in Shai, 2018. In contrast, the RDD method provides reliable analysis for all individuals, regardless of their gender, job section, or situation (Eibich, 2015; Chen et al. 2020; Picchio et al., 2020).

Fuzzy RDD has very similarities to studies that use the eligibility age as an instrumental variable to control for endogeneity of retirement (Gorry et al., 2018; Zhu &He, 2015; Celidoni et al., 2017; Co & Zamarror, 2011). They both require the probability of being retired to raise discontinuously at the eligibility age. However, there is an essential difference in the way these methods specify the models. Studies using the IV approach mostly define a quadratic age trend for the entire age range. On the contrary, the RDD approach suggests letting the age trends to be different on either side of the threshold (Eibich, 2015). This flexibility will capture the non-linearities in older ages that have not been properly considered in the models specified by studies that use an IV approach based on eligibility age.

2.5.3 Setup

The discontinuity increases in the probability of retirement at 62, 65, and 71 was used as a source of exogenous variation in the retirement decision. First, we need to check the validity of

this claim that the probability of being fully retired increases discontinuously at the age of 62, 65, and 71. Figure 1 shows the share of retired individuals by age in my main data set. There is a discontinuity in the percentage of retirees at ages 62, 65, and 71. Although more than 23% are retired before age 62, close to 60% are fully retired after age 65. At age 62, early retirement age, the probability of being retired increases by 40% points. Another discontinuity is observable when moving from age 70 to 71, 11% points, which is lower than an increase in retirees' share at age 62. Yet, it's still considerably higher than change at other ages around 71 (comparing to a 2% decrease in the proportion of retirees from age 69 to 70 and 1% decrease from 71 to 72). Although the retired individual's percentage increases in general after 62, there is not an observed sharp increase till 71. At 65 the proportion of retirees changes smoothly (12% compared to 9% increase in retirees from age 63 to 64 and 4% after 65, age 65-66). Therefore, 65 will not be used for my main analysis. However, since there is a continuous increase in the percentage of retired individuals from 63 to 65, the results of using all three ages will be reported as a robustness check.

Three main assumptions are required for RDD estimation to be valid. First, it is assumed that life satisfaction is a continuous function of the forcing variable (age). This assumption is reasonable since the aging-life satisfaction process should be smooth based on aging being a gradual process (Eibich, 2015). The second assumption is that individuals do not exert control over the value of the assignment variable. This assumption holds by the way data is constructed since the age of individuals is not self-reported. Instead, it is calculated by their date of birth reported in the data set. The last assumption requires the groups around the threshold are exchangeable. In other words, individuals close to the cut-off point should be similar, and the predetermined variables would be continuous over the assignment variable. If they are discontinuous over the age

variable, it can cause doubts over my strategy since the effects could be based on unobservable factors.

The last assumption was tested by checking discontinuity in baseline characteristics. Figure 2 shows the results for four main variables (income, marital status, education, and gender). These graphs can be seen as placebo tests. For example, retirement affects income, so it is expected to see discontinuity in income. But for a variable like marital status, the effect is not apparent. Because it is not predetermined, it can be an outcome of withdrawal (Eibich, 2015). Education and share of women in the sample are the predetermined measures that should not be discontinuous at the threshold since they are not affected by retirement.

As in figure 2, there is a small discontinuity in the log of income after age 62 and no jump after 71. In the case of marital status, there is no noticeable jump. It appears that there is no sharp change in probability of having a BA or higher degree around the threshold of 62 or 71, as we expected. For the women share in the sample, there is a greater variety at older ages, but still, there is no visible discontinuity around cut-off points.

The study also looks for discontinuities in the graphs of scatterplot for independent variables over the age variable (figure 3). A sharp positive change is observable on life satisfaction around age 62 and after age 71. There is also a smaller positive discontinuity in zero drinks, moderate physical activity, sleep quality, grandchild care, and participation in religious service. Except for life satisfaction, overall, the discontinuities look small. But this fact should be considered that the probability of being retired at the threshold increase by 20-30%. So, to estimate the local Average Treatment Effect, the discontinuities in outcomes are weighted by the increase in the probability of retirement (Eibich, 2015). Since these variables are considered potential mechanisms that affect

life satisfaction simultaneously, it is highly expected to observe a lower jump in outcomes over these variables.

2.5.4 The Model

The main models with two discontinuities at age 62 and 71 are presented:

$$retired_{it} = \alpha_0 + \alpha_1 age_{it} + \alpha_2 age62_{it} + \alpha_3 age_{it} * age62_{it} + \alpha_4 age71_{it} + \alpha_5 age_{it} * age71_{it} + g_i + \Omega_t + \varepsilon_{it}$$

$$life\ Satisfaction_{it} = \beta_0 + \beta_1 age_{it} + \beta_2 age_{it} * age62_{it} + \beta_3 age_{it} * age71_{it} + \pi\, \overline{retired_{it}} + c_i + \delta_t + u_{it}$$

In these equations, $age71_{it}$ is the binary variable that is equal to "1" if an individual i in year t is 71 or older and "0" otherwise, $age62_{it}$ is also a binary variable that is "1" if the individual age is the interval of 62<=age<71 in year t. The variable $retired_{it}$is a dummy variable that is "1" if the individual i is fully retired in year t, and $\overline{retired_{it}}$ is the predicted values of treatment from the first stage. Variables g_i and c_i includes the individual-fixed effects, Ω_t and δ_t are the wave dummy variables, and finally ε_{it} and u_{it} are error terms for the first and second stage. The interaction terms allow for a different age trend after each threshold. The analysis was done using STATA 17.

2.6 Results:

2.6.1 The Effects of Retirement on Life Satisfaction

Table 3 shows the results of the first model to estimate the effects of retirement on life satisfaction. The results are displayed for both measures of life satisfaction in two main parts of the table. The first life satisfaction measure is in the primary sample (all participants in HRS), and the second measure belongs to the Leave-Behind Questionnaires sample. First column is the

estimates from my main model for the first measure, which restricts the sample to the individuals who are older than 55 and younger than 80. Second column shows the results for second measure.

The estimated treatment effects in the first column of table 3, restricted sample, suggest that retirement has a positive impact on life satisfaction. Being fully-retired increases the probability of individuals being completely or very satisfied with their lives. As the table points out, both of time-variant controls (income and marital status), are positively correlated with higher life satisfaction.

For the other measure in the second sample, the sign of the retirement coefficient is still positive, but it is not significant. This lack of significance can be because of different levels of questions that have been included in the second measure. To check this possibility, the same model was estimated for every five questions separately. Table 7 in the appendix shows the results separately for each question. The results show a positive relationship between getting retired and three satisfaction questions, if life is close to ideal, if there are excellent conditions in life, and if individuals are satisfied with their lives.

All models in the tables contain the linear age trend, individual fixed effects, and dummy variables for the waves when the interviews have been conducted. The Kleibergen-Paap statistic was mentioned to check for a weak instrument. Their p-values show that the null hypothesis for this test is rejected and imply that the discontinuities are jointly significant in the determination of retirement status.

2.6.2 Mechanisms

In this section, the study looks for the potential mechanisms through which a withdrawal from a job impacts well-being. As it was discussed before, retirement could influence life in many ways.

It changes the opportunity cost of investment on health and well-being, and it increases the amount of leisure time available. The way individuals live and use their time can change the level of their life satisfaction. Thus, it seems possible that retirement affects life satisfaction indirectly through health behavior and how an individual spends her time.

To check the hypothesis of retirement changing the way of living and time use, the effects of retirement on individuals' health behavior were estimated. Table 4 shows the results. Each column represents the results of one separate regression model, where retirement is the dependent variable, and the measures of health behavior and time use are the independent variables. As shown in table 4, retirement is not significantly associated with current smoking. The effects of withdrawal on drinking no alcohol, while positive, is insignificant. The probability of being a regular drinker, drinking three or more days a week, increases by 0.7% points after retirement. Although it has been shown that drinking harms health, it is important to consider that the effect can be more complicated in the case of life satisfaction, since it is shown that consuming 3-4 more drinks can increase life satisfaction (Krekhovets & Leonova, 2013). Physical activity increases significantly after getting retired by about 0.03 standard deviations. BMI increases after retirement significantly. Retirement increases the probability of being well-rested after waking up in the morning by 3.2% points, which is the second strongest effect of getting retired on health behavior. After checking for the retirement effects on health behavior, the impact of retirement on time spent was estimated. As table 4 specifies, participating in religious services increases sharply after retirement. Finally, being out of the labor force does not change the time individuals spend to take care of their grandkids.

These results have some specification problems since there is a chance of reverse causality. Retirement can be affected by any of the health behaviors and the way individuals spend their time.

To partially eliminate this problem, Eibich's method (2015) was used in finding the mechanisms through which retirement affects physical health measures. The same RDD method from the previous section was used, with the health behavior and time use measures as control variables in the model. Then the retirement coefficient from this regression was compared to the model without health behavior and time use[1]. We expect to observe that the coefficient gets smaller with including the controls for potential mechanisms.

Table 5 illustrates the models' results with and without including the indicators for health behavior and time use. The coefficient of retirement is larger compared to the results from table 3 (0.34 compared to 0.27), where a larger sample that considers all individuals was used, regardless of whether data on mechanism variables are available. This coefficient gets smaller, the second column when the health behavior and time use variables were added to the model, 0.28 compared to 0.34. As table 5 shows, changing physical activity habits from zero or once a week to more than once a week, increasing sleep quality, and increasing participation in religious service have a positive effect on life satisfaction (at least those parts that are caused by retirement). These estimates imply that these variables can be the ways that retirement influences the level of life satisfaction.

2.6.3 Robustness Checks Using Another Age Threshold

For robustness checks, the first models were estimated using a third age threshold (65 age). Table 6 shows the first model results to estimate the effects of retirement on first life satisfaction measures in the primary sample (all participants in HRS) with considering two age thresholds of 62 and 65 in the first column. Second column shows the results for the model with age thresholds

[1] With sample limited to individuals whose data on health behavior and time use is available.

62, 71, and 65. As it is shown, the results did not change much when we considered the third age threshold.

2.7 Conclusion

The study results display evidence for the positive effects of getting retired on the individuals' life satisfaction. This is an important result for policy makers since it has been shown that life satisfaction is related to overall health. Although increasing the retirement age directly leads to a lower cost to society, it might add more indirect health expenses. Therefore, it is critical to account for these indirect costs when deciding about a new retirement plan. Instead of decreasing the costs of retiring late directly, we can improve the mechanisms through which retirement can change the level of life satisfaction, which takes us to the second part of this paper.

The most important contribution of my paper is finding the potential paths through which getting retired changes life satisfaction. These findings show an increased physical activity, sleep quality, and social interaction can increase life satisfaction. These channels can be aimed by policies as possible ways to mediate the indirect costs of late retirements to the individuals and the society. There is a need for future studies to examine the frequency and effectiveness of policies that encourage the working population's healthy behavior after they reach the age of 60 or older.

2.8 References

Atalay, K., & Barrett, G. F. (2014). The causal effect of retirement on health: New evidence from Australian pension reform. *Economics Letters*, *125*(3), 392-395.

Angrist, J.D. and Pischke, J.S., 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

Behncke, S., 2012. *Does retirement trigger ill health? Health economics*, 21(3), pp.282-300.

Blake, H. and Garrouste, C., 2019. *Collateral effects of a pension reform in France*. Annals of Economics and Statistics, (133), pp.57-86.

Bonsang, E. and Klein, T.J., 2010. *Retirement and subjective well-being*. Journal of Economic Behavior & Organization, 83(3), pp.311-329.

Bound, J., & Waidmann, T. (2007). Estimating the health effects of retirement. *Michigan Retirement Research Center Research Paper No. UM WP*, *168*.

Celidoni, M., Dal Bianco, C. and Weber, G., 2017. *Retirement and cognitive decline. A longitudinal analysis using SHARE data*. Journal of health economics, 56, pp.113-125.

Chen, S., Geldsetzer, P., & Bärnighausen, T. (2020). The causal effect of retirement on stress in older adults in China: A regression discontinuity study. SSM-population health, 10, 100462.

Coe, N.B. and Zamarro, G., 2011. *Retirement effects on health in Europe*, Journal of health economics, 30(1), pp.77-86

Dave, D., Rashad, I. and Spasojevic, J., 2006. *The effects of retirement on physical and mental health outcomes* (No. w12123). National Bureau of Economic Research.

Dwyer, D.S. and Mitchell, O.S., 1999. *Health problems as determinants of retirement: Are self-rated measures endogenous?* Journal of health economics, 18(2), pp.173-193.

Eibich, P., 2015. *Understanding the effect of retirement on health: Mechanisms and heterogeneity*. Journal of health economics, 43, pp.1-12.

Gorry, A., Gorry, D. and Slavov, S.N., 2018. *Does retirement improve health and life satisfaction?* Health economics, 27(12), pp.2067-2086.

Grossman, M., 1999. *The human capital model of the demand for health (No. w7078*). National Bureau of Economic Research.

Grossman, M., 1972. *On the concept of health capital and the demand for health*. Journal of Political economy, 80(2), pp.223-255.

Hallberg, D., Johansson, P. and Josephson, M., 2014. *Early retirement and post-retirement health.* IZA World of Labor, IZA Discussion Paper No. 8260.

Insler, M., 2014. *The health consequences of retirement.* Journal of Human Resources, 49(1), pp.195-233.

Krekhovets, E. and Leonova, L., 2013. *Alcohol consumption and life satisfaction: Evidence from Russia*. Academic Journal of Interdisciplinary Studies, 2(8), pp.98-98.

Kuhn, A., 2018. *The complex effects of retirement on health*. IZA World of Labor.

Latif, E., 2011. *The impact of retirement on psychological well-being in Canada*. The Journal of Socio- Economics, 40(4), pp.373-380.

Mandal, B. and Roe, B.E., 2007. *Job loss, retirement and the mental health of older Americans*. Retirement and the Mental Health of Older Americans (June 2007).

Picchio, M., & van Ours, J. C. (2020). Mental health effects of retirement. De Economist, 168(3), 419-452.

Shai, O., 2018. *Is retirement good for men's health? Evidence using a change in the retirement age in Israel.* Journal of health economics, 57, pp.15-30.

Johnston, D.W. and Lee, W.S., 2009. *Retiring to the good life? The short-term effects of retirement on health.* Economics Letters, 103(1), pp.8-11.

Zhang, Y., Salm, M., & van Soest, A. (2018). The effect of retirement on healthcare utilization: Evidence from China. *Journal of health economics*, *62*, 165-177.

Zhu, R. and He, X., 2015. *How does women's life satisfaction respond to retirement? A two-stage analysis*. Economics Letters, 137, pp.118-122.

2.9 Tables and Figures[1]

Table 2-1 Summary Statistics, data set 1

Variable	N	Mean	SD	Min	Max	Mean non-retirees N=20,086	Mean retirees N= 28,052
A. Main variables							
Retired	38509	0.560	0.496	0	1	-	-
Life satisfaction	38509	0.721	0.449	0	1	0.708	0.730
B. Health behavior							
Smoking	38283	0.099	0.299	0	1	0.101	0.098
No alcohol	38509	0.441	0.497	0	1	0.405	0.470
Drinking often	38509	0.188	0.391	0	1	0.196	0.183
Physical activity	38509	0.701	0.458	0	1	0.667	0.754
BMI	38097	28.85	5.898	10.3	92.8	28.790	28.896
Sleep quality	37516	0.570	0.495	0	1	0.560	0.578
C. Time use							
Religious services	37449	0.435	0.496	0	1	0.418	0.449
Grandkid care	30648	0.321	0.467	0	1	0.379	0.280
D. Covariates							
Age	38509	68.65	6.72	55	80	65.49	71.136
Female	38509	0.615	0.487	0	1	0.606	0.623
White	38509	0.793	0.405	0	1	0.783	0.8
Married	38509	0.63	0.483	0	1	0.664	0.604
College BA	38509	0.262	0.44	0	1	0.299	0.233
Income	38509	73214.249	126913.3	0	7406316	92321.362	58179.288

[1] The tables and figures are listed in the order they appear in the chapter.

Table 2-2: Summary Statistics, data set 2

Variable	N	Mean	SD	Min	Max	Mean non-retirees	Mean retirees
A. Main variables							
Retired	19688	0.542	0.498	0	1	-	-
Life satisfaction	19688	1.162	1.608	0	5	1.095	1.219
B. Health behavior							
Smoking	19574	0.097	0.295	0	1	0.102	0.092
No alcohol	19688	0.434	0.496	0	1	0.409	0.454
Drinking often	19688	0.193	0.395	0	1	0.196	0.191
Physical activity	19688	0.721	0.449	0	1	0.691	0.756
BMI	19499	28.892	5.886	10.9	92.8	28.833	28.942
Sleep quality	19676	0.590	0.492	0	1	0.580	0.599
C. Time use							
Religious services	19658	0.445	0.497	0	1	0.426	0.464
Grandkid care	16054	0.351	0.477	0	1	0.401	0.313
D. Covariates							
Age	19688	68.046	6.746	55	80	64.949	70.663
Female	19688	0.611	0.488	0	1	0.609	0.613
White	19688	0.83	0.376	0	1	0.825	0.834
Married	19688	0.667	0.471	0	1	0.693	0.646
College BA	19688	0.266	0.442	0	1	0.299	0.239
Income	19688	75056.9	149718.8	0	13569371	92717.8	60130.4

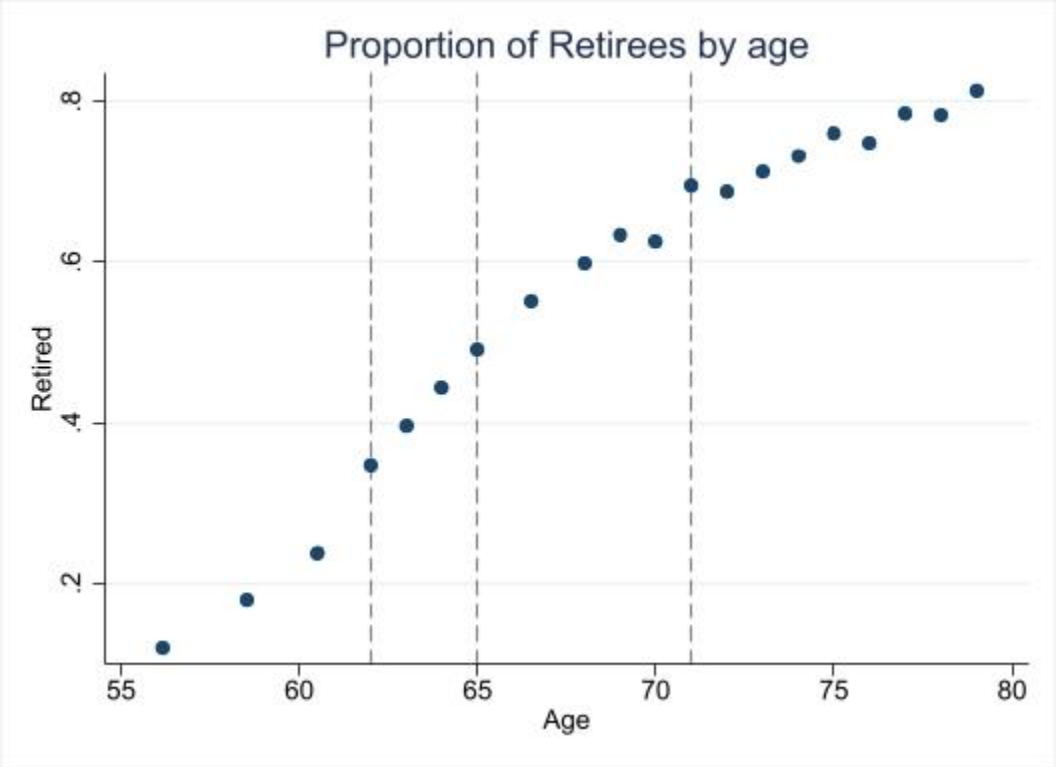

Figure 2-1: Share of individuals retired at given age, main data set

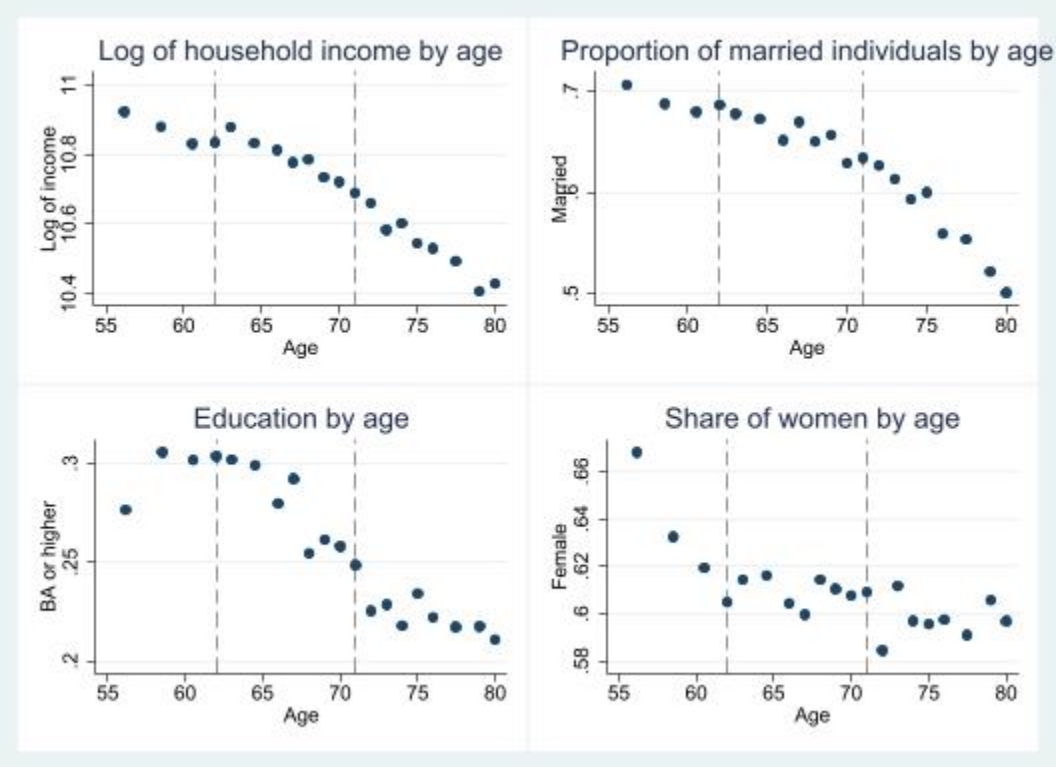

Figure 2-2: Main control variables by age

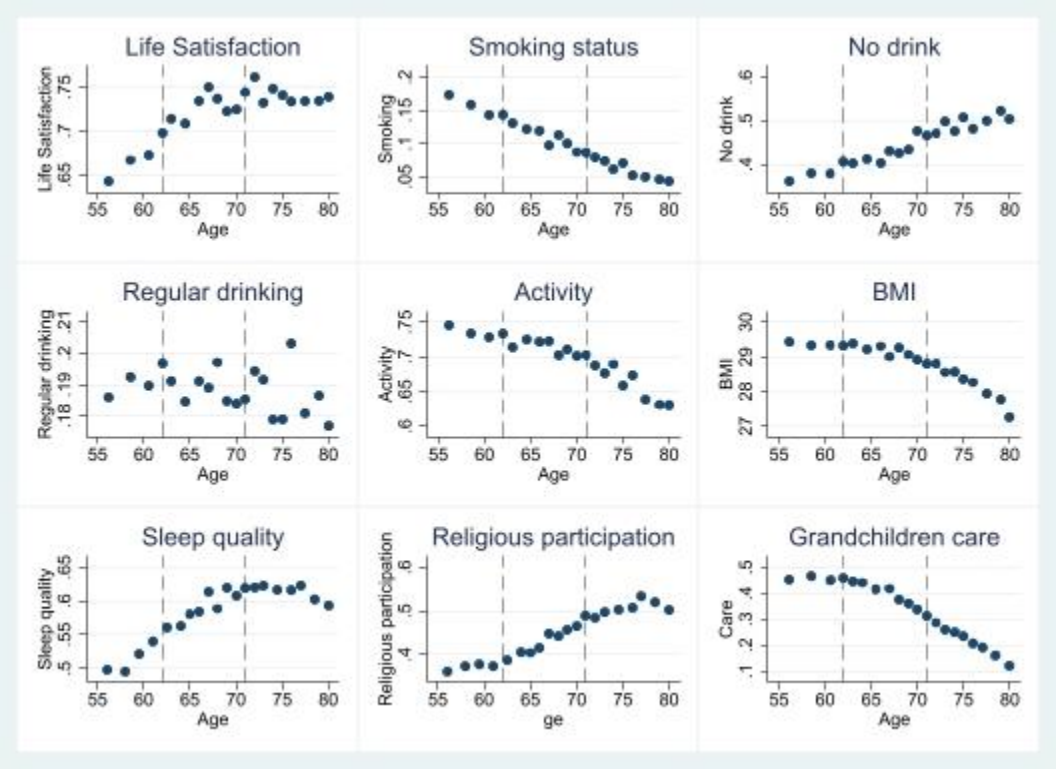

Figure 2-3: Outcome variables over age

Table 2-3: Regression discontinuity estimates

	First Life Satisfaction measure	Second life satisfaction measure
VARIABLES	Main sample 55-80	Second sample 55-80
Retired	0.272***	0.327
	(3.348)	(0.877)
Kleibergen-Paap statistic	74.73	94.79
	0.0001	*0.0001*
Control variables		
Income	0.025***	0.065*
	(3.991)	(1.819)
Married	0.101***	0.168***
	(7.036)	(3.205)
Age	-0.018***	0.033
	(-3.065)	(1.056)
Observations	38,205	19,586

Notes: Robust z-statistics in parentheses. P-values are in italics. All main models include a linear age trend, individual fixed effects, and dummy variables for waves.
* Significance for p<0.1
** Significance for p<0.05
*** Significance for p<0.01

Table 2-4 Estimated effects of potential mechanisms on retirement

	Smoking	No alcohol	Drinking often	Physical activity	BMI	Sleep quality	Religious service	Grandkid care
Retired	-0.002	-0.001	0.007**	0.011**	0.099***	0.032***	0.011**	0.004
	(-0.821)	(-0.147)	(1.987)	(2.571)	(2.625)	(5.537)	(2.510)	(0.598)
Change in standard deviation	-0.030	0.008	-0.013	0.078	0.022	0.003	0.005	-0.061
Observations	49,179	49,481	49,481	49,481	48,935	49,411	49,348	40,202

Notes: Robust t-statistics in parentheses. All models include a linear age trend, individual fixed effects, and dummy variables for waves.
* Significance for p<0.1
** Significance for p<0.05
*** Significance for p<0.01

Table 2-5: RDD estimates for first life satisfaction measure

VARIABLES	Excluding mechanisms	Including mechanisms
retired	0.340***	0.285***
	(3.893)	(2.676)
Drinking often	-	0.002
		(0.149)
Physical activity	-	0.038***
		(4.403)
BMI	-	-0.001
		(-1.162)
Sleep quality	-	0.042***
		(6.230)
Religious service	-	0.026***
		(2.609)
Observations	29,354	29,354

Notes: Robust z-statistics in parentheses. All main models include a linear age trend, individual fixed effects, and dummy variables for waves.
* Significance for p<0.1
** Significance for p<0.05
*** Significance for p<0.01

Table 2-6: Results for models with two age thresholds vs models with three age thresholds

VARIABLES	(1) Main dataset with age thresholds 62 and 71	(2) Main dataset with age thresholds 62 71 and 65
Retired	0.272***	0.301**
	(3.348)	(2.317)
Income	0.025***	0.027***
	(3.991)	(3.001)
Married	0.101***	0.100***
	(7.036)	(6.955)
Age	-0.018***	-0.018***
	(-3.065)	(-2.868)
Observations	38,205	38,205