THREE ESSAYS IN HEALTH ECONOMETRICS

by

Mohamed Ebeid

A Thesis Submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Economics

University of Manitoba

Winnipeg

EXAMINING COMMITTEE

Advisor:	Dr. Julia Wit Economics University of Manitoba
Co-Advisor:	Dr. Umut Oguzoglu Economics University of Manitoba
Internal Examiner:	Dr. Evelyn Forget Economics University of Manitoba
Internal Examiner:	Dr. Mahmoud Torabi Community Health Sciences University of Manitoba
External Examiner:	Dr. Stefan Staubli Economics University of Calgary

DEDICATION

To Mom and Dad, it is impossible to thank you adequately for everything you have done. Thank you for raising me in a tolerant and peaceful family. I could not have asked for better role-models.

To my wife for her kindness and endless support. Your selflessness will always be remembered.

And lastly, to my professor Julia Witt. Thank you for your support

ABSTRACT

This thesis consists of three essays that address questions in health economics using different datasets and econometric approaches. In the first essay, I apply novel non-parametric econometric techniques to estimate the causal effect of retirement on health using the U.S. Health and Retirement Study (HRS) survey. I use a non-parametric Fuzzy Regression Discontinuity Design (RDD) technique for my analysis to avoid restrictive assumptions on a particular functional form and to capture the potential reverse causality from health to retirement (endogeneity issue) by exploiting the exogenous variation in retirement decisions induced by U.S. pension eligibility ages at 62 and 65. The results show that retirement is associated with an 8% decline in the cognitive functioning score of retirees, and 0.42 points increase in the CESD depression scale. Retirees also are 13 percentage points less likely to report good general health status, 8.8 percentage points less likely to be drinkers, and they are 4 percentage points less likely to consume alcohol more than three times per week. In the second essay, I use an administrative database to investigate the impact of the timing of first exposure to maternal depression on a comprehensive measure of children's school readiness that incorporates multidimensional developmental domains that underlie school class adaptation and later success. I find that exposure to maternal depression is associated with developmental vulnerability in emotional, physical, social, and cognitive domains. The strongest adverse effects on development are from exposure to depression during pregnancy, followed by exposure during the preschool period. In the third essay, I examine the impact of Type I Diabetes Mellitus (TIDM) during childhood on educational attainment and labor market outcomes in adulthood using the National Health Interview Survey (NHIS). The results show that individuals who developed Type I diabetes early in life are 7 to 17 percentage points less likely to be employed, work fewer hours (3 to 11 hours less per week) and are 5 to 10 percentage points more likely to receive social welfare assistance than non-diabetics. In addition, Type I diabetics experience less educational attainment than non-diabetics.

INTRODUCTION

This thesis consists of three essays that address questions in health economics using different datasets and econometric approaches. In the first essay, I apply novel non-parametric econometric techniques to estimate the causal effect of retirement on health status and health-related behavior using the last ten waves of the U.S. Health and Retirement Study (HRS) survey from 1996-2014. I use a non-parametric Fuzzy Regression Discontinuity Design (RDD) technique to avoid restrictive assumptions on a particular functional form and to capture the potential reverse causality from health to retirement (endogeneity issue) by exploiting the exogenous variation in retirement decisions induced by U.S. pension eligibility ages at 62 and 65. I check the robustness of the nonparametric RD estimates in the analysis of the effect of retirement on health using a variety of distinct procedures. I demonstrate that the RD estimates of the effect of retirement do not change across different weighting schemes that give higher weight to the observations in the neighborhood of the cut-off point but away from the cut-off. I also check the robustness of the results under different polynomial procedures, linear and quadratic, and there is no significant change in the results. The robustness of the regression is verified by estimating the effect of retirement at different bandwidths around the cut-off point, i.e., RD treatment effect estimated for 50 to 200 percent of the Cross-Validation optimal bandwidth (half and twice of the optimal bandwidth) with 5 percentage point increments. To check the validity of the RD, I run a procedure called the "donut hole" approach, which excludes observations that are very close to the cut-off point to investigate how sensitive the results are to this exclusion. The RD estimation and inference analysis are then repeated using the remaining sample. I also run a falsification test to check for the continuity of the running variable and the sensitivity of the RD to different methods of the

chosen bandwidths and show that the estimates were robust to different bandwidths and weighting kernel functions. I perform two tests to see what will happen to the estimated RD treatment effect when the value of the running value is a little bit further away from the cut-off point, which is known as the Treatment Effect Derivative (TED). This is another novel approach in this chapter. Finally, the parametric fuzzy RD estimation is introduced as a sensitivity check for using more information from the observations far away from the neighborhood of the cut-off point and the results were consistent. In general, I find that retirement has a significant negative impact on physical and mental health status. The non-parametric fuzzy RD results show that the significant jump in retirement probability is associated with an 8% decline in the cognitive functioning score of retirees, and 0.42 points increase in the CESD depression scale. Retirees also are 13 percentage points less likely to report good general health status. Regarding health-related behavior, retirees are 8.8 percentage points less likely to be drinkers relative to non-retirees and they are 4 percentage points less likely to consume alcohol more than three times per week. However, they are 3 percentage points more likely to be smokers relative to non-retired individuals. Furthermore, retirement has a significant heterogeneity effect across socioeconomic and retired gender groups.

The second and third essays address questions related to the potential short and long-term consequences of reduced childhood human capital formation, including both cognitive and non-cognitive skills. These two essays focus on two potential disruptions to the formation of human capabilities, exposure to maternal depression (essay 2), and early childhood-onset Type I diabetes (essay 3). According to the human capital formation model, individuals are born with heterogeneous endowments of capabilities (e.g., cognitive capabilities, non-cognitive capabilities, and stock of mental and physical health at birth) and the technology of skill formation at any point in the child's life is affected by skills attainment in previous stages. Some studies show that the

most sensitive period of child development spans from the time of conception through the fifth birthday (Houston, 2014; Georgiadis et al., 2016). It is characterized by rapid physical, motor, cognitive, and socio-emotional development (Britto et al., 2011; Couperus & Nelson, 2006), which allows cognitive and socio-emotional abilities to start developing and establishes the foundation for a child's short and long-term success (Deoni et al., 2015; Knudsen, Heckman, Cameron, & Shonkoff, 2006). In addition, some studies show that human capital stock during childhood predicts adult human capital and earnings (Currie and Thomas 1999; McLeod and Kaiser 2004). Therefore, disruption to skills formation in early childhood play an important role at school entry age and in the formation of human capabilities in early adulthood.

In the second essay, I investigate the impact of the timing of first exposure to maternal depression on a comprehensive measure of children's school readiness that incorporates multidimensional developmental domains that underlie school class adaptation and later success. Using the administrative database at the Manitoba Center for Health Policy, 59,413 children with an Early Development Instrument (EDI) score are linked to their mothers and followed over time from five years before the child's birth to the child's 5th birthday. The focus on the first exposure to maternal depression is necessary to help isolate the effects resulting from each exposure period and to control for the issue of overlapping repeated depression episodes. I assign each mother-child dyad to 1 of 4 mutually exclusive exposure timing categories. Specifically, these are: children not exposed to maternal depression at any point (reference group); children first exposed to maternal depression during pregnancy; children first exposed to maternal depression during the postnatal period (birth - 12 months); children first exposed to maternal depression during the toddler period (12 months - 36 months); and children first exposed to maternal depression during the preschool period (36 months - 60 months). I find that exposure to maternal depression is

associated with developmental vulnerability in emotional, physical, social, and cognitive domains. When controlling for health (measured by major Adjusted Diagnosis Groups (ADGs) and minor ADGs, and hospital admission frequency) of the child at birth, and through early childhood, and mother's health prior to pregnancy, however, the effects of exposure to maternal depression on children's abilities in the emotional, physical, and social domains were attenuated across the different exposure periods. This effect almost disappeared for cognitive and communication skills. That is, although maternal depression is a risk factor for children's school readiness, children's health and socioeconomic adversity remained an important factor for early child development. Exposure to depression during pregnancy has the strongest effect on developmental vulnerability, followed by the preschool period. Emotional maturity is the most sensitive domain across the different exposure periods. In contrast, cognitive and communication domains are the least sensitive to depression. Finally, there is gender and marital status heterogeneity in the effect of maternal depression on the emotional, physical, and social domains. These findings underscore the need for early detection of maternal depression, ideally by obstetricians during pregnancy, and in programs that focus on the mother and child together. Intervention programs should commence prior to the start of school to mitigate early developmental difficulties, which exacerbate if they are not addressed.

In the third essay, I examine the impact of Type I Diabetes Mellitus (TIDM) during childhood on educational attainment and labor market outcomes in adulthood using the last ten waves of the National Health Interview Survey (NHIS). Individuals are identified as Type I diabetics during childhood if they were diagnosed for the first time with diabetes at age less than 15 years and have been taking insulin. To avoid selection bias, a Tobit model is used to account for zero working hours and earnings. The results show that individuals who developed Type I

diabetes early in life are 7 to 17 percentage points less likely to be employed, work fewer hours (3 to 11 hours less per week) and are 5 to 10 percentage points more likely to receive social welfare assistance than non-diabetic individuals. In addition, Type I diabetics experience less educational attainment than non-diabetics. They are 3 to 7 percentage points more likely to drop out of high school and 5 to 9 percentage points less likely to get a university degree. I also find lower wages for type I diabetic persons. They can conservatively expect to lose more than $3,050 annually compared to their peers without Type I diabetes. The results show that there is socioeconomic heterogeneity in the impact of Type I diabetes on educational attainment and labor market outcomes. Individuals of parents with less than high school and who are in a low-income group have the worst educational attainment and labor market outcomes in adulthood. The socioeconomic status of the family has a positive impact on the long-run consequences of Type I diabetes, mitigating the negative effects of diabetes on their children.

ACKNOWLEDGMENTS

Many people deserve a big thank you for their helpful advice, guidance, and support throughout the completion of this project.

First and foremost, I am profoundly indebted to my advisors, Dr. Julia Witt and Dr. Umut Oguzoglu for their continuous guidance and unconditional support in each step towards the completion of my thesis throughout my entire time in the Ph.D. program in the Department of Economics at the University of Manitoba. Thank you for sharing and handing down your wisdom and experiences in academia and daily life. Your passion for research has greatly inspired me. It is my fortune to have been your student.

I owe my deepest gratitude to my committee members, Dr. Evelyn Forget and Dr. Mahmoud Torabi, for not only providing me the guidance to conduct my thesis but for constantly providing support and advice. I have been fortunate enough to be surrounded by an incredible advisory committee. Thank you to Dr. Forget for her valuable advice and continuous professional insights to advance my work. Thank you to Dr. Torabi for the great academic experience that I get from his Biostatistics course which helped me in my work.

I would like to gratefully acknowledge the invaluable comments and advice that I received from Dr. Stefan Staubli, my external examiner at the University of Calgary. His excellent guidance and insights helped me to enrich my work.

Another warm acknowledgment goes out to Mrs. Heather Prior at the Manitoba Centre for Health Policy (MCHP). I received countless amounts of help from her. Thank you for your generosity in giving information. I would like also to thank Charles Burchill for the generous information he provided to me during my SAS course at MCHP. I am grateful to Mrs. Theresa

Daniuk for her help getting various approvals from different authorities to get access to the datasets at MCHP.

Financial support provided by the University of Manitoba Graduate Fellowship is also gratefully acknowledged. I am also very grateful for the financial support received from my advisor Dr. Julia Witt through her CIHR grant which allowed me to get access to the Repository database at Manitoba Center for Health Policy and to complete my second project in my thesis.

And last but not least, thank you to my family and friends for all of their support and encouragement along the way!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

The Effect of Retirement on Health and Health-related Behavior: Evidence from Non-parametric Fuzzy Regression Discontinuity Design (RDD)

Abstract

This paper estimates the causal effect of retirement on health status and health-related behavior using the last ten waves of the Health and Retirement Study (HRS) survey in the U.S. during the period 1996-2014. The non-parametric and parametric Fuzzy Regression Discontinuity Design (RDD) techniques are applied to capture the potential reverse causality from health to retirement (endogeneity issue) by exploiting the exogenous variation in retirement decisions induced by the U.S. pension eligibility ages at 62 and 65. In general, retirement has a significant negative impact on physical and mental health status. The non-parametric fuzzy RD results show that the significant jump in retirement probability is associated with an 8% decline in the cognitive functioning score of retirees, and 0.42 points increase in the CESD depression scale. Retirees also are 13 percentage points less likely to report good general health status. Regarding health-related behavior, retirees are 8.8 percentage points less likely to be drinkers relative to non-retirees and they are 4 percentage points less likely to consume alcohol more than three times per week. However, they are 3 percentage points more likely to be smokers relative to non-retired individuals. Furthermore, retirement has a significant heterogeneity effect across socioeconomic and retired gender groups. To address the issue of Medicare eligibility at age 65, I conduct different falsification tests of the RD analysis and the results remained valid. The parametric fuzzy RD estimation is introduced as a sensitivity check for using more information from the observations far away from the neighborhood of the cut-off point and the results were consistent. The estimates also were robust for using different bandwidths and different weighting kernel functions.

1.1. Introduction[1]

Demographic change in the U.S shows that the proportion of retired individuals has substantially increased during the last few decades, which has also coincided with an increase in the life expectancy of Americans. This change means that individuals will spend a large portion of their life in retirement, during which a different lifestyle can affect the health of retirees (Bonsang, Adam, and Perelman, 2012). Theoretically, retirement may change individuals' health in opposing ways. For instance, retirement may reduce work-related stress and increase leisure time, which can be used to invest in health through physical exercise and more sleep (Behncke, 2009; Eibich, 2014). In addition, retirees are more likely to participate in social activities and reduce their alcohol and tobacco consumption (Midanik et al., 1995). In contrast, retirement is a fundamental change in an individual's life, which may lead to depression and social isolation. For instance, individuals who are very satisfied with their work and have good work relations and contacts may see retirement as a stressful event that takes them away from their social networks and friends. Consequently, the retirement period may be accompanied by the emotional impacts of loneliness or feeling old (Behncke, 2009).

[1] All empirical estimation included in the tables in this chapter has been done using the statistical package "rdrobust" which has been written and developed by a group of professors at University of Michigan, U.S (Calonico, Cattaneo and Titiunik, CCT hereafter). Also, I used the statistical package "rddensity" which has been developed by Cattaneo, Ma and Jansson at University of Michigan and UC Berkeley, respectively. Actually, my estimation would have not been done without their seminal work on the non-parametric RD design and their codes to implement this design. Thanks for making the packages available for public access. All the RD estimation presented in the figures have been done based on my own estimation by applying the conventional RD estimator and using the Cross-validation optimal bandwidth as a kind of robustness to see how the RD results based on MSE-optimal bandwidth developed by CTT are consistent with the RD results based on the CV-optimal bandwidth. The Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points increment. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

Understanding the net and causal effects of retirement on health is important, especially as life expectancy is increasing and populations around the world are aging. There has been significant growth in the proportion of individuals aged 65 and older in the American population since the 1960s. Moreover, there are over 40 million Americans age 65 and over, comprising 14% of the population, and this proportion is projected to increase to 20% by 2030 (U.S. Census Bureau, 2016). Governments have thus realized that future contributions to the public pension program by the labor force may not be enough to cover their pension obligations to retirees, which could result in either a rise in the contribution rate by the employed or a decrease in pension payments. A common and easy route that politicians can use to relieve some of the pressure on the public pension system is to increase the official working age to keep individuals in the labor force longer. In fact, many developed countries have started to or are planning to raise the eligibility age of pension receipt to induce individuals to postpone their retirement (Hering and Klassen, 2010). For instance, the earliest eligibility age for receiving retirement pension in the U.S. is 62 and the official age to claim full social security is 65. However, the government plans to increase the normal retirement age from 65 by two-month increments to 67 by 2025. The success of policies that delay the pension eligibility age in reducing government expenditure on social security programs will depend partially on the health effect of delaying retirement. For instance, if retirement has a negative effect on health, any retirement policy that supports late retirement will preserve health and labor force participation of older people, thereby reducing health care utilization and expenditure, in addition to relieving the government's financial pension burden. In contrast, if retirement has positive effects on health, policies that increase age eligibility may increase the cost of health care utilization, and this, in turn, can offset savings in the government's pension burden. Therefore, policymakers need to know the

effect of retirement on health, so that they can design programs that eliminate or mitigate negative effects.

The literature has a large number of studies that focus on investigating the causal effect of retirement on health. However, the empirical results of these studies are mixed (Eibich, 2014). Some studies find that retirement has a positive effect on health (e.g., Ekerdt et al. 1983; Evenson et al. 2002; Charles 2004; Neumann 2008; Coe and Lindeboom 2008; N. B. Coe & Zamarro, 2008; Midanik et al., 1995), while other studies find that retirement has a negative effect on health (Behncke, 2009; Dave et al. 2008; Behncke 2012; Calvo et al. 2013). Some studies find no effect (Bound and Waidmann, 2007).

Although these studies use varying methodologies, samples and health outcomes, many of them use data from the same country. Consequently, the mixed results of these studies can not be explained by differences in institutional settings, whether there is universal healthcare or some other health insurance scheme. However, the inconclusive and conflicting results of the effect of retirement on health may be due to an endogeneity problem and heterogeneity effects (Eibich, 2015). Specifically, a reverse causality problem can exist when individual health has direct impacts on retirement decisions. Empirical studies find that workers with poor health are more likely to retire ((Disney, Emmerson and Wakefield 2006; Jones, Rice, and Roberts 2010 for the UK); (Riphahn, 1999; Lechner and Vazquez-Alvarez, 2011, for Germany); (Bound et al., 1998 for the US) and (McGarry, 2004)). Therefore, an endogeneity issue may arise when researchers estimate the effect of retirement on health by comparing the health of retirees before and after retirement and ignoring the possibility that workers with negative health experiences are more likely to retire (Bound & Waidmann, 2007). If this reverse causality of health on retirement is not captured, the results will be negatively biased (Eibich, 2014) and methods such as ordinary

least squares or fixed-effects panel estimations cannot be used. Another potential explanation of the conflicting results of previous studies is the heterogeneity effect, where the retirement effect may depend on the gender or educational attainment of the retired individuals (Behncke, 2009).

This study contributes to the current literature in four ways. First, further evidence on the effect of retirement on health is provided. The analysis estimates the causal effect of retirement not only on health status but also on health-related behaviors, such as smoking and alcohol consumption, which are inputs to health in Grossman's health capital model (Grossman, 1972). This sheds light on the channels through which retirement can affect health. Specifically, instead of focusing on a narrow set of health outcomes, I will estimate the effect of retirement on more than one outcome. Applying the same methodology and using the same individuals in the estimation of the impact of retirement on multiple health outcomes provides a clearer picture of the retirement effect and generates results that are precise and robust. Second, this study proposes an alternative econometric technique to quantify the impact of retirement on health and health-related behavior. A non-parametric and a parametric approach of Regression Discontinuity Design (RDD) are applied to capture the retirement endogeneity problem. To my knowledge, this is the first study that uses the non-parametric approach to quantify the effect of retirement on health. Using non-parametric and parametric methods is a useful way to show that the estimation results do not rely on the approach that was used. The study also investigates the effect of retirement on health at multiple discontinuities, which increases the validity of the results. In addition, when the parametric approach is applied, this paper uses a method called bivariate likelihood. Most previous parametric studies of the relationship between retirement and health used instrumental variables (IV) to capture the retirement endogeneity problem by applying different instrumental variables. Although the instrumental variable estimation methods

are widely used in econometrics to address endogeneity, for parametric estimation bivariate likelihood estimation methods are theoretically and empirically superior to the traditional two-stage instrumental variable when outcomes are binary (Marra et al., 2014; Bhattacharya et al., 2006; Wooldridge, 2010). Third, this study examines the impact of retirement on health and health behavior by stratifying the models by gender, and socioeconomic status to assess a potential heterogeneity effect for retirement. Finally, I applied two different techniques to estimate the optimal bandwidth that is used to estimate the RD treatment effect: Cross-validation optimal bandwidth (CV) and the Mean Square Error- optimal bandwidth (MSE), which developed by Calonico, Cattaneo, and Titiunik (2014, CCT hereafter), as a kind of robustness check of the consistency between RD results based on MSE-optimal bandwidth and those based on the CV-optimal bandwidth. I also applied three different procedures to estimate the RD treatment effect to ensure the robustness of the RD results: (i) the traditional RD estimates with conventional variance estimator (conventional inference); (ii) Bias-corrected RD estimates with a conventional variance estimator (Bias-corrected inferences); (iii) and bias-corrected RD estimates with a robust variance estimator (Robust RD inferences), proposed by CCT.

Specifically, this study addresses the following research questions:

- What is the impact of retirement on health, where "health" is defined not exclusively as the absence of disease or injury but includes physical, mental and social well-being?
- How does retirement affect individuals' health behavior? Specifically, what is the effect of retirement on alcohol consumption and smoking (participation and intensity)?
- Is there socioeconomic heterogeneity in the effect of retirement on health and health-related behavior?
- Is there gender heterogeneity in the impact of retirement on health?

The main issue in addressing these questions is to find exogenous variation in retirement status, which eliminates the potential reverse causality problem from health shocks to retirement. That is, individuals can retire when they become ill or injured, and as a result, poor health may cause retirement. This study uses the Regression Discontinuity Design (RD) to exploit exogenous variation in retirement decisions induced by pension eligibility ages and identifies the causal effects of retirement on health outcomes and behavior using the last eight waves (2000-2014) of the American Health and Retirement Study survey (HRS).

The other issue that affects the RD estimators is the drastic change in health care coverage at age 65. Specifically, the effect of retirement on health at age 65 could potentially be driven by Medicare eligibility and health care utilization, not necessarily be driven by retirement status. To address this issue, I conducted a falsification test of the RD analysis using several checks. First, I conducted a similar Fuzzy RD analysis at age 62, where a significant jump in retirement is also observed because this is the earliest age at which Americans can receive social security payment. Second, I run the RD analysis on the low-income subgroup. This group includes individuals who are in the lowest 20% of the income distribution in the dataset or individuals whose income is below the federal poverty line after adjustment for the number of persons in the family (HHS, 2018)[2]. This subgroup will include the persons who may be eligible for the Medicaid health insurance program before age 65. Medicaid is a joint federal and state program that helps low-income adults and people with certain disabilities cover health care costs. Eligibility to Medicaid is based on income and family size. Although the HRS includes a question about whether persons are enrolled in the Medicaid program, there are some cases in

[2] According to the U.S. federal poverty guidelines used to determine eligibility for certain federal programs, the 2018 poverty guidelines in 48 states are as follow: $12,140 for one person family, $16,460 for two person family, $20,780 for three person family, $25,100 for four person family, $29,420 for five person family, $33,740 for six persons family, $38,060 for seven person family, $42,380 for 8 persons family, and $4,320 for each additional person for families with more than 8 persons.

which individuals did not report any information about their health insurance coverage before age 65. Therefore, using the previous income criteria may help in identifying persons who may be eligible to have Medicaid insurance coverage to get a reasonable sample size to this subgroup. Third, I run the RD analysis at age 65 for the subgroup of individuals who were eligible for health care coverage before 65 or individuals who have other health care coverage, private or public, for at least five years before they retired. This, then, includes individuals who already had health insurance coverage before reaching age 65, and who continue to have health care coverage after age 65, whether the coverage is public or private.

The results of this study show that retirement has a significant negative impact on health status. The non-parametric fuzzy RD results report that the significant jump in retirement probability is associated with a negative and highly significant jump in the reported total word recall test score. Using the MSE-optimal bandwidth, the non-parametric fuzzy RD estimator suggests that retired individuals experience a drop in their cognitive test score by about 0.8 points, which is equivalent to an approximately 8% decline in the cognitive functioning score of retired individuals. Retirement also has a significant negative impact on mental health status as measured by the CESD depression scale. The non-parametric fuzzy RD estimator suggests that retired individuals experience dramatic increases in the CESD scale, approximately 0.42 points higher (95% confidence interval [0.294554, 0.550295]). Additionally, retirement is associated with a significant negative impact on self-reported general health. Retired individuals are 13 percentage points less likely to report that they are in excellent, very good or good health status than non-retired individuals. The effects of other covariates included in the study are consistent with previous literature. Moreover, the effects of retirement on the three health status measures

(cognitive functioning, mental health, and self-reported health) are not sensitive to different bandwidth selections.

Regarding health-related behavior, retirement has a negative effect on alcoholic consumption participation and intensity, but it has a positive impact on smoking participation. The non-parametric fuzzy RD estimates indicate that retirees are 8.8 percentage points less likely to be drinkers relative to non-retirees and that they are 4 percentage points less likely to consume alcohol more than three times per week, for those who continue drinking after retirement, relative to non-retirees. However, the effect of retirement on alcohol consumption intensity per week is small and not significant for any bandwidth choices and across the three different non-parametric fuzzy RD estimators. Also, retired individuals are 3 percentage points more likely to be smokers relative to non-retired individuals. However, the impact of retirement on smoking participation is small in magnitude and very sensitive to bandwidth choice.

The results of this paper also suggest that there is significant heterogeneity across gender and socioeconomic groups. In general, the transition from working to retirement has larger negative impacts on males than females. Across different socioeconomic groups, retirement has a larger negative impact on less-educated individuals compared to high-educated individuals. Well-educated individuals may have better health outcomes because they are more efficient at making health investments (Grossman and Kaestner, 1997). For health-related behavior, retirement has a larger negative impact on drinking participation and drinking intensity on retired males compared to retired females. The non-parametric fuzzy RD results based on the CV-optimal bandwidth selection are consistent with the RD results based on the MSE-optimal bandwidth selection. However, the CV-optimal bandwidth length, for all estimations, is larger than the MSE-optimal bandwidth. These findings are expected because the adjusted MSE-

optimal bandwidth by CCT gives more weight to the variance and the chosen bandwidth could be larger to minimize both the bias and the variance under the given bandwidth. In addition, the non-parametric RD results are supported by the parametric RD results for the same health outcomes, which ensures the results are precisely estimated and robust.

The remainder of this chapter is organized as follows. Section 2 summarizes the related literature; section 3 describes the data and presents summary statistics for the sample; section 4 provides the empirical econometric specification and methodology; sections 5 and 6 show the results of the non-parametric and parametric regression discontinuity design (the recursive bivariate probit model); section 7 includes the discussion and conclusion.

1.2. Literature Review

There exist several studies in the literature studying the effect of health status on the decision to retire. Although this decision is determined by many factors, including eligibility for social security pension payments, access to health insurance, financial resources, and spouse's retirement status, the health status of the individual plays a significant role. Persons who are in poor health, for example, those suffering from chronic health conditions, are more likely to retire early than those who are healthy (Belgrave et al., 1987). McGarry (2004) finds that individuals in poor health are less likely to continue working than individuals in good health, and the decision to retire is significantly associated with a change in individuals' health status. Gustman and Steinmeier (1986) find that the onset of a long-term health problem increases the probability of retirement. Additionally, access to health care services has been documented as a crucial factor in retirement decisions: several studies find a large positive association between retiree health insurance offers and early retirement (Madrian, 1994; Hurd and McGarry, 1993; Karoly and Rogowski, 1994).

Complementing this literature are studies that focus on the effect that the decision to retire has on individual health outcomes (instead of the effect of health status on retirement). The empirical evidence on the causal effect of retirement on health outcomes is inconclusive (Eibich, 2014). Some studies show that retirement has a positive effect on health (N. B. Coe & Zamarro, 2008; Midanik et al., 1995; Insler, 2014). Coe and Zamarro (2011) used early and full retirement ages across European countries as instruments and found a significant positive effect of retirement on self-reported health. Neuman (2008) used early and official retirement ages in the U.S. as instruments and finds similar results. Similarly, Gorry et al. (2016), using the same set of instruments as Neuman (2008), finds that retirement improves self-reported health and life satisfaction in the U.S. Some studies used different identification strategies and also reported a positive association between retirement and self-reported health. For example, using the Health Survey for England in a regression discontinuity setting, Johnston and Lee (2009) find positive effects of retirement on self-reported health and mental health. The RD design used by Eibich (2015) exploited variation in age eligibility to pension benefits in Germany to estimate the effect of being retired on different dimensions of health and the mechanisms through which retirement can affect health. He finds that retirement improves self-reported health and that there is gender heterogeneity. Coe and Zamarro (2008), using the Survey of Health, Aging, and Retirement in Europe (SHARE), found that retirement has a significant health-preserving effect on the overall health status of people who retire at age 65. Contrarily, they found that retirement has no significant health-preserving effect for those retiring at a younger age.

However, some studies have found that retirement has a negative effect on health (Behncke, 2009). Retirement was associated with lower life satisfaction (Bossé et al. 1987), depression (Portnoi, 1983), increases in the risk of chronic health conditions and severe

cardiovascular diseases (Behncke, 2009), lower well-being (Atchley and Robinson 1982, Grâce et al. 1994, among others), increased difficulties related to mobility and performing daily activities, and mental health decline (Dave et al., 2007). For example, by using the ELSA survey in England, Behncke (2012) applied an instrumental variable (IV) model with propensity score matching and concluded that retirement increases the probability that an individual is diagnosed with a chronic health condition. Dave, Rashad, and Spasojevic (2008) used the US Health and Retirement Study for the period 1992-2005 and found that full retirement causes a 5 to 16 percent increase in difficulties associated with mobility and daily activities, a 5 to 6 percent increase in illnesses, and a 6 to 9 percent decrease in mental health. Numerous other studies show a significant negative effect of retirement on cognitive functioning (Rohwedder and Willis, 2010; Mazzonna and Peracchi, 2012; Bonsang et al, 2013; Bingley and Martinello, 2013). Results suggest that promoting labor force participation among older workers could delay cognitive decline (Bonsang et al., 2012; Mazzonna & Peracchi, 2010). Ayyagari (2014) used data from the U.S. Health and Retirement Study (HRS) to examine the relationship between retirement and smoking decisions from 1992 through 2008. He exploits eligibility for Social Security benefits at age 62 to account for the endogeneity problem of retirement. He found that retirement increases the probability of smoking among smokers, and he also found evidence of heterogeneity in the impact of retirement. Other studies, however, reported no evidence of negative health effects related to retirement (Bound and Waidmann, 2007). Coe and Lindeboom (2008) found no negative effect of early retirement on men's health, and Latif (2012) found that, in a Canadian sample, retirement has no significant effect on self-assessed health.

These inconclusive and conflicting results of the effect of retirement on health may in part due to an endogeneity problem and heterogeneity effect (Behncke, 2009). Specifically,

reverse causality can exist when individual health directly affects retirement decisions. In fact, empirical studies have found that workers who experience a decline in their health are more likely to retire (Disney, Emmerson and Wakefield 2006; Jones, Rice, and Roberts 2010 for the UK; Riphahn, 1999; Lechner and Vazquez Alvarez 2011, for Germany; Bound et al., 1998 for the US; and McGarry, 2004). Therefore, an endogeneity issue may arise when researchers estimate the effect of retirement on health by comparing the health of retirees before and after retirement and ignoring the possibility that workers with negative health experiences are more likely to retire (Bound & Waidmann, 2007). If this reverse causality of health on retirement is not captured, the results may be negatively biased (Eibich, 2014) and we cannot use methods such as ordinary least squares or fixed effect panel estimations.

To correct this bias, economists have attempted to use different techniques, including cross-country variations, early and unexpected retirement incentives, and official pension age eligibility. Regardless of the identification strategy, most economists have been looking for an exogenous source of variation in retirement decisions to address the endogeneity issue. Therefore, a recent development in the health economics literature is to use age eligibility for social security benefits to estimate the effect of retirement on health. For instance, in the UK, most retirement income programs become available at age 65 for men and 60 for women (Bound and Waidmann, 2007). In the United States, individuals can retire early at age 62 and receive some social security retirement benefits, and receive their full retirement social security at age 65 (Bonsang et al., 2012). In recent years, the instrumental variables (IV) approach has been widely used by researchers to address the endogeneity problem. Most commonly, instrumental variables that affect the endogenous retirement decision but that do not affect health outcomes are applied, and usually, these are related to the social security and pension systems in the country. For

instance, Charles (2004) estimated the effect of retirement on wellbeing as measured by "feeling depressed" and "feeling lonely". He used a group of dummy variables that define age 62, 63, 70, and 72 as instrumental variables by applying a 2SLS. He found a negative correlation between retirement and subjective well being. To examine the mortality effects of retirement, Fitzpatrick and Moore (2018) assess whether there is a discontinuity in mortality when individuals are eligible for social security at age 62 in the U.S. using a sample of decedents born between 1930 and 1948. They find that retirement is associated with a two percent discontinuity in male mortality immediately after age 62 and the discontinuity in the female mortality is smaller. They also document that the increase in males mortality is related to the change in their lifestyle.

The effect of retirement on mental health and cognition has been extensively studied. Kofi Charles (2004) was one of the first to examine the relationship between retirement and mental health. He used both the age of Social Security benefit eligibility and the change in laws that affect when Social Security can first be withdrawn. He found that retirement reduced loneliness and depression, as a measure of mental health. Rohwedder and Willis (2010) argued that countries with a larger proportion of their labor force working later in life exhibit smaller differences in cognitive performance between older and younger men. Using cross-country variation in eligibility ages for early and official public pension benefits, they found that retirement reduces cognitive scores by nearly 1.5 standard deviations. Similar studies that used eligibility ages for public pension as instruments also found negative effects on cognitive function (Bonsang, Adam, and Perelman, 2012; Mazzonna and Peracchi, 2012; and Tumino et al., 2016). Mazzonna and Peracchi (2012) investigated the relationship between aging, cognitive abilities, and retirement by using a survey on health and retirement in Europe (SHARE) for individuals older than 50 years. They used an instrumental variables approach to exploit

variation in the eligibility age for retirement between and within countries. They found that there was an increase in the rate of cognitive ability decline after retirement and that education plays an important role in explaining heterogeneity in the level of cognitive abilities. Bonsang et al (2012) used the same dataset (SHARE) and instrumental variables. Their results confirm the significant negative effect of retirement on cognitive functioning and suggest that promoting labor force participation delays cognitive decline. Midanik et al., (1995) assessed the short-term effect of retirement on mental health and health behaviors of members of a health maintenance organization (age 60-66). After controlling for age, gender, marital status, and education, they found that retired individuals were more likely to have lower stress levels and engage in regular exercise compared to those who did not retire during the study period. They also found that retired women were more likely to report no alcohol problems compared to non-retired women. There were no differences between the groups on self-reported mental health status, coping, depression, smoking, alcohol consumption, and frequency of drunkenness. Ekerdt, Bosse, and LoCastro (1983) showed that health may improve after retirement due to the elimination of work-related stress. Additionally, retirement may give individuals more time to invest in health through activities that promote health improvement. However, Coe and Zamarro (2011) did not find any effects. Coe et al. (2012) also did not find an effect using an exogenous variation of early retirement design.

The literature on cognitive functioning argues that individuals' activities, educational attainment, and personality could affect cognitive reserve (Stern 2002, 2003), which is formed through social activities, lifestyle and social networks. Individuals who engage in brain-stimulating activities may prevent their cognitive functioning or mitigate the negative effects on their cognitive abilities, surrounding environment and social networks may play the same role

(e.g., Flicker 2009; Hertzog et al. 2008; Salthouse 1991, 2006; Scarmeas and Stern 2003; Fratiglioni et al. 2004; Börsch-Supan and Schuth 2013). Some studies found an association between individual personality traits, like risk aversion, patience, and self-esteem, and cognition (e.g., Frederik 2005; Benjamin et al. 2006; Dohmen et al. 2007; Midanik et al. 1995; Yates 1990). Also, some studies documented the significant role of education in cognitive abilities (Banks and Mazzonna 2012; Maurer 2010; McFadden 2008; Evans et al. 1993).

The impact of retirement on health-related behaviors has been given the least attention in the literature. Moreover, studies focusing on health-related outcomes come to mixed conclusions. Insler (2014) used individual retirement age from the Health and Retirement Survey (HRS) as an instrument and found that retirement increases exercise and decreases smoking. Using pension eligibility age in Germany, Eibich (2015) showed that retirement increases activity, sleep, and leisure time activities, and that it decreases smoking rates and BMI. Motegi, Nishimura, and Terada (2016) found that retirement in Japan increases physical exercise and reduces drinking but does not change smoking rates among Japanese retirees. In contrast, Müller and Shaikh (2017) used the Survey of Health, Ageing and Retirement in Europe (SHARE) for 19 European countries to investigate the impact of the partner's retirement on the other partner's health behavior. They found a significant increase in the frequency and intensity of alcohol consumption, combined with a significant decrease in moderate physical activities as a response to the partner's retirement. However, they found that own retirement has significant positive effects on engaging in moderate and vigorous physical activities, but that it also led to a significant increase in the frequency of alcohol consumption. In sum, subjective health is negatively affected by spousal retirement and positively by own retirement. Stancanelli & Soest (2012) use a similar identification strategy to exploit the earliest age retirement laws in France.

Using data from the French Time Use Survey, they find that retirement not only affects own housework time but also affects the time allocation of the other partner.

Overall, although many studies have investigated the impact of retirement on health outcomes, there is inconclusive evidence on the effect of retirement on health. Also, little attention has been paid to the reverse-causality problem of retirement in many of the existing studies, which raises the possibility of biased and inconclusive results. Little attention has also been paid to the effect of retirement on health-related behavior, such as smoking and alcohol consumption. To address these gaps this study applies regression discontinuity design to solve the endogeneity problem and to obtain more comprehensive results of the impact of retirement on health status, as well as on health-related behavior, such as drinking and smoking. Non-parametric and parametric methods are used in the RD design estimates to show that the estimation results do not rely on the chosen strategy. Also, the study investigates the effect of retirement on health at multiple discontinuities, which increases the validity of the results. Moreover, when using the parametric RD design, bivariate likelihood estimation methods will be applied, which are theoretically and empirically superior to the traditional two-stage instrumental variable approach in the case of binary outcome settings (Bhattacharya et al., 2006; Wooldridge, 2010; Marra et al., 2014).

1.3. Data and Study Variables

The empirical analysis in this study uses the last ten waves (1996 - 2014) from the Health and Retirement Study (HRS). The HRS is conducted by the survey research center at the University of Michigan. The original HRS data is combined with RAND HRS data, which is a longitudinal data set that includes a cleaner version of the most frequently used HRS variables for all health outcomes and covariates in this study. The RAND dataset is "a research-friendly

version of a subset of the Health and Retirement Study (HRS)" (St Clair, 2008). This survey collects rich information and detail on respondents' health, health behavior, economic and socio-demographic characteristics, education and labor market dynamics, which allows controlling for individual heterogeneity and addressing the endogeneity problem of retirement and health. The HRS sample has been built up over time to create a sample representative of the American population over age 50 and includes Americans who were born between 1931 and 1941 and their partners since 1992. In 1998, three birth cohorts were added to the original cohort in the HRS: respondents of the Assets and Health Dynamics Among the Oldest Old (AHEAD) Study who were born before 1924; children of the Depression Age, born between 1924 and 1930; and the war babies, born between 1942 and 1947. In addition, the early Baby Boomers, born between 1948 and 1953, were added in 2004; the mid Baby Boomers, born between 1954 and 1959, were added in 2010; and the late Baby boomers, born between 1960 and 1964, were added in 2016 (HRS, 2017).

The HRS collects detailed information about the age of the respondents. The dataset includes information about the month and year of respondents' birth and their ages in months and years. Also, the month and the year of the interview are included in the dataset. This information provides a precise, continuous measure of age in months, which is a crucial element in the RD design when choosing optimal bandwidths and optimal bin sizes to visualize the discontinuity graphically. There are two sample restrictions for this study. First, individuals between 50 and 75 years are included for flexibility in estimating the optimal bandwidth. Second, individuals who reported a disability or not working for mental or physical illness reasons are dropped from the sample. The final sample includes 94,131 person-year

observations. The full sample is divided into subgroups to investigate heterogeneity in the effect of retirement on health.

1.3.1 Retirement

The main variable in this study is the retirement status of individuals. There are many definitions of this in the retirement literature (Coe and Zamarro, 2011 and Insler, 2014), and the one used here follows those in French (2005); Bonsang et al. (2012); Mazzonna and Peracchi (2012), and Eibich (2014). An individual is considered retired if she/he is not in the labor force (has exited the labor market). This includes individuals who report that they are ‘‘not in the labor force’’ at the time of the interview. Individuals working or looking for work (unemployed) are considered to be in the labor force (Mavromaras et.al, 2014). Following Eibich (2014), I use self-reported retirement as a treatment variable where behavior adjustment usually happens when individuals regard themselves as retired. That is, respondents are classified as being retired if they are out of the labor force with the intention of staying out permanently. Therefore, the study excludes the following respondents from the sample: respondents who reported that they had never worked because retirement implies a change in respondent lifestyle; respondents who report that they are not in the labor force because they are disabled or homemakers; respondents who reported that they left their jobs before age 50; and respondents who report that they are retired, they just left their career job, but continue to do some paid work.

1.3.2 Measuring Health (Health Outcome Variables)

For the empirical analysis, this paper uses three variables to capture physical and mental health as outcome variables.

1.3.2.1 Self-reported Health (General Health)

Since objective measures of health usually ask about specific diseases, they may not be complete measures of health. Therefore, health economists usually use self-reported general health (GH) to get a more precise picture of the respondent's overall well-being (Coe and Zamarro, 2011). Farmer and Ferraro (1997) and Mackenbach et al. (2002) found that respondents' subjective assessment of their health is related to objective health measures. some literature argues that self-reported measures of health may suffer from a "justification bias" in the context of retirement, that is people may report that they are in worse health upon retirement to justify that they are retired (Currie and Madrian, 1999). Black et al., 2017 document that men and women overstate their disability level to justify their access to disability welfare payments, justification bias. The effect of this bias can be captured partially, in this study, by dropping individuals who reported that they are disabled (see Benitez-Silva et al., 2004). This issue also can be captured in the regression discontinuity setting by using different bandwidths in the estimation and by applying different validation tests to validate the RD estimates. Moreover, to compare self-reported and objective measures of health in retirement models, Bond (1991) constructed statistical models to incorporate the self-reported and objective measures of health. He finds that self-reported measures of health perform better than many researchers believe.

In HRS, respondents are asked how they would describe their current health status in general on a 5-point scale. I dichotomized these responses: the best three categories (excellent, very good, and good) are recorded as "*in good health"* and the worst two categories (fair and poor) are recorded as *"not in good health"*, as in previous studies (Kunst et al., 2005; Mackenbach et al., 2002). That is, self-reported health status is captured by a dummy variable that is "1" if "*in good health"* and "0" if "*not in good health*".

1.3.2.2 Mental Health Status

Affecting functioning, an individual's mood and emotional health, is a significant domain of an individual's overall health and wellbeing (Steffick, 2000). Research has shown that the causal relationship between affective functioning and individual health status runs in both directions. For instance, Barefoot and Schroll (1996) found that individuals who experienced high levels of depressive symptomatology are more likely to develop myocardial infarction (MI), heart attack, and mortality. In contrast, Hachinski (1999) found that physical illness may increase the likelihood of developing depression and anxiety. He found that, among 100 ischaemic stroke survivors, the incidence of major depression increased over time from 6% initially, to 11% at year one, to almost 18% at 18 months.

To measure mental health status, I use the HRS depression symptoms measure, which is a subset of the Centre for Epidemiologic Studies Depression (CESD) scale. The CESD scale includes 20 domains to assess the level of depressive symptoms. Starting from wave 2, respondents in the HRS are asked to rate the frequency of eight symptoms of psychological distress. The HRS shortened version of the CESD scale is the sum of five negative domains plus three positive domains.

The negative domains measure whether the individual reported having the following symptoms all or most of the time, during the last week: depression, felt alone, felt sad, felt that everything is an effort, sleep is restless. The positive domains measure whether the individual felt happy, felt hopeful about the future and enjoyed life, all or most of the time, during the last week. Positive domains are reverse-scored, with lower scores given to the response "most of the time", so that higher numbers indicate more negative emotions and worse mood. The HRS shortened version of the CESD scale is estimated by adding up the eight domains, which yields a

total score ranging from zero to 8. The CESD scale has been widely used as a measure of depressive symptoms and tested for reliability and validity for a variety of subpopulations (Steffick, 2000).

1.3.2.3 Cognition Functioning

It is well-documented in the literature that heterogeneity in individual-level cognitive functioning and the rate of age-related change in cognitive functioning are associated with an individual's lifestyle, such as diet, frequency of social interactions, physical activity, and the degree of engagement in mentally stimulating activity (Salthouse, 2006). The mental-exercise hypothesis (which states that keeping mentally active will prevent age-related mental decline) is widely used by researchers to argue that individuals have some control over their cognitive functioning and that policy intervention can induce individuals to participate in mentally stimulating activities to save resources that are otherwise used to treat issues related to cognitive function decline (WHO, 2002).

The HRS has included a variety of cognitive functioning measures (including single and indices measures) that assess different aspects of the cognitive domain. In this study, I use the total word recall index as a measure of episodic memory. In the HRS, the episodic memory test consists of two stages, which provides two measures: immediate word recall and delayed word recall. For the immediate word recall measure, the interviewer read a list of 10 words (e.g., car, lake, book, etc.) to the respondent, and asked the respondent to recall as many words as possible from the list in any order. Correct responses were scored as one point, thus 10 was the highest attainable score. To measure delayed word recall, after around 5 minutes of other survey questions, respondents were asked to recall the words that were read to them for the immediate word recall question. Correct responses were again given one point, with 10 being the highest

attainable score. The total word recall test score is calculated by counting the number of immediate and delayed words that were recalled correctly. The test score ranges from 0 to 20 and has a normal distribution with a sample mean of 10.5 and a standard deviation of 3.38.

1.3.3. Health-related Behavior

The HRS contains a rich set of variables that measure individual health-related behavior. This study uses smoking participation and alcohol consumption participation and intensity (number of drinks) as outcome variables. There are two questions in the HRS that can be used to identify smoking behavior after retirement: *"have you ever smoked*" and "*are you currently smoking*". The ever-smoked question is usually only asked at the respondent's first interview, and so the answer is carried forward for subsequent waves. If during any wave a respondent reports that she/he currently smokes cigarettes, the ever-smoked question is set to yes in that survey and all subsequent waves. Therefore, smoking status is captured by a binary variable ("smoking") and equals "1" if the individual currently smokes and "0" otherwise. For alcohol consumption, there are two questions in the HRS that capture drinking status and frequency. Drinking participation is captured by the binary variable "currently drinking", which takes a value of "1" if the respondent is currently drinking any kind of alcohol and "0" otherwise. If the respondent is identified as a drinker, the following question is asked about the last 3 months: "*In the last three months, on average, how many days per week have you had any alcohol to drink? (For example, beer, wine, or any drink containing liquor.)"*. The binary variable "at least 3 days per week" takes a value of "1" if drinking frequency is at least 3 days or more per week and "0" otherwise.

1.3.4. Control Variables

I used control variables to capture heterogeneity effects, including education, age, household income, marital status, and gender. Marital status is represented by two dummy

variables. “married” which is one if the individual is married and zero otherwise, and “others” which is one if respondent reports that she/he is separated, divorced, or widowed and “0” otherwise, “Single” (reference group). The dummy variable for “female” takes value “1” if the respondent is female and “0” if male. The race is captured by two dummy variables: white (reference category), black, and other.

Respondent’s educational attainment is captured by two dummy variables: high school and post-secondary degree, a university degree or postgraduate degree, and less than high school (reference category). Income level is captured by two dummy variables: low income (reference category), middle income, and high-income group. A dummy variable “spouse retired” is created to capture whether the respondent is living with a partner or not. Some studies found that the behavior of retired individuals may be significantly affected by the characteristics of the individuals with whom they spend time. “children number” is a continuous variable to capture the number of children in the respondent’s household. Children may affect an individual’s lifestyle, and thus individual health behaviors. Finally, different geographical regions in the US are captured by four dummy variables: Midwest, South, West, Others, and Northwest (reference group).

1.4. Methodology: The Regression Discontinuity Design Theoretical Framework

In this section, I introduce the fuzzy regression discontinuity (RD) design. I begin by framing the traditional single-forcing variable RD design in the context of the Rubin causal model (Holand, 1986; Rubin, 1974) and then describing the empirical estimation of RD treatment effect in the context of Hahn et al., (2001). After that, I introduce the extension of CCT to the traditional RDD. Regression Discontinuity Design (RDD) was first conceptually developed and applied by Thistlethwaite and Campbell (1960) in a study of merit impact awards on the future academic outcomes of students. By the end of the 1990s, RD methods were widely used and extended by many academic researchers in economic studies (Lee, 2007; Card teal., 2006; Chay and Green Stone, 2005; Dinardo and Lee, 2004; Van Der Klaavw, 2002; Black, 1999; Angrist and Lavy, 1999).

RD design is a quasi-experimental design with the main characteristic that the probability of receiving treatment changes discontinuously as a function of one or more underlying variables (Hahn et al., 2001; Coe and Zamarro, 2011). Generally, RD requires an assignment variable (Forcing or score variable) which classifies individuals below or above a threshold point, and individuals above this cut-off receive treatment (treatment group), while individuals below this cut-off will not receive treatment (control group). Direction and magnitude of the jump at the threshold can be used as a measure of the causal effect of the treatment for individuals close to the cut-off point (Lee & Lemieux, 2010; Jacob et al., 2012).

Under the standard Robin causal model (RCM), it is assumed that there is a random sample $\{Y_i(1), Y_i(0), X_i\}_{i=1}^{n}$, where $Y_i(1)$, and $Y_i(0)$ are potential outcome variables when individual i is treated and untreated, respectively, and X_i is an observed running variable (sometimes called forcing or assignment variable), which used in determining treatment status.

Define $D_i \in \{0, 1\}$ as a binary decision rule where D_i is determined partially or completely based on the value of the running variable X_i. Assignment to treatment is assumed to follow the rule $T_i = \mathbb{1}(X_i \geq \bar{x})$. That is, all units whose score are above the cutoff point, $\bar{x}$, are assigned to treatment and all units whose score is less than the cutoff point, $\bar{x}$, are assigned to the control group. In the literature, there are two distinctive types of RD, the sharp and the fuzzy RD. In the sharp RD setting, the binary treatment variable D_i is a deterministic function of the running variable X_i such as $T_i = D_i = \mathbb{1}(X_i \geq \bar{x})$ with the cutoff value $X_i = \bar{x}$. In contrast, in fuzzy RD, the relationship between treatment and running variable is stochastic such that individuals do not necessarily comply with the assignment to treatment when $X_i \geq \bar{x}$ (imperfect compliance). That is, there are some units for which $T_i \neq D_i$, i.e., assignment to treatment (T_i) does not necessarily mean that treatment is actually received by these units. Therefore, the probability of receiving the treatment, $P(X) = E(D_i|X_i = x)$, becomes discontinuous at $X_i = \bar{x}$ for a value less than one.

Lee and Lemieux (2010) defined the discontinuity setting as designs and not just as estimation methods because the running variable x must be boosted by empirical evidence and institutional information. In this study, the design is partially given by the official age of retirement in the USA where early retirement age (62 years) and official retirement age (65 years) represent the threshold score which forces individuals to be above the threshold value and claim their financial benefits.

In this theoretical setup, I am interested in estimating the treatment effect D_i (retirement) on the outcome variable Y_i (health indicator) by defining the individual treatment effect (TE) as $TE_i = Y_i(1) - Y_i(0)$, where TE_i denotes the difference between the value of the outcome variable when individual i is treated ($Y_i(1)$) and the value of the same outcome variable when

the same individual i is untreated ($Y_i(0)$). However, the fundamental problem in estimating the treatment effect is the unobservability of the same individual i in both the treated and untreated status (Rubin, 1975, 1997). Since it is impossible to observe both factual and counterfactual status for individual i at the same time, the identification of the treatment effect relies on comparing outcomes for different individuals with different treatments. Specifically, the focus will be on the average treatment effect (ATE) over a sub-population rather than on the unit level effect (Imbens and Lemieux, 2008). We could then identify the average treatment effect for all individuals in the data set by

$$ATE_i = \mathbb{E}[\, Y_i(1) - \; Y_i(0)],$$

where the expectation $\mathbb{E}[\cdot]$ is taken over individuals. Therefore, the observed sample is $\{Y_i, X_i\}_{i=1}^{n}$, where the observed outcome of individual i, denoted, Y_i , can be written as follows:

$$Y_i = \; Y_i(0) + \; D_i[\, Y_i(1) - \; Y_i(0)] = \begin{cases} Y_i(0) & if\ D_i = 0 \\ Y_i(1) & if\ D_i = 1 \end{cases} \qquad (1)$$

Equation (1) is called the potential outcome model (POM), and it is the baseline equation that links unobservable with observable outcomes (Cervlli, 2015). The following standard assumptions are sufficient conditions to identify the conditional average treatment effect (Hahn, Todd, and Van der Klaauw, 2001, (HTV) hereafter):

Assumption 1: (i) $\lim_{x \downarrow \bar{x}} E[D_i \mid X_i = \bar{x}] = d^+$ *and* $\lim_{x \uparrow \bar{x}} E[D_i \mid X_i = \bar{x}] = d^-$ *exist and (ii)* $\lim_{x \downarrow \bar{x}} E[D_i \mid X_i = \bar{x}] \neq \lim_{x \uparrow \bar{x}} E[D_i \mid X_i = \bar{x}]$

This assumption of RD means that the probability of receiving treatment (retirement) is discontinuous at the cut-off point, $\bar{x}$. That is, the probability of receiving the treatment may not change from zero to one at the cut-off, which gives

$$0 < \lim_{x \downarrow \bar{x}} pr\,[D_i \mid X_i = \bar{x}] - \lim_{x \uparrow \bar{x}} pr[D_i \mid X_i = \bar{x}] < 1$$

Sometimes treatment is not fully determined by the assignment variable. That is, some individuals take treatment when they are not assigned, and some go beyond the cut-off point and do not get treatment, due to noncompliance (Trochim, 1984). In this study, retirement is not fully determined by age, but the probability of retirement increases at the cut-off points of early and official retirement age. As shown in figure 1.1, when individual approaches age 62 and 65 this will increase the chance of retirement discontinuously, but not completely determine the retirement decision in this study. Therefore, I use a "Fuzzy" RD design that allows for a jump at the cut-off point with probability greater than zero and less than one (Lee & Lemieux, 2010; Stancanelli, 2012).

Assumption 2: *The two potential outcomes* $E[Y_i(1)|X_i = x]$ *and* $E[Y_i(0)|X_i = x]$ *are continuous functions in the running variable* X_i *over the support of* $\bar{X}$*,* $Supp\ (\bar{X})$*, i.e., in the neighborhood of the threshold value* $\bar{X}$.

That is, outcome variables have to be a continuous function around the cut-off in the absence of the treatment[3]. Continuity of the conditional expectation of counterfactual outcomes in the running variable is a sufficient condition for identification in RD design (McCrarry, 2008).

Since the difference in the potential outcome $Y_i(1) - Y_i(0)$ is unobservable, we need representative quantities that can be estimated with the data. Therefore, if the conditional mean functions $\mathbb{E}[Y_i(1)|X_i = x]$ and $\mathbb{E}[Y_i(0)|X_i = x]$ are continuous, we can contrast the right limit of the conditional mean, $Y^+ = \lim_{x \downarrow \bar{x}} E[Y_i \mid X_i = x]$, and the left limit of the conditional mean,

[3] Potential outcome must be continuous at the cut-off (Hahn et al., 2001; Imbens and Lemieux, 2007; Lee and Lemieux, 2009):

$$\lim_{x \uparrow \bar{x}} E[Y_i(0)|X_i = x] = \lim_{x \downarrow \bar{x}} E[Y_i(0)|X_i = x]$$
$$\lim_{x \uparrow \bar{x}} E[Y_i(1)|X_i = x] = \lim_{x \downarrow \bar{x}} E[Y_i(1)|X_i = x]$$

If one or both potential outcomes are discontinuous at the cut-off, the RD estimates would be biased.

$Y^- = \lim_{x \uparrow \bar{x}} E[Y_i \,|X_i = x]$, of observable outcomes to estimate the average treatment effect (Otsu, Xu, Matsushita, 2015).

Theorem: (HTV (2001)) suppose assumption 1 and 2 hold; then, the quantity

$$\tau_{fuzzy} = \frac{Y^+ - Y^-}{d^+ - d^-}$$

measures the average treatment effect at $X = \bar{x}$*, i.e.* $\tau_{fuzzy} = E[Y_i(1) - Y_i(0)|X_i = \bar{x}]$*,* $i \in C$*, where C is the compliers set.*

That is, to estimate the causal effect of treatment, we estimate the discontinuity (jump) in the conditional mean of the outcome variable Y_i at the threshold and scale it by the discontinuity (jump) in the conditional treatment probability D_i at the threshold point, i.e.

$$\tau_{fuzzy} = \frac{\lim_{x \downarrow \bar{x}} \mathbb{E}[Y_i \mid X_i = x] - \lim_{x \uparrow \bar{x}} \mathbb{E}[Y_i \,|X_i = x]}{\lim_{x \downarrow \bar{x}} \mathbb{E}[D_i \mid X_i = x] - \lim_{x \uparrow \bar{x}} \mathbb{E}[D_i \,|X_i = x]} \qquad (2)^4$$

It is very important to note that regression discontinuities are calculated at a single cutoff point on the support of the continuous assignment variable which leads to a very local causal effect. That is, RD estimates are, by construction, local and not informative for units away from the cutoff point in the absence of additional assumptions.

Empirically, the local average treatment effect, τ_{fuzzy}, can be estimated non-parametrically by running a one-sided Kernel regression at the cut-off point $\bar{x}$, and obtain the point-wise expectation for the four limits value in equation (2). However, and like other non-parametric methods, the one-sided Kernel regression presents some problems at the boundaries regarding the speed of convergence (Porter, 2003). Alternatively, a consistent estimation procedure to implement the LATE in formula (2) can be constructed nonparametrically by using

[4] See appendix A

the local polynomial regression estimators restricted to a window around the cut-off point. Local polynomial regression estimates of different order approximate the four unknown regression functions restricted to a window around the cut-off point. These estimators have been widely known by their excellent boundary property that makes them a well-suited candidate for estimation and inference in RD design[5] (Cheng, Fan and Marron, 1997).

Following Hahn et al., (2012) and Calonico et al., (2014), the local average treatment effect at the cut-off point is estimated by a local polynomial regression with the same bandwidth for estimation of the discontinuity in the outcome variable (health and health-related behavior) and treatment (retirement) regression. By defining the conditional means of both the outcome variable and the treatment variable as a function of the running variable $\Upsilon(x)$, we can identify four values on both sides of the cutoff as follows:

$$\Upsilon_{+}^{i} = \lim_{x \downarrow \bar{x}} \Upsilon^{i}\ (x) \qquad , \quad \Upsilon_{-}^{i} = \lim_{x \uparrow \bar{x}} \Upsilon^{i}\ (x) \qquad ; \quad i \in \{Y, D\}$$

where Y and D denote the outcome variable (health indicators) and treatment variable (retirement), respectively. The estimated intercept from the local weighted regression $\widehat{\Upsilon}_{+}^{i}$ and $\widehat{\Upsilon}_{-}^{i}$ is an estimate of the LATE

$$\widehat{LATE}_{fuzzy} = \frac{\widehat{\Upsilon}_{+}^{Y} - \widehat{\Upsilon}_{-}^{Y}}{\widehat{\Upsilon}_{+}^{D} - \widehat{\Upsilon}_{-}^{D}} \qquad (3)$$

Specifically, the four unknown regression functions in equation (2) can be approximated by local polynomial methods, which overestimate the global polynomial method, by using only observations that are between $\bar{x} - h$ and $\bar{x} + h$, where $h > 0$ is the chosen bandwidth. Moreover,

[5] Global polynomial approach delivers good approximation overall, but poor approximation at the boundary point (problem known as Runge's phenomenon in approximation theory). Also, local polynomial approach employs a low-order polynomial approximation, usually linear or quadratic, and discard all observations which are extremely far from the cutoff. Therefore, it is more robust and less sensitive to boundary-related problems (for more details, see Catten et al., 2018)

the observations close to the cut-off are given more weight than observations that are far from it by a Kernel function $K(\cdot)$.

In econometric form,

$$\Upsilon_{+}^{i}(h) = I_0' \hat{\beta}_{+}^{i}(h) \;\; and \;\; \Upsilon_{-}^{i}(h) = I_0' \hat{\beta}_{-}^{i}(h) \;\; ; \;\; i \in \{Y, D\}$$

where RD estimators on both sides of the discontinuity are given by

$$\hat{\beta}_{+}^{i}(h) = arg\,min \sum_{\bar{x} \le X_i \le \bar{x}+h} \mathbb{1}(X_i \ge \bar{x}) \left[Y_i - \beta_{+,0} - \beta_{+,1}(x_i - \bar{x}) - \cdots - \beta_{+,p}(x_i - \bar{x})^p\right]^2 K_h\left(\frac{x_i - \bar{x}}{h}\right)$$

$$\hat{\beta}_{-}^{i}(h) = arg\,min \sum_{\bar{x}-h \le X_i < \bar{x}} \mathbb{1}(X_i < \bar{x}) \left[Y_i - \beta_{-,0} - \beta_{-,1}(x_i - \bar{x}) - \cdots - \beta_{-,p}(x_i - \bar{x})^p\right]^2 K_h\left(\frac{x_i - \bar{x}}{h}\right)$$

where $I_0 = (1,0,0,0,\ldots..0) \in \mathbb{R}^{p+1}$ is the first unit vector, $\mathbb{1}(\cdot)$ is an indicator function that equals one if the condition in parentheses is true and zero otherwise. $K(\cdot)$ is the kernel function, (h) is the bandwidth, and (p) is the degree of the polynomial function. The magnitude of the discontinuity in the outcome regression is estimated by $\hat{\tau}_Y = \hat{\beta}_{+}^{Y}(h) - \hat{\beta}_{-}^{Y}(h)$, and the magnitude of the discontinuity in the treatment regression is estimated by $\hat{\tau}_D = \hat{\beta}_{+}^{D}(h) - \hat{\beta}_{-}^{D}(h)$. Therefore, the estimated effect $\hat{\tau}_{FRD}(h)$ using local polynomial RD estimator is the ratio of the two discontinuities

$$\hat{\tau}_{FRD}(h) = \frac{\hat{\tau}_Y(h)}{\hat{\tau}_D(h)} = \frac{\hat{\beta}_{+}^{Y}(h) - \hat{\beta}_{-}^{Y}(h)}{\hat{\beta}_{+}^{D}(h) - \hat{\beta}_{-}^{D}(h)} \quad (5)$$

Hahn et al. (2001) showed that the non-parametric estimates of equation (5) are numerically equivalent to the simple IV estimation. Therefore, the treatment effect in the RD setting can also be estimated by a parametric approach through the IV/2SLS method (Imbens and Lemieux, 2008; Lee and Lemieux, 2010).

Since the fuzzy regression discontinuity estimator given by formula (5) is a special case of the local Wald estimator, RD estimation and inference can be parametrically run by a simple IV

regression of Y on D and using T as an instrument for D. Consequently, and with no loss of generality, we can estimate equation (5) by using a local linear regression that requires running four regression on left (L) and right (R) side of the cut-off point. The following two equations can be used to estimate the outcome variable on the left and the right side of the cut-off

$$Y_L = \alpha_L + \delta_L(x_i - \bar{x}) + \varepsilon_{L,i} \quad i \in \{\bar{x} - h \leq X_i < \bar{x}\} \quad (6)$$

$$Y_R = \alpha_R + \delta_R(x_i - \bar{x}) + \varepsilon_{R,i} \quad i \in \{\bar{x} \leq X_i \leq \bar{x} + h\} \quad (7)$$

By combining the previous two equations into a unique local pooled linear regression, using the POM equation (1) and substituting T for D, we get:

$$Y_i = \alpha_L + (\alpha_R - \alpha_L) \cdot T_i + \delta_L(x_i - \bar{x}) + (\delta_R - \delta_L) \cdot T_i \cdot (x_i - \bar{x}) + \varepsilon_i \quad (8)$$

The equation estimates the discontinuity in the outcome at the cut-off point as the difference between the left and right intercepts of the regression, which is given by the coefficient of T.

By the same logic, the following two equations can be used to estimate the probability of retirement on the left and right side of the cut-off

$$D_L = P(D_i = 1 | x_i = \bar{x}) = \mu_L + \pi_L(x_i - \bar{x}) + \eta_{L,i} \quad i \in \{\bar{x} - h \leq X_i < \bar{x}\} \quad (9)$$

$$D_R = P(D_i = 1 | x_i = \bar{x}) = \mu_R + \pi_R(x_i - \bar{x}) + \eta_{R,i} \quad i \in \{\bar{x} \leq X_i \leq \bar{x} + h\} \quad (10)$$

In addition, in a fuzzy RD setting the compliers are all the units i that satisfy the rule

$$if\ T_i = 1 \Rightarrow D_i = 1$$

$$if\ T_i = 0 \Rightarrow D_i = 0$$

where $T_i = \mathbb{1}(x_i \geq \bar{x})$, assignment to treatment, and $D_i \in \{0,1\}$, receiving treatment.

Then the previous two equations can be combined into a unique pooled linear regression using the equation of observed treatment $D_i = T_i \cdot D_i(1) + (1 - T_i) \cdot D_i(0)$,

$$D_i = \mu_L + (\pi_R - \pi_L) \cdot T_i + \pi_L(x_i - \bar{x}) + (\pi_R - \pi_L) \cdot T_i \cdot (x_i - \bar{x}) + \eta_i \quad (11)$$

This equation estimates the discontinuity in the treatment (retirement) at the cut-off point as the difference between the left and right intercepts of the regressions, which is given by the coefficient of T, $\pi_R - \pi_L$. Therefore, a consistent estimator for the LATE in RD setting is

$$\widehat{LATE} = \frac{\hat{\alpha}_R - \hat{\alpha}_L}{\hat{\pi}_R - \hat{\pi}_L} \quad (12)$$

Equation (8) and (11) can be used as a reduced form two-equation structural system in which Y and D are endogenous and T is exogenous. By implementing a local polynomial regression, equation (12) can be estimated using the RD setting by the IV-2SLS method as follows (Hahn et al., 2001)

$$Y_i = \alpha_L + ATE \cdot D_i + \sum_{p=1}^{P} \delta_{L,p}(x_i - \bar{x})^p + T_i \sum_{p=1}^{P} (\delta_{R,p} - \delta_{L,p}) \cdot (x_i - \bar{x})^p + \mathbb{Z} + \varepsilon \quad (13)$$

$$D_i = \mu_L + (\mu_R - \mu_L) \cdot T_i + \sum_{p=1}^{P} \pi_{L,p} (x_i - \bar{x})^p + T_i \sum_{p=1}^{P} (\mu_{R,p} - \mu_{L,p}) \cdot (x_i - \bar{x})^p + \mathbb{Z} + \eta$$

where ATE is the parametric RD average treatment effect, D_i is a dummy variable where $D_i = 1$ if the respondent received the treatment (i.e., is retired), T_i is an instrument for D, p is the polynomial order of the regression, and $\mathbb{Z}$ is a vector of covariates in the model.

Although instrumental variable estimation methods are widely used in econometrics to capture the endogeneity issue, the bivariate likelihood estimation methods are theoretically and empirically superior to the traditional two-stage instrumental variable in the case of binary outcome settings (Marra et al., 2014; Bhattacharya et al., 2006; Wooldridge, 2010). The recursive bivariate probit model is a convenient setting for estimating the effect of retirement as an endogenous binary covariate ("retirement") on the binary health outcome indicators. The setup of this estimation is based on two latent variable equations:

$$retired_i = \mathbb{1}(\beta_0 + \beta_1\, cutoff_i + \beta_2 G_i + \beta_3 G_i * cutoff_i + \gamma\, X_i + u_i > 0);$$

0, otherwise (14)

$$health_i = \mathbb{1}(\alpha_0 + \alpha_1\, cutoff_i + \alpha_2 G_i + \alpha_3 G_i * cutoff_i + \pi\, retired_i + \psi\, X_i + \varepsilon_i > 0);$$

0, otherwise

where $\mathbb{1}(\cdot)$ is an indicator function that equals one if the condition in parentheses is true and zero otherwise, X_i is a vector of individual characteristics and demographic variables, and π is a consistent estimator of the Local Average Treatment Effect (LATE) on health among retirees. In the model, age is normalized to be $G_i = (age_i - C)$, where C = 65, so that discontinuity arises at $G_i = 0$. The variable $cutoff_i$ is equal to one if $G_i \geq 0$ and zero otherwise. That is, it is a dummy variable that takes on a value of "1" if individual "i" is on the right of the cut-off point of the treatment. The interaction term $G_i * cutoff_i$ allows different slopes on either side of the discontinuity. That is, age is centered at the cut-off point to interpret any shift at the cut-off point as a shift in the intercept and the interaction term as a change in the slopes on either side of the cut-off point. The first equation determines whether or not treatment is received, while the second equation describes the health outcome as a function of a binary treatment ("retirement") and latent error. The recursive bivariate probit model is used to estimate the impact of retirement on health behavior outcome variables.

It is very important to note that including fixed effects in the specification of RD design is not necessary. Lee and Lemieux (2010); Van Der Klaauw (2008); Imbens and Lemieux (2007) show that the RD analysis can be conducted for the entire pooled cross-section dataset, and the within-individual correlation of the errors over time can be captured by using clustered standard errors. They also show that the source of identification in the RD setting is a comparison between those just below and above the cut-off point, and this can be done with a

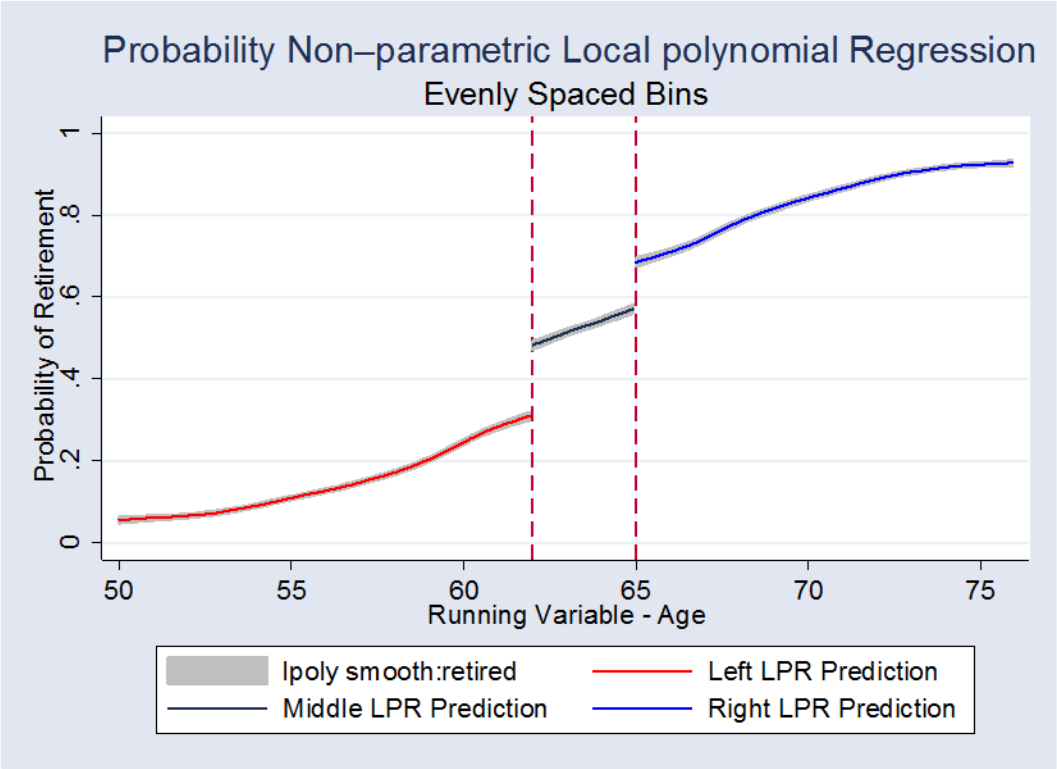
single cross-section. Consequently, imposing a specific dynamic structure introduces more restrictions without any marginal gain in identification (Lee and Lemieux 2010).

1.5. Regression Discontinuity Visualization:

One of the distinctive features of RD design is the possibility to visualize the discontinuity graphically. Graphical illustration associated with formal estimation, inference, and falsification tests enriches the transparency of the analysis by visualizing the observations used in the estimation. Although individuals in the U.S. are not forced into retirement, the official pension claims schemes provide incentives to retire at certain ages, i.e. early and full retirement age at 62 and 65, respectively.

I first visually inspect if there is an empirical discontinuity in retirement at the official age of retirement. In the following visualization, I will show the conditional probability of being assigned to the treatment given the forcing variable, $pr[T_i = 1 \,|X_i = \bar{x}]$, for different values of the running variables x_i. Recall that in the fuzzy RD design, treatment assignment and treatment received are not identical, and the figure reflects treatment take up, not treatment assignment. Figure (1.1) shows the share of retirement at every age between 50 and 75 in the observed data set. The solid dots indicate the local sample mean of retirement over non-overlapping 3-month bin partitions of the actual data. The sample means help visualize the dispersion of data which can be used to detect any potential discontinuity. The figure shows that there is a marked increase in retirement at age 65, which is the full retirement age (the official age of retirement). There is also a noticeable jump in the probability of retirement at age 62, the early retirement age.

Figure 1.1: Age profile of retirement using non-parametric Local Polynomial Regression



The optimal choice of the number of bins is based on IMSE- optimal evenly spaced partitions bins.

Figure 1.1 is constructed by using the RD plot that has evenly spaced bins that mimic the underlying variability of the actual data. The global polynomial is constructed using a 4^{th} degree polynomial. The graph shows how the local binned sample means approximate the underlying regression function. In this approach, the optimal number of bins is selected to balance squared bias and variance. The graph shows distinctive discontinuities at age 62 and age 65. The estimated 95% confidence intervals do not intersect at the threshold points, which is strong visual evidence for the validity of discontinuities. RD design is a research strategy that can be applied when selection into treatment is highly determined by specific levels of particular variables, called running or forcing variables, which are used to set up a threshold point to identify treated and untreated units. Therefore, ages 62 and 65 are good candidates for the discontinuity in the retirement decision.

Figure 1.2: The change in the retirement probability using the fixed effect estimator

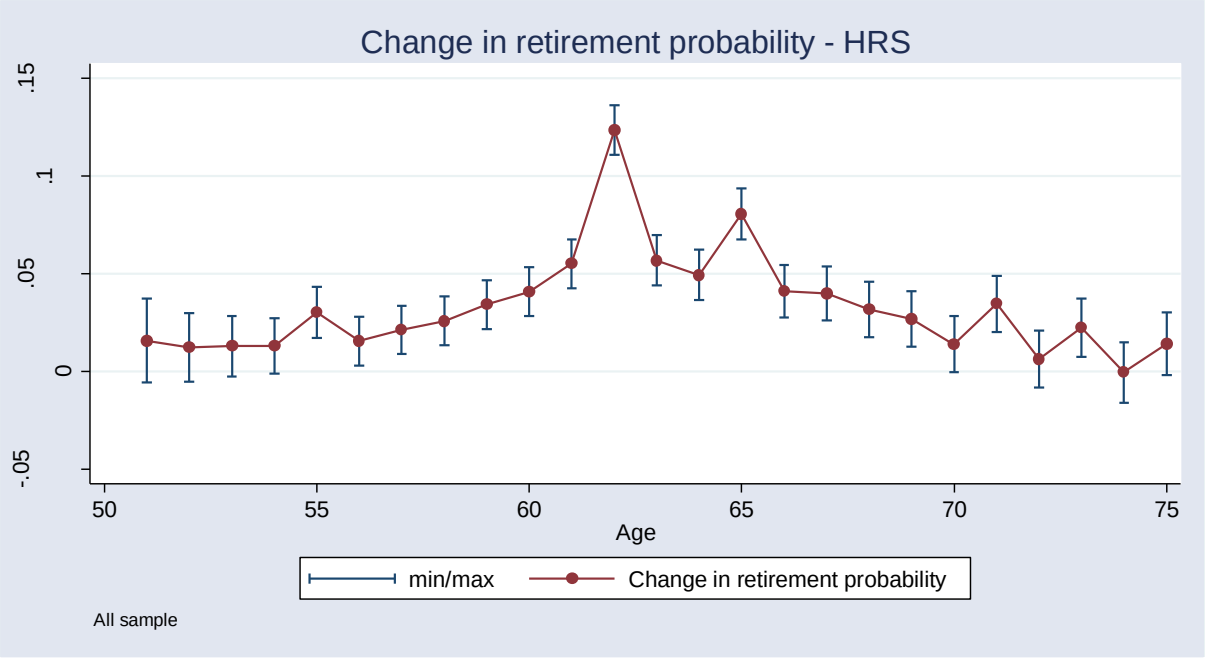


Figure 1.2 represents the estimated change in retirement probability between age 50 and 75 after controlling for individual fixed effect. Following Bonsang et al. (2012), I estimated the following model $\Re_{it} = \rho_i + \sum_{g=51}^{75} \phi_g Z_{it}^g + \xi_{it}$, where $\Re_{it}$ is the retirement dummy, $Z_{it}^g = 1[age_{it} \geq g]$, ρ_i is the individual fixed effect, and ξ_{it} is the error term. The figure reports the estimates of $\phi_g$with 95% confidence intervals. The figure shows that there is a significant increase in the probability of retirement at age 62, the early retirement age, and there is a noticeable increase in the probability of retirement at age 65, the official age of retirement.

In general, Figures 1.1 and 1.2 show that the probability of retirement increases as individuals get older. In addition, there are significant discontinuities in the probability of retirement at ages 62 and 65 for the full sample. The sample means inside bins and the local polynomial smoothing of the data shows that there is a greater jump in the retirement probability at age 62 (early retirement age) than there is at age 65 (the official age of retirement). Since there is no documented reason to believe that individuals' health would change drastically when they

turn 62 or 65, this suggests that factors such as social norms or financial incentives provided by pensions plans might cause these jumps. According to Thistlethwaite and Campbell (1960), Lee (2007), Card et al. (2006), Chay and Greenstone (2005), and Lee & Lemieux (2010), the validity of the RD design does not depend on the cause of the jump but on whether there are true jumps in the probability of retirement at the cut-off points.

The next step is to investigate if there are discontinuities in the outcome variables. At first glance, the relationship between the outcome health variables and the running variable can be illustrated by simply constructing a scatter plot of the observed outcome health variable against the score variable. However, it is hard to see jumps or discontinuities in the outcome-score relationship by simply looking at the raw data. The more useful approach in RD design is to smooth the data before plotting. Following Calonico et al., (2017), global polynomial fit and local sample means are combined to show the RD plot. Global polynomial fit is a smooth approximation to the unknown regression function that fits the original observed outcome variable against the score variable above and below the cut-off separately using the raw data. In contrast, the local sample means are created by determining the number of bins of the running variable and calculating the mean of the health outcome for all observation within each bin and then plotting the average outcome in each bin against the midpoint of the bin. Suppose that the total number of bins ($\mathcal{W}_j$) is $\mathcal{J}_-$ and $\mathcal{J}_+$ for the number of bins chosen to the left and right of the cut-off point, respectively. Each local sample mean ($\bar{Y}$) for each bin ($\mathcal{W}_j$) is computed as follows:

$$\bar{Y}_{-,j} = \frac{1}{\#\{x_i \in \mathcal{W}_{-,j}\}} \sum_{i:x_i \in \mathcal{W}_{-,j}} Y_i \text{ and } \bar{Y}_{+,j} = \frac{1}{\#\{x_i \in \mathcal{W}_{+,j}\}} \sum_{i:x_i \in \mathcal{W}_{+,j}} Y_i \quad (14)$$

where $\mathcal{J} = 1, 2, 3, \cdots, j$ is the number of bins

In RD design, there are two different approaches to construct the RD plot, and both of them address two issues regarding the location and optimal numbers of the bins. The first issue is how to determine the location of the bins. There are two types of bins that can be used to construct the RD plot: evenly-spaced bins (ESB), where the entire support of the cut-off point is partitioned into non-overlapping intervals with the same length; and quantile-spaced bins (QSB), where the entire support of the cut-off point is partitioned into non-overlapping intervals with the same number of observations. According to CCT, right and left bins can be defined as follows

$$\mathcal{W}_{-,j} = \begin{cases} [x_l\ , w_{-,1}) & j = 1 \\ [w_{-,j-1}\ , w_{-,j}) & j = 2, \cdots\cdots, \mathcal{J}_- - 1 \\ [w_{-,j_- -1}\ , \bar{x}) & j = \mathcal{J}_- \end{cases}$$

$$\mathcal{W}_{+,j} = \begin{cases} [\bar{x}, w_{-,1}) & j = 1 \\ [w_{+,j-1}\ , w_{-,j}) & j = 2, \cdots\cdots, \mathcal{J}_+ - 1 \\ [w_{+,\mathcal{J}_+ -1}\ , x_r] & j = \mathcal{J}_+ \end{cases} \qquad (15)$$

where $\cup_{j=1}^{\mathcal{J}_-} \mathcal{W}_{-,j} = [x_l, \bar{x})$, $\cup_{j=1}^{\mathcal{J}_+} \mathcal{W}_{+,j} = [\bar{x}, x_r]$, and the set $[x_l, x_r]$ is the support of the running variable centered at the cutoff $\bar{x}$.

Partitioning estimators of the evenly-spaced bins is given by

$$w_{-j} = x_l + \frac{j(\bar{x} - x_l)}{\mathcal{J}_-} \; and \; w_{+j} = \bar{x} + \frac{j(x_r - \bar{x})}{\mathcal{J}_+} \qquad (16)$$

and the partitioning estimators of the quantile-spaced bins is given by

$$w_{-j} = X_{-,}\left(\left\lfloor \frac{j}{\mathcal{J}_-} \right\rfloor\right) \; and \; w_{+j} = X_{+,}\left(\left\lfloor \frac{j}{\mathcal{J}_+} \right\rfloor\right) \qquad (17)$$

where $X_{-,}(\cdot)\; and \; X_{+,}(\cdot)$ denotes the i^{th}quantile of the control and treatment subsample, respectively, and $\lfloor \cdot \rfloor$ denotes the floor function.

The second issue is how to determine the number of bins that can be used in the case of ES and QS. There are again two approaches to determine the optimal number of bins. The first approach is the Integrated Mean Square Error (IMSE) of the local mean estimator. According to this approach, when we choose a large number of bins, we get a small bias, but larger numbers of bins may lead to more variability within bins due to the smaller number of observations inside each bin. Therefore, according to IMSE, the optimal number of bins is the one that balances squared-bias and variance and minimizes IMSE. The IMSE optimal numbers of bins for $\mathcal{J}_-$ and $\mathcal{J}_+$ are given by the following selectors

$$\mathcal{J}_- = \lceil \varphi_- n^{1/3} \rceil \; and \; \mathcal{J}_+ = \lceil \varphi_+ n^{1/3} \rceil \quad (18)$$

Where n is the total number of observations, and $\lceil \cdot \rceil$ denotes the ceiling operator. The constant φ_- and $\varphi_+$are unknown and their values depend on whether ES or QS partitioning selectors of bins are used. The unknown constants are estimated using the CCT procedure, which is based on some features of the underlying data generating process[6].

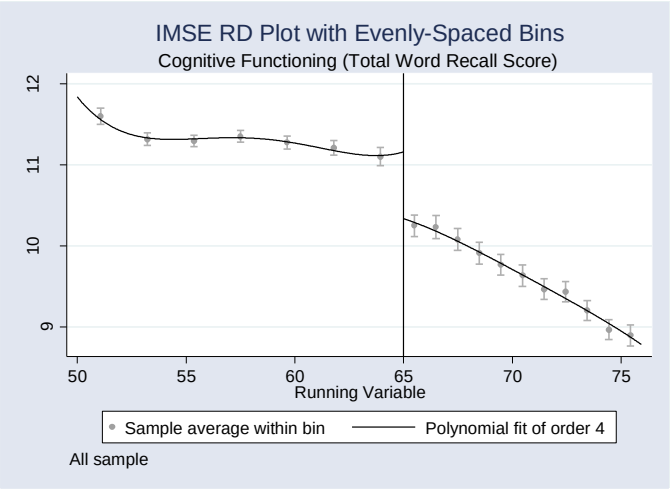
The second approach is the mimicking variance approach, where the optimal number of bins is chosen so that the overall variability of the binned means mimics the overall variability in the raw data (Calonico et al., 2017). The mimicking-variance selectors of the number of bins use the same approach of the IMSE selection, except that the form of the constant φ_-^{MV} and φ_+^{MV} are different from those included in the IMSE optimal selector. The MV procedure leads to a larger number of bins than the IMSE procedure for both ES and QS. Therefore, the number of dots, which represents the local mean inside each bin, is higher and this can give a clearer picture of the variability of the data.

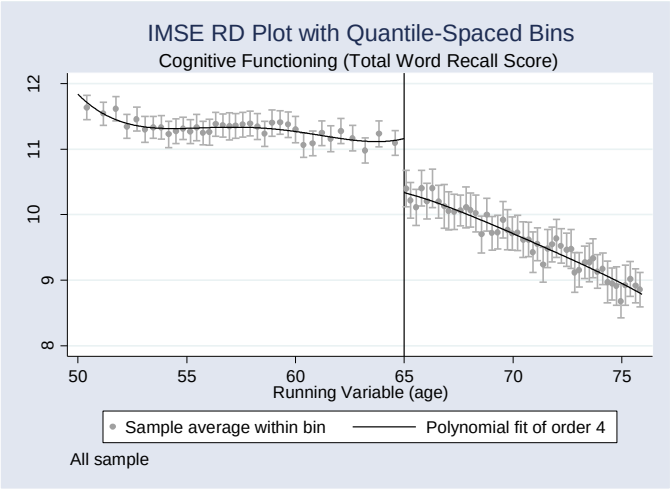
[6] All details regarding how the selectors of the number of bins for control and treatment groups, and the rule of thumb estimates of the unknown constants can be found in Calonico, Cattaneo, and Titiunik (2015).

In Figure 1.3, the global fit shows that the observed regression function is close to a linear relation. The local average of the outcome 'cognitive function' in each bin shows how the average total word recall score behaves around the global fit. The plot visualizes the discontinuity at the cut-off point for complicity and helps in understanding the shape of the underlying regression function over the support of the running variable. The local means of total word recall in panels A and B are similar. Since the observations are uniformly distributed on the support of the running variable, the local mean in each bin is precisely calculated and the local average will be directly comparable in terms of variability. The plot also reveals a negative jump in the value of the total word recall index at the cut-off point. That is, the cognitive functioning of retired individuals is lower than that of non-retired individuals within the same age group. Figure 1.4 shows that there is a significant positive jump in the CESD depression scale at the cut-off point for compliances for both Evenly and Quantile spaced methods. The RD plots give an initial intuition about the negative effect of retirement on the mental health of retirees. Figure 1.5 shows a negative jump in self-reported health for retirees in the neighborhood of the cut-off point. That is, the proportion of retired individuals who reported to be in good health is less than the proportion of non-retired individuals who reported to be in good health: approximately 85% of non-retirees and 73% of retirees.

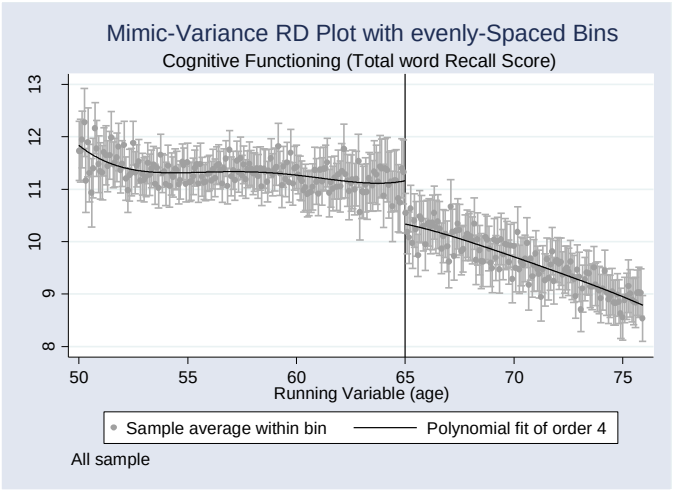
Figure 1.6 shows a significant negative jump in the proportion of retired individuals who drinking alcohol. The figure shows that retired individuals are around 5 percentage points less likely to drink. In contrast, the RD plot (Figure 1.7) shows a positive jump in the proportion of retired individuals who smoke relative to non-retirees.

Figure 1.3: Cognitive Functioning (Total Word Recall Scale) – RD plot
Panel A: Integrated Mean Square Error (IMSE) approach





Panel B: Mimicking Variance (MV) approach



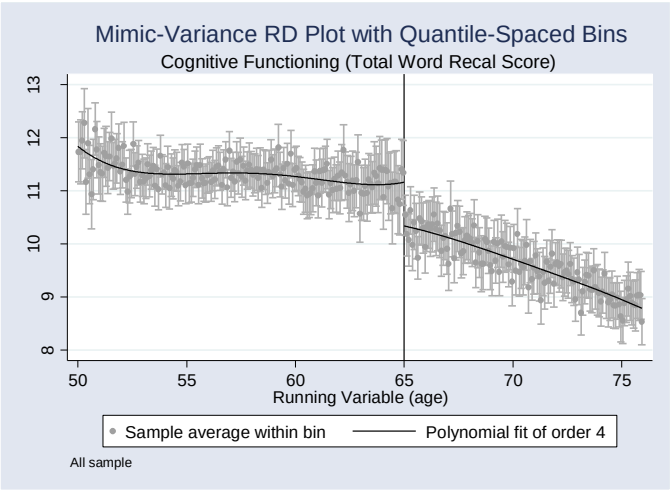
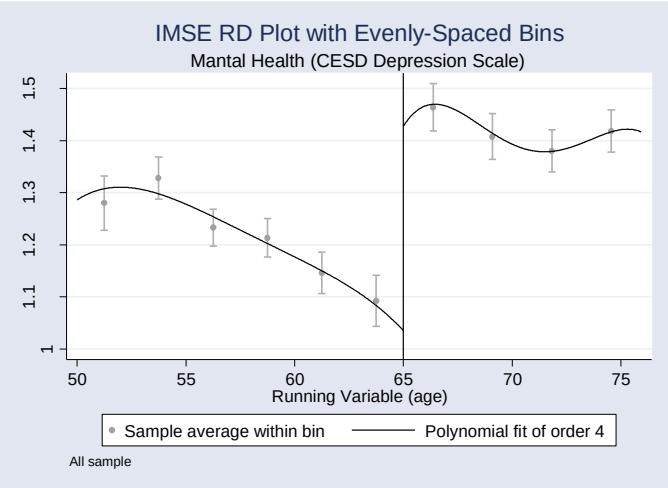
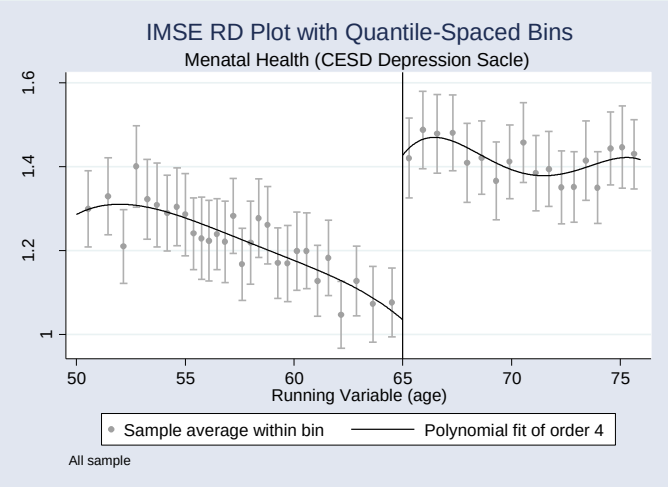
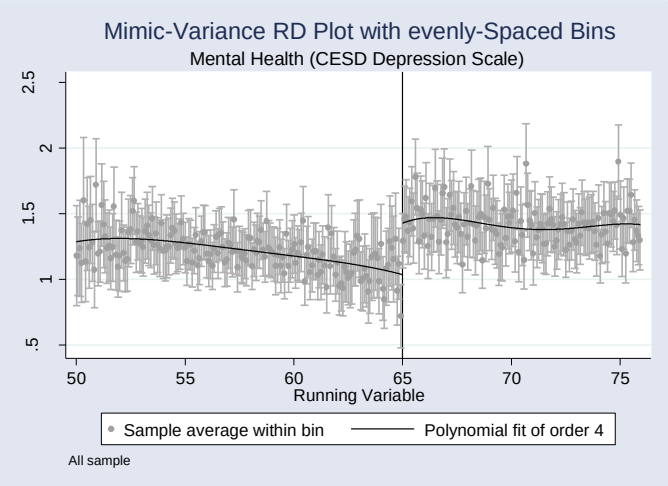


Figure 1.4: Mental Health – CESD depression scale – RD plot
Panel A: Integrated Mean Square Error (IMSE) approach





Panel B: Mimicking Variance (MV) approach



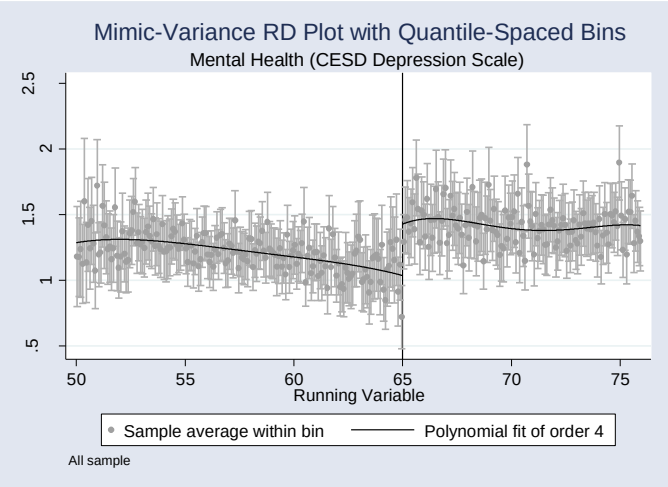
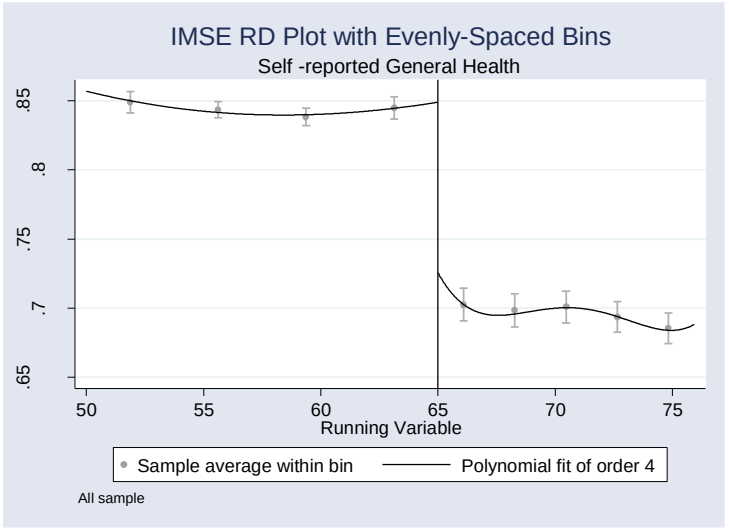
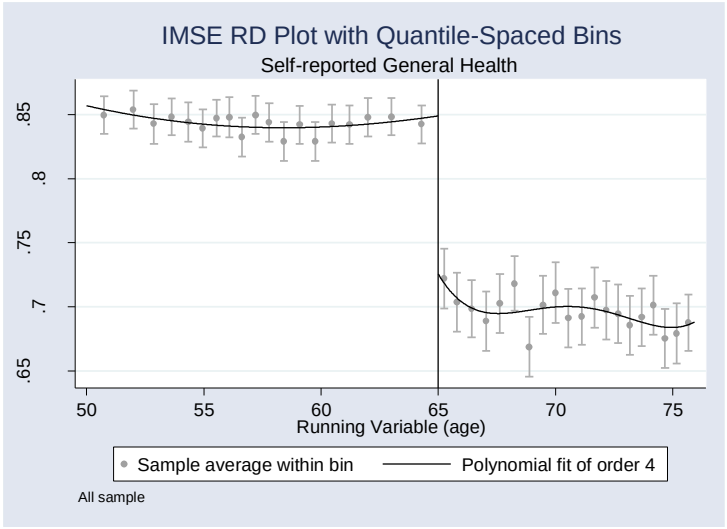
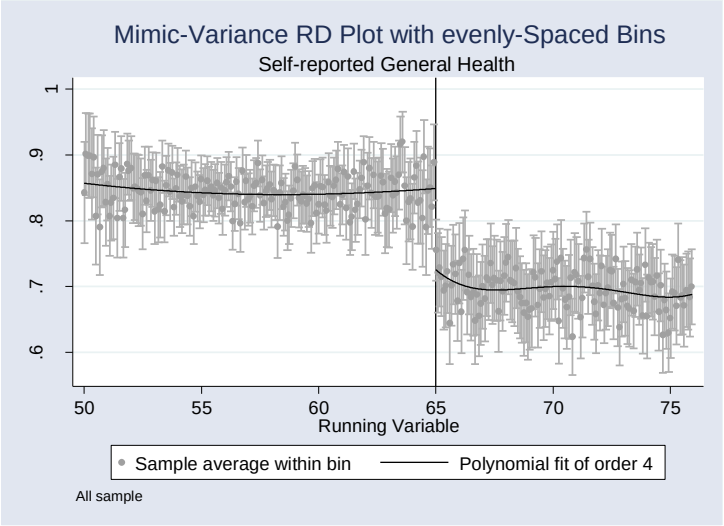


Figure 1.5: Self-reported General Health – RD plot
Panel A: Integrated Mean Square Error (IMSE) approach





Panel B: Mimicking Variance (MV) approach



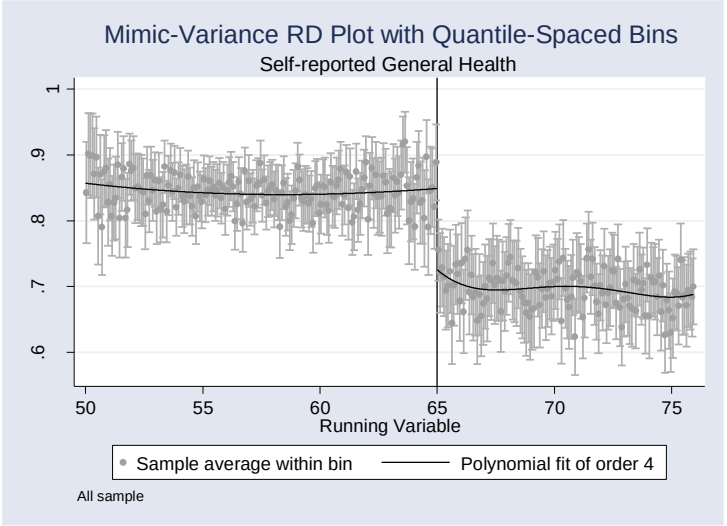
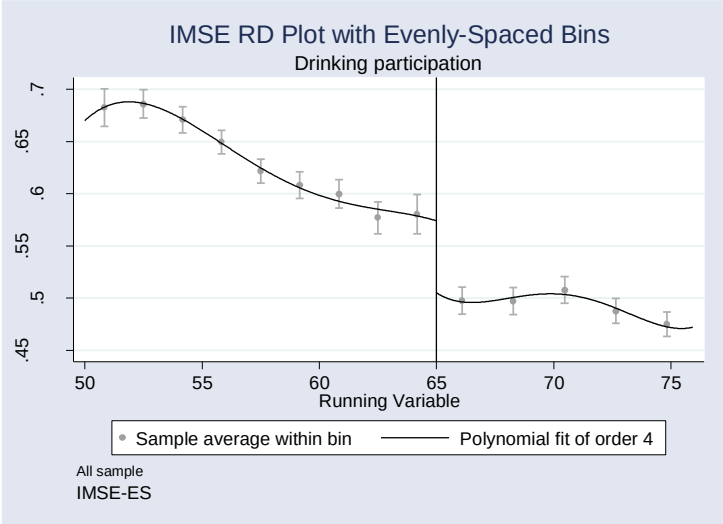
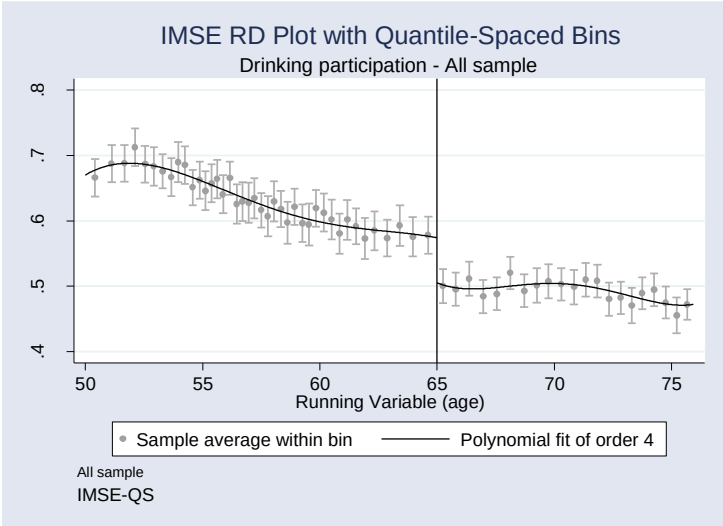


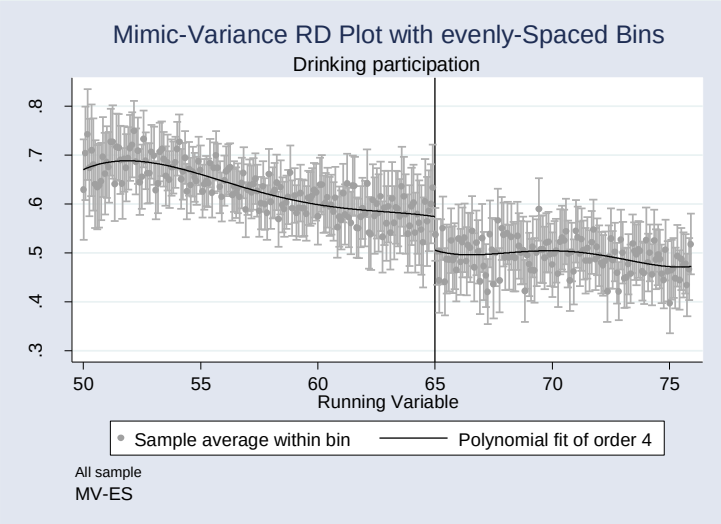
Figure 1.6: Drinking participation – RD plot

Panel A: Integrated Mean Square Error (IMSE) approach





Panel B: Mimicking Variance (MV) approach



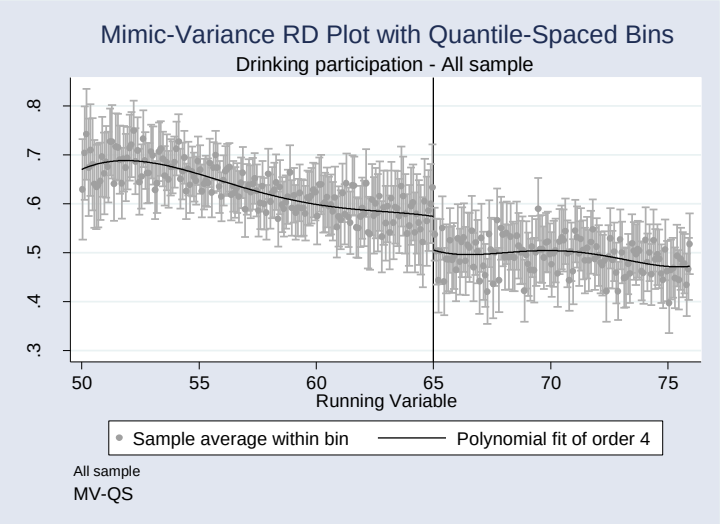
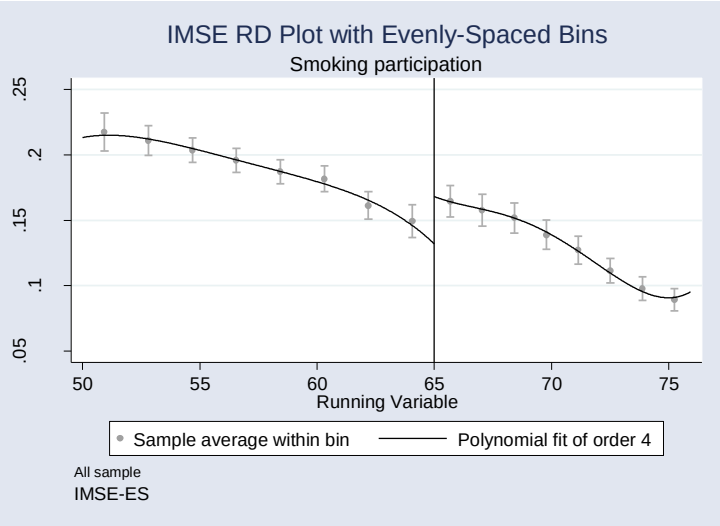
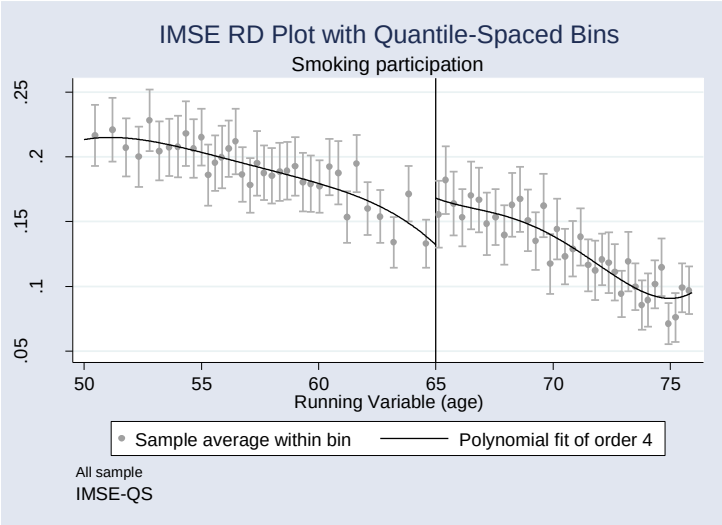


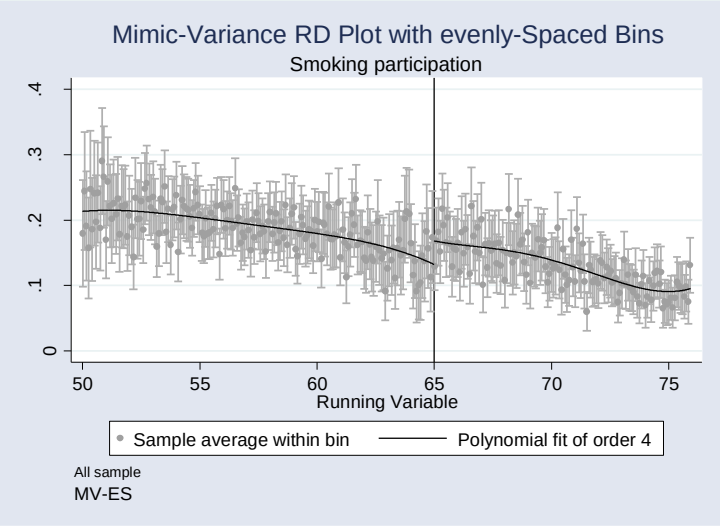
Figure 1.7: smoking participation – RD plot

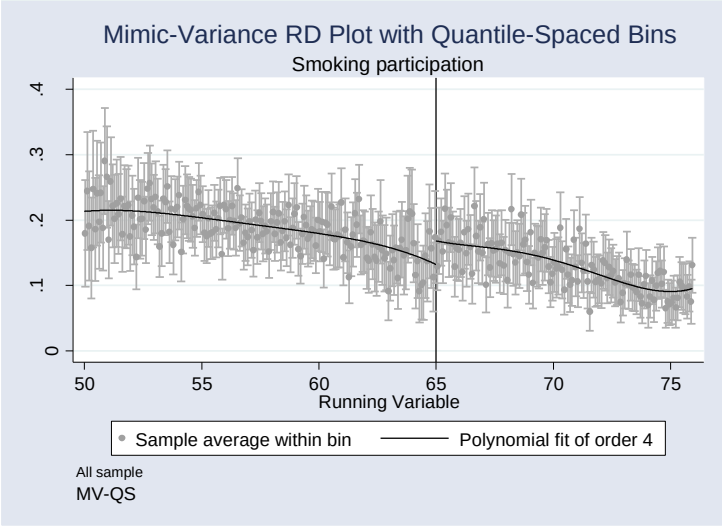
Panel A: Integrated Mean Square Error (IMSE) approach





Panel B: Mimicking Variance (MV) approach





1.6. Non-parametric Estimation Results of the Fuzzy Regression Discontinuity Design (FRD)

1.6.1 Non-Parametric RD Results of Health

Table 1.1 presents the main results of the non-parametric RD estimation of the impact of retirement on cognitive functioning, measured by the total word recall score (TWC). Each column shows the local linear point estimate using the bandwidths indicated in the first row of each column. The first column reports the RD treatment effect using the optimal bandwidth, which is selected according to the Mean Square Error (MSE) bandwidth selector. The estimates in the first column show that the jump in retirement probability was associated with a negative jump in the reported word recall test score. That is, the effect of retirement on cognitive function is negative and highly significant. The RD estimator suggests that retired individuals experienced a drop in their cognitive test score by about 0.8 points with a 95% confidence interval [-1.03411, -0.593943]. This decline in cognitive functioning is equivalent to an approximately 8% decline in the cognitive functioning score of retired individuals. Moreover, the effect of retirement on cognitive functioning is not sensitive to the change in the selected bandwidth for all reported bandwidths.

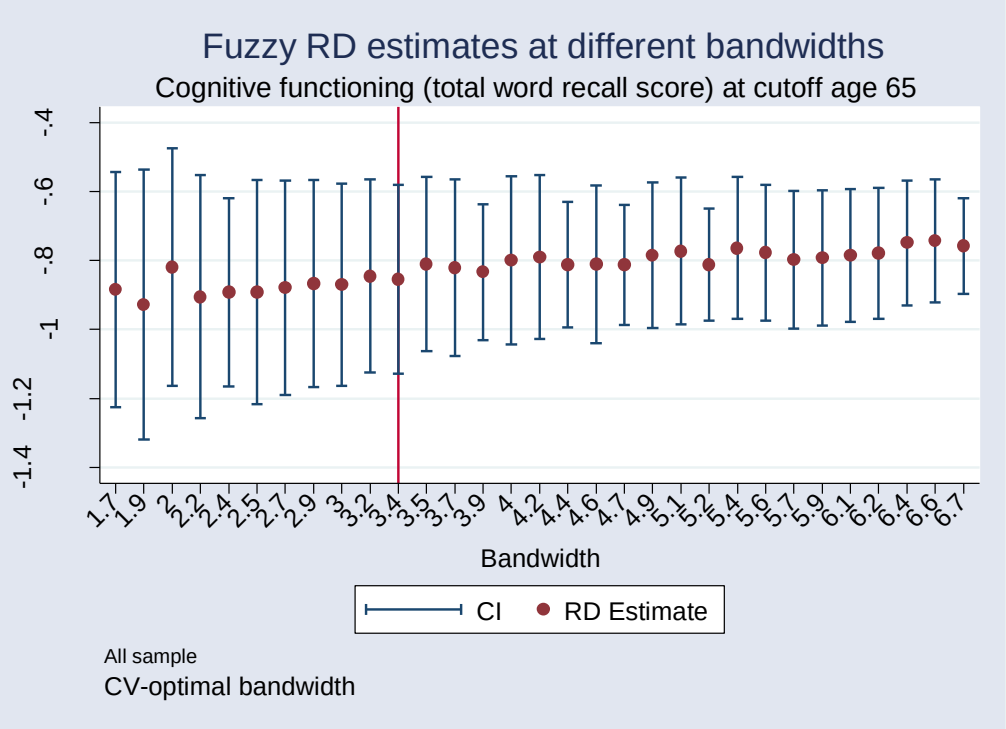
The table also shows results for three different methods for the RD treatment effect estimator: the traditional RD estimates with conventional variance estimator (conventional inference), Bias-corrected RD estimates with a conventional variance estimator (Bias-corrected inferences), and bias-corrected RD estimates with a robust variance estimator (Robust RD inferences), which is proposed by Calonico, Cattaneo, and Titiunik (2014, CCT hereafter) and Calonico et al., (2016) to give more robust confidence intervals by using a bias-corrected RD treatment effect estimator.

Table 1.1: Fuzzy RD estimates - Cognitive functioning (Total word recall score) – All sample at age 65

	MSE-Optimal bandwidth	Bandwidths (h)				
	4.8	1	2	3	4	5
Conventional	-0.81403***	-0.99274***	-0.88438***	-0.87222***	-0.82654***	-0.8142***
	(0.11229)	(0.29964)	(0.19173)	(0.15071)	(0.12601)	(0.11054)
$CI^{c}_{95\%}$	[-1.03411, -0.593943]					
Bias-Corrected	-0.86067***	-1.4924***	-0.92883***	-0.92134***	-0.9233****	-0.86472***
	(0.11229)	(0.29965)	(0.19173)	(0.15071)	(0.12601)	(0.11054)
$CI^{bc}_{95\%}$	[-1.08073, -0.640564]					
Robust	-0.86065***	-1.4924***	-0.92883***	-0.92134***	-0.9233***	-0.86472***
	(0.13365)	(0.47549)	(0.28369)	(0.22048)	(0.18557)	(0.16322)
$CI^{rbc}_{95\%}$	[-1.1226, -0.598694]					
Obs L|R	38927| 28150	38927|28150	38927|28150	38927|28150	38927| 28150	38927| 28150

Notes: (i) All estimates are computed using a triangular kernel (ii) The first column shows three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) Standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.8: Non-parametric Fuzzy RD estimates at different bandwidths – cognitive function



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for cognitive functioning. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. Bootstrapped standard errors are based on 120 simulations to construct 95% CI.

The first two rows of the table show the conventional RD estimator $\hat{\tau}_{FRD}^{c}\ (h)$, which is based on local polynomial non-parametric estimators and a 95% confidence interval. This approach is arguably the most commonly used in practice (Calonico, 2014). The bias-corrected approach ($\hat{\tau}_{FRD}^{bc}\ (h)$) is introduced to remove the potentially large effect of unknown leading bias of the RD estimator (Calonico, 2014). The robust bias-corrected approach ($\hat{\tau}_{FRD}^{rbc}(h)$) offers an alternative confidence interval $CI_{95\%}^{rbc}$on the bias-corrected RD treatment effect estimator and the adjusted variance that account for the additional variability in the bias-corrected RD. That is, the variance formula used is constructed to account for the variability of the original RD treatment effect estimator and the bias correction term[7]. The results show that the RD effect of retirement on cognitive functioning is negative and stable across the three RD inference methods with a small difference in the standard error in the robust bias-corrected approach relative to the other two approaches.

The choice of the neighborhood of the cut-off point in RD estimation $[\bar{x}-h, \bar{x}+h]$ is crucial in getting consistent estimates for the treatment effect. Theoretically, the bandwidth choice in RD design was first introduced by Imbens and Kalyanaraman (2009, IK hereafter). They developed the MSE optimal bandwidth choice, which is extended further to fit the RD estimation using local polynomial estimators, clustered data, and inclusion of covariates, in a sequence of recent papers (Calonico et al., 2014b; Calonico et al., 2016; Cattaneo and Vazquez-Bare, 2016; Bartalotti and Brummnt, 2017). Generally, smaller bandwidth tends to produce lower bias and higher variance. In contrast, larger bandwidth tends to produce higher bias and lower variance. Moreover, the optimal bandwidth in the case of fuzzy RD design will tend to be

[7] For further details about the theoretical implication of these alternative RD approaches see Calonico, Cattaneo, and Farrell (2014

larger due to the additional variance that arises from the estimation of the jump in the conditional mean of the treatment (retirement). Put all together, larger bandwidth in fuzzy design may lead to additional bias which, in turn, depends on the curvature of the conditional mean functions. Therefore, the fuzzy estimates at different bandwidths could help investigate how the RD treatment effect estimates depend on the chosen bandwidth.

Figure 1.8 facilitates visualizing whether or not the RD estimates depend on the chosen bandwidth. The optimal bandwidth choice is based on the IM selector, a fully data-driven optimal bandwidth[8]. Generally, in this diagram, the MSE is defined as

$$MSE^i\,(h) = E[\hat{\tau}^i_{RD} - \tau^i_{RD}]^2 = E\left[\left(\hat{\jmath}^i_+ - \jmath^i_+\right) - \left(\hat{\jmath}^i_- - \jmath^i_-\right)\right]^2 \;; \quad i \in \{Y, D\} \qquad (19)$$

The optimal bandwidth that minimizes the MSE criterion is

$$h^i{}_{IM} = \arg\min MSE^i(h)$$

That is, the optimal bandwidth (h_{IM}) is the one that minimizes the mean square of the difference between the estimated $\hat{\tau}^i_{RD}$ and actual value of τ^i_{RD}, where the observations on the left-hand side of the cut-off point, ($X_i < \bar{x}$) used in the regression are $\bar{x} - h \leq X < \bar{x}$, and the observations on the right-hand side of the cut-off point ($X_i \geq \bar{x}$) used in the regression are $\bar{x} \leq X \leq \bar{x} - h$. In the FRD setting, there are four non-parametric regressions, two for the treatment variable and two for the outcome variable, on both sides of the cut-off point. Following Imbens and Kalyanaraman (2009), we have to choose the optimal bandwidth that minimizes $MSE^Y(h)$ for the outcome variable and minimizes $MSE^D(h)$ for the treatment variable. Therefore, the optimal bandwidth for fuzzy RD is given by

$$h^*_{IM} = min \quad \{arg\,min\,MSE^Y(h) \;\; , arg\,min\,MSE^D(h)\} \qquad (20)$$

[8] See Imbens and Kalyanaraman (2009) for all the details related with the formulas and algorithms of choosing the optimal bandwidth in RD setting.

The figure shows that the estimated effect of retirement on cognitive functioning is negative and significant at all bandwidths in a range between 50% and 200% of the estimated optimal bandwidth with 5 percentage points increments. The graph also confirms that variability among retired individuals is consistent across the different bandwidths, which increases the efficiency of the estimated effect.

One challenge in the non-parametric RD setting is the choice of the kernel function: how observations close to the cutoff point are given a relatively high weight and observation far from the cut-off point a relatively low weight. I used three different kernel functions: the triangular kernel function, which is the most common form in the RD setting; a uniform kernel function; and the Epanechnikov kernel function. The goal of estimating the RD treatment effect under different kernel functions is to investigate if the relative weight given to the neighborhood observations around the cut-off affects the estimated treatment effect. Table 1.2 summarizes a variety of information organized in five columns. The first two columns show the RD treatment effect using uniform and Epanechnikov kernel functions, respectively. The estimation is conducted using local linear estimators ($\rho = 1$). The results in the first two columns show that retirement has a significant negative effect on retirees' cognitive functioning. The results show a drop in retirees' cognitive score by 0.82 and 0.79 points when using the uniform and Epanechnikov kernel functions, respectively. Moreover, when the uniform and Epanechnikov kernel functions are used with the robust bias-corrected RD estimators, there was a drop in retirees' cognitive score by 0.86 and 0.84 points with robust 95% confidence intervals [-1.15481,-0.575584] and [-1.09997,-0.585606], respectively. The results with different kernel functions show that the retirement effect on cognitive functioning is consistent with the RD estimation using a triangular kernel.

Table 1.2: RD estimates for different kernel and different polynomial orders – Cognitive functioning (Total word recall score) at cut-off age 65

RD estimator	1	2	3	4	5
Conventional	-0.82047***	-0.79777***	-0.88619***	-0.89449***	-0.87022***
Conventional Std. Err.	(0.12288)	(0.1105)	(0.14525)	(0.17594)	(0.14827)
Robust Bias-corrected	-0.8652***	-0.84279***	-0.92422***	-0.89063***	-0.90864***
Robust Std. Err.	(0.14776)	(0.13122)	(0.16724)	(0.20248)	(0.17029)
Robust 95% CI	[-1.15481, -0.575584]	[-1.09997, -0.585606]	[-1.25201, -0.596431]	[-1.28748, -0.493782]	[-1.2424, -.57489]
Kernel Type[9]	Uniform	Epanechnikov	Triangular	Uniform	Epanechnikov
Order Loc policy (p)	1	1	2	2	2
BW local poly (h)	3.4	4.6	6.5	4.1	5.9
BW Type	mserd	mserd	mserd	mserd	mserd
Observations	67077	67077	67077	67077	67077
L|R	38927|28150	38927|28150	38927|28150	38927|28150	38927|28150
effective no of obs	6100|8235	9434|11319	15020|16053	7891|9879	13198|14518

Notes (i) First column report two different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (ii) Standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

[9] In Triangular Kernel function, the weight is maximized at the cutoff point, $X_i = \bar{x}$, and decline symmetrically and linearly as the value of the score gets farther from the cutoff. In Epanechinkov Kernel function, the quadratic weight is maximized at the cutoff point and decline symmetrically as the value of the score gets farther from the cutoff. In Uniform Kernel function, all observations have the same weight.

The other challenge in the RD setting is the order of the polynomial function. Although a local linear estimator has become the standard choice in the RD setting literature since the seminal work of Hahn et al., (2001), it was not the best strategy in all social studies (Card and Pei, 2014). Therefore, I tried to be more flexible in the choice of the polynomial order p. Following recent findings in RD design, Gelman and Imbens (2014) argue that RD estimators for causal effects, which rely on using high-order polynomial approximations for the conditional mean functions (third, fourth or higher), have been proven to be sensitive to the order of the polynomial and conventional inference has poor performance in these settings. They also point out that higher-order polynomials may assign a very large weight to observations far from the threshold point, which is an undesirable property of high-order polynomials[10]. However, I checked different polynomial orders for the running variable and found that there is no improvement in the MSE after the second-order polynomial, which is in line with the argument of Gelman and Imbens (2014). Therefore, I reported the results of the RD under the linear and quadratic polynomial smoothing function. This choice is consistent with the work of Bonsang et al., (2014), who assumed that second-order degree form accounts for the normal cognitive aging decline process, i.e. quadratic age may allow cognitive functioning to decline at an increasing rate with age. The last three columns report the RD estimates using local quadratic estimators ($p = 2$) with different kernel functions. Overall, the RD treatment effect is stable across the three different kernel functions. Retirement has a significant negative effect on retirees' cognitive score, which ranges from 0.86 to 0.92. That is, the decrease in retirees' cognitive score maybe around 8.6% and 9.2% according to the nonparametric RD local quadratic estimator

[10] For more details, see Gelman, A., and Imbens, G. (2014). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs.

under the conventional and robust bias-corrected variance. Moreover, the value of the RD estimate, using the quadratic estimate is similar to the linear estimate.

Up to this point, the RD treatment effect estimator has assumed that the non-parametric local polynomial fit includes only the running variable, age, as the regressor. Although, according to many researchers (Imbens & Lemieux, 2008; Lee & Lemieux, 2010), this specification is sufficient in the RD setting, Calonico, Cattaneo, Ferrell and Titiunik, 2018 (hereafter CCFT) suggested a formal framework for estimation and inference in RD designs when covariates are included in the local polynomial procedure. They argue that augmented covariates in RD design may achieve substantial efficiency gain relative to the standard RD estimator.

Table 1.3: Covariate-adjusted RD estimates – Cognitive functioning (Total word recall score) whole sample at cut-off age 65

	Without covariates	With covariates
	Standard Fuzzy RD	Covariate-Adjusted Fuzzy RD
Inference with h/p unrestricted		
RD treatment effect	-0.86065***	-0.70257***
Robust 95% CI	[-1.1226, -0.598694]	[-.948205, -0.456944]
CI length change (%)		-6.2
robust p-value	0.000	0.000
h/p	0.645	0.641
Inference with h/p restricted		
RD treatment effect	-0.92134***	-0.77384***
Robust 95% CI	[-1.35347, -.489204]	[-1.15646, -0.391214]
CI length change (%)		-11.45
robust p-value	0.000	0.000
$h\|p$	2|2	2|2
$n-\|n+$	38927|28150	38927|28150
observations	67077	67077
effective obs	9895|11793	5650|7864

Notes: (i) All estimates are computed using a triangular Kernel and nearest neighbor heteroskedasticity-robust variance estimator as suggested by CCT framework. (ii) Bandwidth used for *h and b* is data-driven MSE-optimal for either standard RD estimator or covariate-adjusted RD estimator. (iii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.3 shows the main results of the standard and covariate-adjusted RD estimate, which includes covariates. The covariate-adjusted RD framework imposes the same adjustment below and above the threshold point, which means that the covariate at the cut-off should have an equal conditional expectation limit from below and above at the threshold.

The first row of the table reports the RD treatment effect using the corresponding MSE optimal bandwidth h. The next three rows report 95% robust bias-corrected confidence intervals, the percentage length change of the covariate-adjusted confidence interval relative to the unadjusted confidence interval, and the p-value associated with the hypothesis that the RD treatment effect equals zero.

The previous covariate-adjusted and unadjusted RD inference are estimated twice, once when bandwidth h, for the RD estimator, and bandwidth b, for the bias-corrected estimator, are chosen separately, and then when they are chosen to be equal, $h = b \; or \; h/b \; = 1$. The point estimates of RD range from -0.702 to -0.860 and are statistically significant at 1% in all cases. The results show that including covariates in the RD estimator according to the CCT framework does not dramatically change the point estimate of retirement on the cognitive function score. That is, including covariates in the RD framework, which are truly predetermined, do not substantially affect the RD point estimate (Calonico et al., 2018). However, augmenting the standard RD estimator with covariates, which is restricted to be equivalent below and above the cut-off, can achieve substantial efficiency gains compared to the unadjusted RD estimator. The rows labeled CI length change show 6.2% and 11.5% efficiency gains when bandwidths are unrestricted and restricted, respectively, and optimally chosen using MSE with covariates. That is, including covariates in our RD estimator leads to inference improvements and precise point estimates for the impact of retirement on cognitive function.

Theoretically, there are two dominant theories of intelligence that were widely used by cognitive psychologists to conceptualize and measure intelligence: the Cattell-Horn models of fluid and crystallized intelligence, and the theory of General Cognitive Ability (Postlethwaite, 2011). According to the Cattell-Horn model of fluid intelligence and crystallized intelligence, different aspects of intelligence, which interact with each other to form the individual's overall intelligence, can be grouped into two domains. The first domain, "fluid intelligence", involves the ability of thinking, reasoning abstractly, and solving problems. This ability is independent of prior education and experience, which is closely related to biological and physical components (Mzzonna and Peracchi, 2012). The second domain, "crystallized intelligence", includes the ability to use the knowledge acquired during an individual's life from education and other life experience. Cognitive psychologists reported that fluid intelligence and crystallized intelligence have different age trajectories over the lifecycle. Fluid intelligence reaches a peak in early life (Anderson and Craik, 2000), and is subject to continuous decline as a person gets older (Anderson and Craik, 2000; Prull et al. 2000). In contrast, crystallized intelligence can be maintained at older ages, and is subject to a lower rate of decline after reaching its peak in middle age, or it can even improve with age, based on an individual's exposure to engaged cognitive stimulating activities (Herzog and Wallace, 1997; Hertzog et al., 2008).

Based on the theoretical framework of measuring cognitive function, the decline in cognitive function after retirement can be explained by the "disuse" or "reuse it or lose it" hypothesis, which states that individuals can reduce the rate of cognitive decline by engaging in cognitively demanding activities which keep their mind active "engaged lifestyle". Conversely, the disuse of cognitive activities may accelerate the process of cognitive decline (Roberts et al., 2011; Rohwedder and Willis, 2010). When retirees move from employment, which requires

regular engagement in cognitive-based activities, to a new lifestyle without cognitive activities this may cause mental retirement (Rohwedder and Willis, 2010) and hence their cognitive ability may gradually decline (Hultsch et al., 1999). However, the effect of retirement on cognitive function may be not immediate, because the retiree may experience a "period of honeymoon" in which they engage in activities that they did not do while they were working (Bonsange et al., 2012). These activities may still stimulate their cognitive abilities and delay their decline, especially that of crystallized intelligence (Salthouse, 2006), and thus mitigate the effect of retirement on cognitive function (George and Maddox, 1977). However, presumably the "honeymoon" will not last long and cognitive ability will start to decline, due to the decline in both crystallized and fluid intelligence.

In the previous RD analysis, I have implicitly assumed that the effect of retirement on cognitive functioning is constant across gender and level of education. Since the types of engaged cognitive-stimulated activities may vary by gender and education level, the change in lifestyle after retirement may have heterogeneous effects on cognitive abilities. To test this heterogeneity in the impact of retirement on the cognitive score, I run the RD estimation separately for four subgroups: low educated, high educated, males, and females.

It is well documented that there are gender differences in cognitive ability and females perform better in different cognitive tasks (Upadhayay & Guragain, 2014). Tables 1.4 and 1.5 report the RD estimation for females and males, respectively. The effect of being retired on the cognitive score is negative and significant for both males and females, separately. Retired females experienced a drop in cognitive ability by about 0.79 points, which corresponds to an approximately 7.1% decrease in cognitive score compared to the female subsample average score, 11.12, with 95% confidence interval [-1.09583, -0.49206]. The estimated drop in the

cognitive score for retired females is not sensitive to the choice of bandwidth for being retired at least one year. Moreover, the RD treatment effect of retirement on the cognitive score is consistent and stable across all RD point estimation procedures and their associated bandwidths.

On the other hand, retired males experience a higher drop in the cognitive score compared to retired females across all bandwidth choices. The drop in retired males' cognitive score is approximately 1.01 points, which corresponds to a 10.18% decrease in cognitive score compared to the sample average (9.92). One explanation that the RD treatment effect is higher for males than females is that work centrality is significantly lower among females than males (Mannhein, 1997; Sharabi and Harpuz, 2011), and family centrality is higher among females than males (Cousins and Tang, 2004). In sum, retirement has a larger and significant negative impact on the cognitive function of retired males compared to retired females, which means that retirement has a gender heterogeneity effect on the cognitive function of retired individuals.

Table 1.4: Fuzzy RD estimates for cognitive functioning (Total word recall)-Females only 65 age

	optimal bandwidth	bandwidths (h)				
	4.6	1	2	3	4	5
Conventional	-0.79395***	-0.92262**	-0.79233***	-0.8244***	-0.78901***	-0.80096***
	(0.15402)	(0.41039)	(0.258)	(0.20284)	(0.16933)	(0.14844)
$CI^{C}_{95\%}$	[-1.09583, -0.49206]					
Bias-corrected	-0.84551***	-1.3943***	-0.78419***	-0.83521***	-0.85251***	-0.79733***
	(0.15402)	(0.41039)	(0.258)	(0.20284)	(0.16933)	(0.14844)
$CI^{bc}_{95\%}$	[-1.14739, -0.543627]					
Robust	-0.84551***	-1.3943**	-0.78419**	-0.83521***	-0.85251***	-0.79733***
	(0.18492)	(0.64054)	(0.38499)	(0.29748)	(0.25038)	(0.22013)
$CI^{rbc}_{95\%}$	[-1.20795, -0.483067]					
L|R	37468	21889|15579	21889|15579	21889|15579	21889|15579	21889|15579
effective	37468	655 |1408	1566 |2657	2574|3957	3995|5284	5429|6738

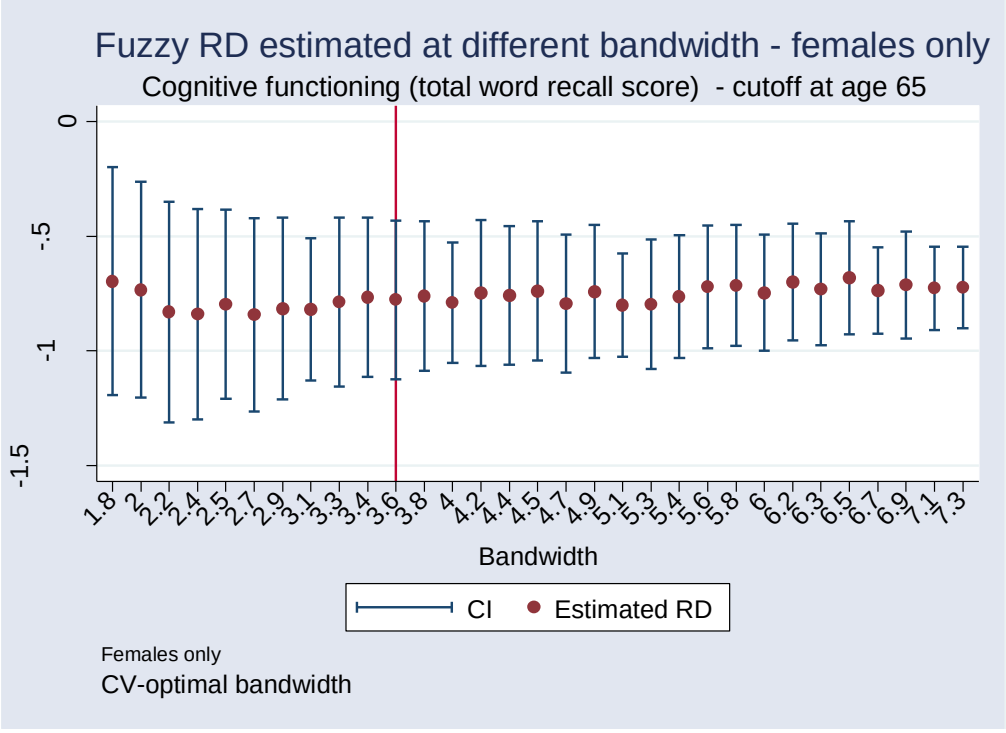
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

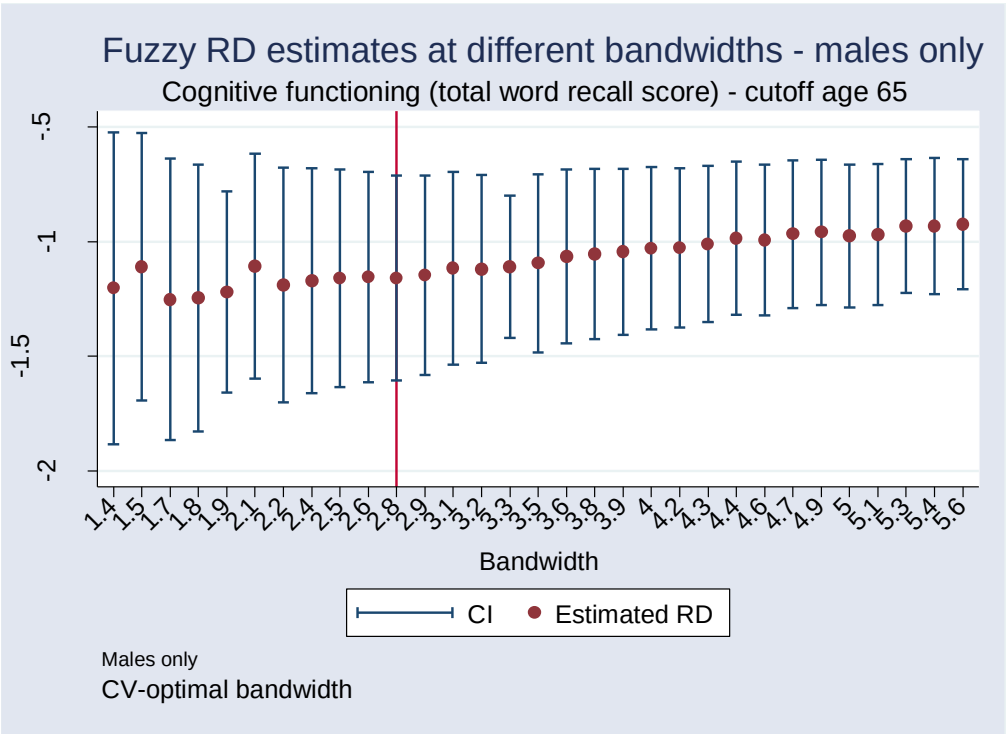
Table 1.5: Fuzzy RD estimates for cognitive functioning (Total word recall)-Males only at age 65

	optimal bandwidth	bandwidth				
	4.3	1	2	3	4	5
Conventional	-1.0101***	-1.2543***	-1.2118***	-1.136***	-1.0471***	-0.97553***
	(0.17425)	(0.42873)	(0.27888)	(0.21845)	(0.18233)	(0.15976)
$CI^{C}_{95\%}$	[-1.35165, -0.668619]					
Bias-corrected	-1.082***	-1.871***	-1.2794***	-1.2464***	-1.2348***	-1.1695***
	(0.17425)	(0.42873)	(0.27888)	(0.21845)	(0.18233)	(0.15976)
$CI^{bc}_{95\%}$	[-1.42348, -0.740449]					
Robust	-1.082***	-1.871***	-1.2794***	-1.2464***	-1.2348***	-1.1695***
	(0.20675)	(0.69149)	(0.40968)	(0.31984)	(0.2689)	(0.23616)
$CI^{rbc}_{95\%}$	[-1.48719, -0.676743]					
L|R	17038|12571					

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.9: RD estimates of Cognitive functioning for females and males separately





Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for cognitive functioning by gender. Each point is a separate non-parametric regression discontinuity point. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percentage points of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage point increments. Bootstrapped standard errors are based on 120 simulations to construct 95% CI.

To check whether the effect of retirement on cognitive functioning may vary by level of the education level of retirees, I run the RD estimation separately for a subsample of low-educated individuals (less than high school) and for a subsample of high-educated individuals.

Tables 1.6 and 1.7 report the RD treatment effect of retirement on cognitive functioning for the low- and high-educated samples, respectively. At first glance, the results show that there is a clear and dramatic education gradient, where retirement has a negative effect on the cognitive score for both subgroups, but the decline in the total word recall score for low-educated retirees is almost twice the decline in the cognitive score of high-educated retirees. The conventional RD estimator suggests that low-educated retirees experience a drop in their total word recall score by almost 1.7 points with a 95% confidence interval [-3.06743, -0.266735] compared with a 0.74 point drop in the cognitive score of high-educated retirees with 95% confidence interval [-1.13995, -0.357512]. This negative jump in total word recall score is equivalent to an approximately 18% and 6.4% decline in the cognitive functioning score of low- and high-educated retired individuals, respectively. Moreover, the effect of retirement on cognitive functioning is sensitive to the change in the selected bandwidth for all reported bandwidths. The results show that there is a dramatic negative jump in total word recall for low-educated retirees and this decline gets less as time since retirement increases, which means that longer retirement duration mitigates the cognitive functioning decline of low-educated retirees. High-educated retirees show a smaller jump in the cognitive score, which could be explained by higher acquisition rates of a cognitive reserve during their lifetime through engagement in -stimulating cognitive activities, which mitigates the effect of the change in lifestyle after retirement. Neuropsychologists define cognitive reserve as the individual's capacity to use brain networks to perform tasks efficiently (Coe and Zamarro, 2011; Stern, 2002). That is, although a

great deal of the accumulated empirical evidence reports that aging is accompanied by a systematic decline in the performance of different domains of cognitive tasks, this process of age-cognitive decline profile is not unavoidable. Certain life experiences, such as educational and occupational attainment, can increase an individual's level of reserve, which manifest as a set of skills that increase the capacity of individuals to keep learning and adapt to age-related challenges (Baltes and Baltes, 1990), hence slowing cognitive decline associated with normal aging or Alzheimer's disease (Adam et al., 2007). Some studies argue that individuals with a high level of education have lower risks of developing dementia relative to individuals with a low level of education (Letenneur et al., 1999; Stern et al., 1994). In addition, functional imaging studies argue that individuals' common response in normal situations, when facing tasks with increasing difficulty, is the recruitment of an additional brain area to work on that task (Gur et al., 1988; Grady et al., 1996; Grasby et al., 1994). Stern (2002) argues that if we have two individuals with the same level of brain reserve but one of them has a more cognitive reserve, this individual has the ability to tolerate or cope with a larger negative effect on the brain before suffering from any functional impairment. Based on that, when individuals face the same task difficulty, high-educated retirees, who are presumably more highly skilled, experience lower task-related recruitment than low-educated retirees. That is, retired persons with a high cognitive reserve are able to cope with the normal decline in cognitive ability after retirement and still maintain effective functioning. Another explanation is that low-educated individuals, who usually work continuously in uninspiring jobs, have experienced a wearing effect in cognitive abilities, which negatively affects their cognitive reserve, i.e. when individuals do the same routines in their jobs or do less cognitive-stimulated activities for several years.

Table 1.6: Fuzzy RD estimates - Cognitive functioning (Total word recall) Low educated at age 65

	Optimal bandwidth	Bandwidth				
	2.6	1	2	3	4	5
Conventional	-1.6671**	-3.2211**	-1.7525**	-1.6461**	-1.451***	-1.0737**
	(0.71448)	(1.461)	(0.88911)	(0.65724)	(0.53937)	(0.4731)
$CI^{c}_{95\%}$	[-3.06743, -0.266735]					
Bias-corrected	-1.9322***	-3.4752**	-2.1898**	-1.8758***	-1.9355***	-1.9872***
	(0.71448)	(1.4610)	(0.88911)	(0.65724)	(0.53937)	(0.4731)
$CI^{bc}_{95\%}$	[-3.33252, -0.531826]					
Robust	-1.9322**	-3.4752*	-2.1898*	-1.8758*	-1.9355**	-1.9872***
	(0.80332)	(2.0288)	(1.2748)	(1.0026)	(0.81971)	(0.70522)
$CI^{rbc}_{95\%}$	[-3.50665, -0.357696]					
Obs	4135	4135	4135	4135	4135	4135

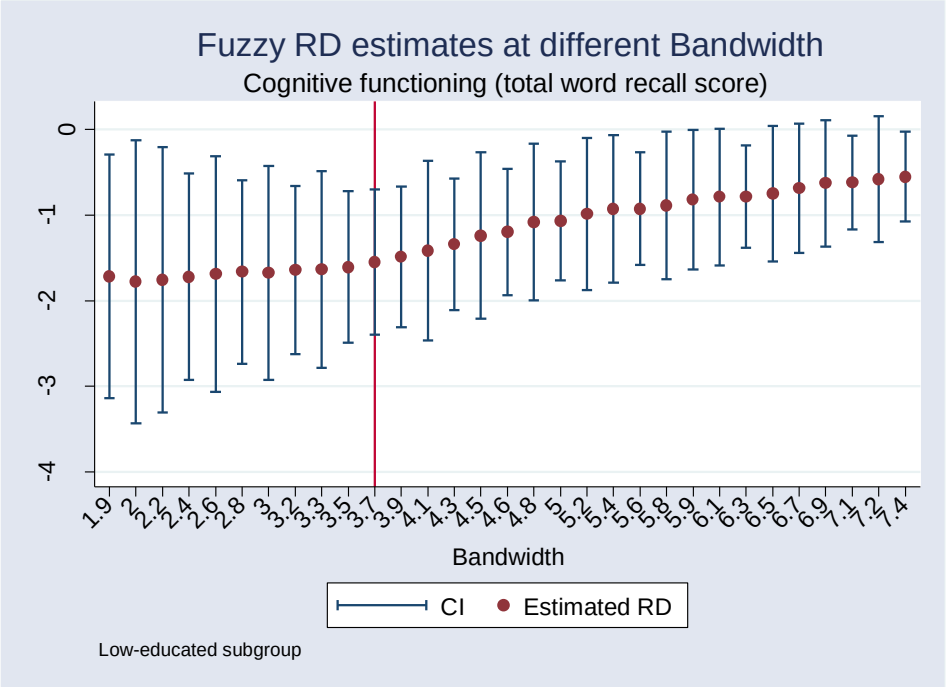
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

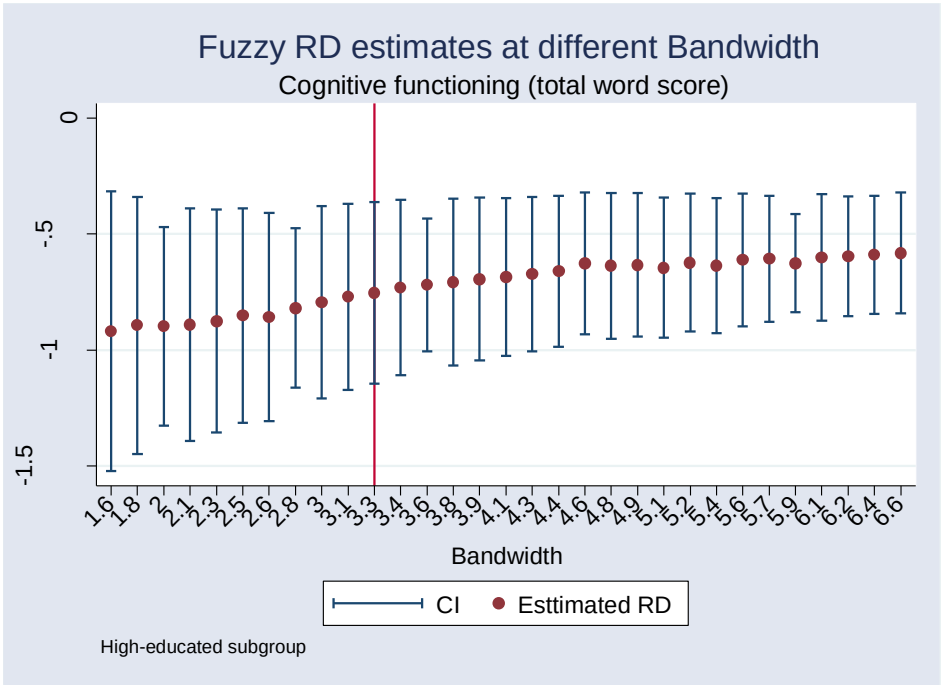
Table 1.7: Fuzzy RD estimates - Cognitive functioning (Total word recall) high educated at age 65

	Optimal bandwidth	Bandwidth				
	3.2	1	2	3	4	5
Conventional	-0.74873***	-1.0079**	-0.89764***	-0.78655***	-0.69182***	-0.64618***
	(0.19961)	(0.41278)	(0.26721)	0.21098	0.17692	0.15589
$CI^{c}_{95\%}$	[-1.13995, -0.357512]					
Bias-corrected	-0.81332***	-1.0465**	-0.89933***	-0.95907***	-0.89298***	-0.80091***
	(0.19961)	(0.41278)	(0.26721)	(0.21098)	(0.17692)	(0.15589)
$CI^{bc}_{95\%}$	[-1.20454, -0.422098]					
Robust	-0.81332***	-1.0465*	-0.89933**	-0.95907***	-0.89298***	-0.80091***
	(0.23674)	(0.63324)	(0.38869)	(0.30422)	(0.25787)	(0.2276)
$CI^{rbc}_{95\%}$	[-1.27731, -0.349323]					
Obs	31787	31787	31787	31787	31787	31787

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.10: Non-parametric Regression discontinuity robustness of bandwidth choice for cognitive functioning by educational level





Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for cognitive functioning by educational level. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points increment. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

Figure 1.10 shows the effect of retirement on cognitive functions at different bandwidths, and that the negative jump in the total word recall score gets smaller as the duration of retirement increases. Moreover, the negative jump in cognitive score will converge to stable values for both less-educated and high-educated retirees. However, the improvement in cognitive abilities for low-educated retirees will be at a remarkable rate after 6 years of retirement. The improvement in low-educated retiree's mood or mental health after a long period of change in lifestyle may mitigate the dramatic drop in their cognitive abilities (Grip et al., 2015) and drive down the RD treatment effect of retirement due to including more retirees in the neighborhood of the cut-off point.

Table 1.8 shows the main results from the non-parametrical RD estimation of the impact of retirement on mental health as measured by the CESD depression scale. Each column reports the local polynomial point estimates using the corresponding bandwidth.

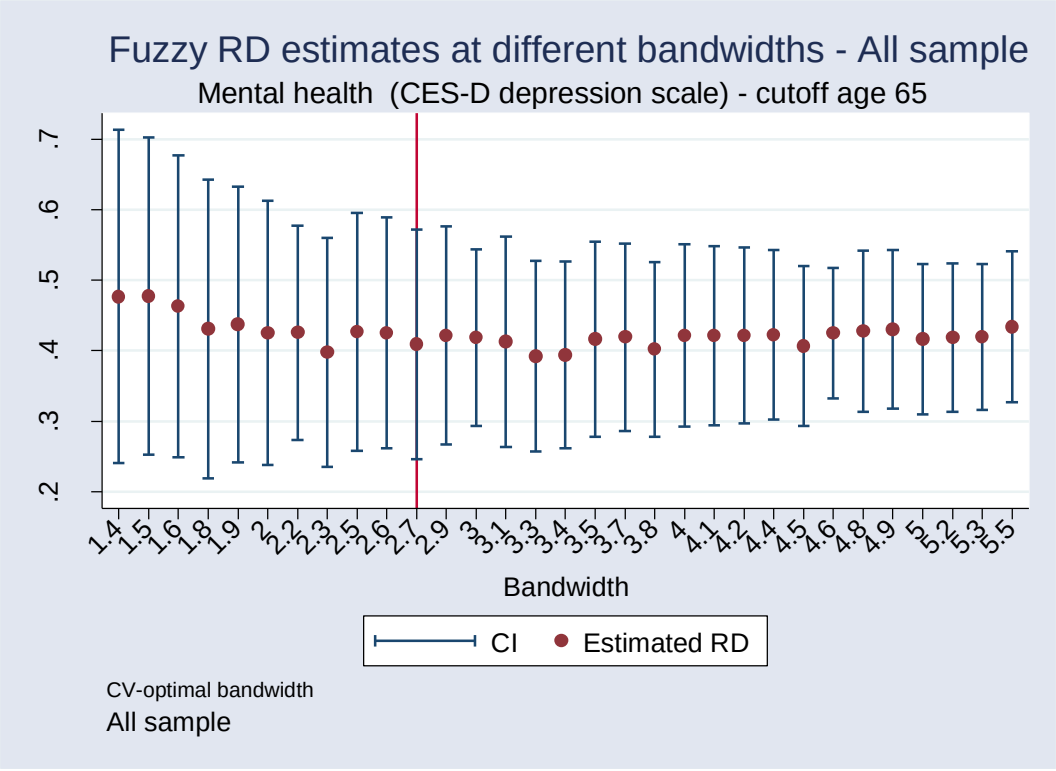
The first column shows that the jump in retirement probability was associated with a positive jump in the CESD depression scale. That is, the impact of retirement on mental health is negative and highly significant. The RD estimation suggests that retired individuals experienced a dramatic increase in the CESD scale by approximately 0.42 points with a 95% confidence interval [0.294554, 0.550295]. The positive jump in the CESD depression scale is significantly high relative to the sample average. In addition, the RD estimation is not sensitive to changes in the bandwidth for all bandwidths. The RD effect of retirement on mental health is stable across the three RD inference methods with a small difference in standard error in the robust-bias corrected approach. In contrast to the cognitive function score, transition into retirement status has an immediate and highly significant impact on mental health without a time lag.

Table 1.8: Fuzzy RD estimates - Mental Health (CES-D Depression scale) all sample at age 65

	Optimal Bandwidth	Bandwidth				
	3.9	1	2	3	4	5
Conventional	0.42242***	0.52908***	0.42687***	0.41852***	0.42215***	0.43155***
	(0.06524)	(0.14331)	(0.09692)	(0.07672)	0.06452	0.05677
$CI^{c}_{95\%}$	[0.294554, 0.550295]					
Bias-corrected	0.4142***	0.97218***	0.54744***	0.45831***	0.42194***	0.40951***
	(0.06524)	(0.14331)	(0.09692)	(0.07672)	(0.06452)	(0.05677)
$CI^{bc}_{95\%}$	[0.286333, 0.542073]					
Robust	0.4142***	0.97218***	0.54744***	0.45831***	0.42194***	0.40951***
	(0.07797)	(0.21599)	(0.13627)	(0.1098)	(0.0938)	(0.08284)
$CI^{rbc}_{95\%}$	[0.26139, 0.567016]					
L|R	43266|31261	43266|31261	43266|31261	43266 |31261	43266|31261	43266|31261
Obs	74527	74527	74527	74527	74527	74527

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.11: Robustness of RD estimates of retirement effect on mental health at different bandwidths



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for mental health. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points increment. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

Table 1.9: Fuzzy RD estimates - Mental Health (CES-D Depression scale) Females only at age 65

	Optimal Bandwidth	Bandwidth				
	3.1	1	2	3	4	5
Conventional	0.39627***	0.49977**	0.39081***	0.39739***	0.41704***	0.43872***
	(0.10784)	(0.2073)	(0.14045)	(0.11185)	(0.09416)	(0.08285)
$CI^{c}_{95\%}$	[0.184901, 0.607639]					
Bias-corrected	0.3712***	0.9208***	0.53028***	0.40431***	0.37436***	0.37743***
	(0.10784)	(0.2073)	(0.14045)	(0.11185)	(0.09416)	(0.08285)
$CI^{bc}_{95\%}$	[0.159827, 0.582565]					
Robust	0.3712***	0.9208***	0.53028***	0.40431**	0.37436***	0.37743***
	(0.12647)	(0.30843)	(0.19653)	(0.15904)	(0.13632)	(0.12077)
$CI^{rbc}_{95\%}$	[0.123317, 0.619074]					
L|R	24328|17488	24328|17488	24328|17488	24328|17488	24328|17488	24328|17488
Obs	41816	41816	41816	41816	41816	41816

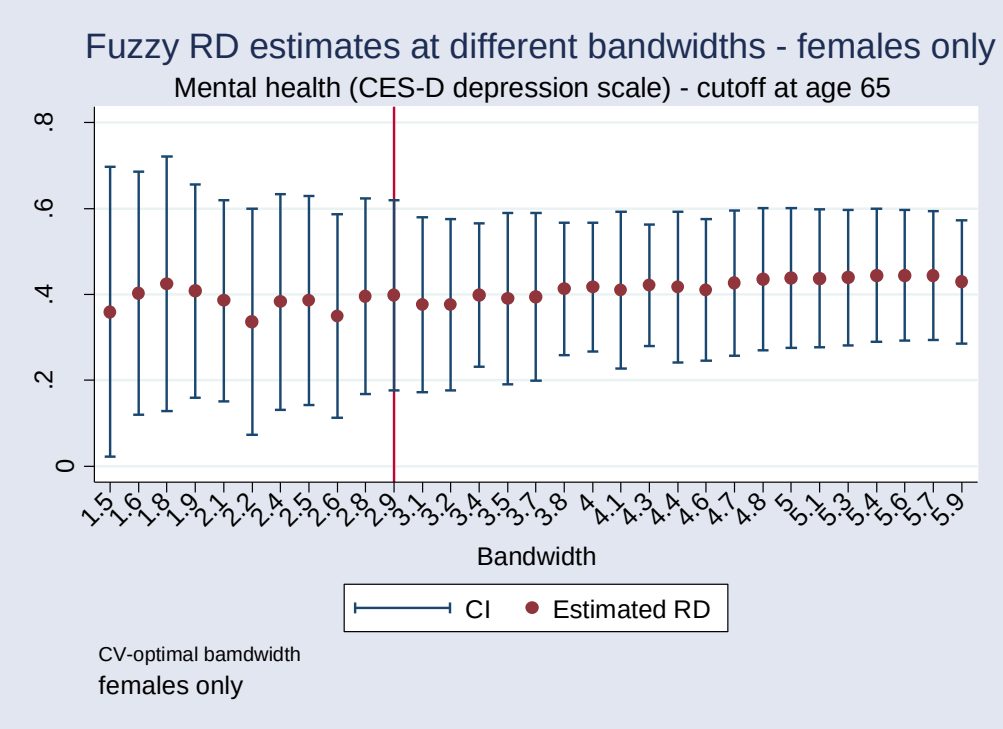
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

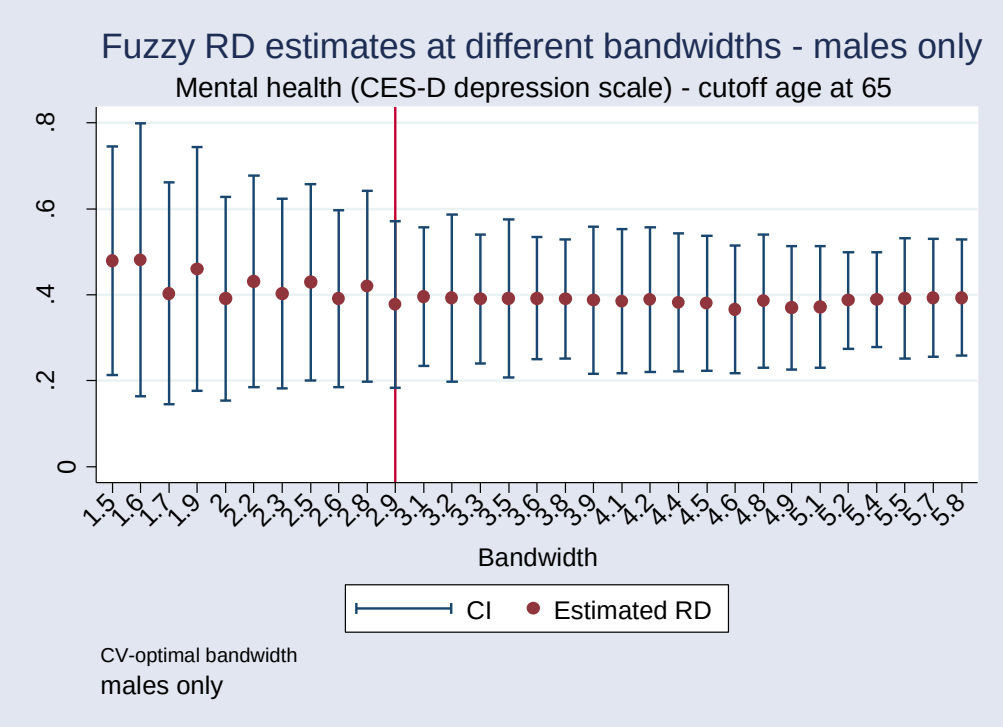
Table 1.10: Fuzzy RD estimates - Mental health (CES-D Depression scale) Males only at 65

	Optimal bandwidth	Bandwidth				
	4.06	1	2	3	4	5
Conventional	0.38498***	0.5365***	0.42347***	0.39686***	0.38573***	0.38509***
	(0.08491)	(0.19751)	(0.13161)	(0.10274)	(0.08581)	(0.07522)
$CI^{c}_{95\%}$	[0.218565, 0.551399]					
Bias-corrected	0.3799***	0.97749***	0.53709***	0.47665***	0.42908***	0.39809***
	(0.08491)	(0.19751)	(0.13161)	(0.10274)	(0.08581)	(0.07522)
$CI^{bc}_{95\%}$	[0.213482, 0.546316]					
Robust	0.3799***	0.97749***	0.53709***	0.47665***	0.42908***	0.39809***
	(0.10304)	(0.29729)	(0.18787)	(0.14965)	(0.12672)	(0.11116)
$CI^{rbc}_{95\%}$	[0.177954, 0.581844]					
Observation	32711	32711	32711	32711	32711	32711
L|R	18938|13773	18938|13773	18938|13773	18938|13773	18938|13773	18938|13773

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.12: Robustness of RD estimates of retirement effect on mental health at different bandwidths for females and males separately





Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for mental health. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points increment. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

The RD estimator shows that there is a significant increase in the CESD depression scale of retirees who have spent around one year in retirement and this effect is mitigated by the increase in retirement duration. The positive jump in the CESD depression scale indicates that when individuals move from the labor force to retirement, they may be more likely to experience one or more of the following mental health issues: depression, feeling lonely, feeling sad, seeing everything as an effort, or suffering from sleeping issues.

To investigate whether the effect of retirement on mental health status is heterogeneous across individuals, RD estimation is run separately for different subsamples (gender, education, and income). Tables 1.9 and 1.10 show the RD estimates for males and females, respectively. The results show that there is no significant difference in the effect of retirement on mental health between males and females. Retirement has a negative effect on the mental health status of both females and males. Figure 1.12 indicates that there is a positive jump in the value of the CESD depression scale by about 0.4 for both males and females and this estimated RD jump in the CESD depression scale is stable for different bandwidths for both males and females.

The highly significant increase in the CESD depression scale for males and females by around 0.4 points is equivalent to 27 percentage points for females and 35 percentage points for males with a 95% confidence interval [0.123317, 0.619074] and [0.177954, 0.581844] for females and males, respectively. That is, retirement has a larger negative effect for males than females, i.e., retired males are more likely to experience more depressive symptoms relative to retired females. The low relative negative effect of retirement on female's mental health compared to males may be explained by the high and large prevalence of depression symptoms among females during their life (Nolan-Hoeksema et al., 1999), which turn the absolute equivalent jump in CESD into lower relative effects among females.

Table 1.11: Fuzzy RD estimates - Mental health (CES-D Depression scale) Low educated at age 65

	Optimal bandwidth	Bandwidth				
	4.4	1	2	3	4	5
Conventional	0.52164***	0.58505***	0.52603***	0.5193***	0.52052***	0.53098***
	(0.08745)	(0.21163)	(0.14178)	(0.11163)	(0.0935)	(0.08195)
$CI^{c}_{95\%}$	[0.350237, 0.693041]					
Bias corrected	0.52413***	1.0089***	0.64628***	0.56237***	0.5256***	0.50886***
	(0.08745)	(0.21163)	(0.14178)	(0.11163)	(0.0935)	(0.08195)
$CI^{bc}_{95\%}$	[0.352728, 0.695533]					
Robust	0.52413***	1.0089***	0.64628***	0.56237***	0.5256***	0.50886***
	(0.10651)	(0.33076)	(0.20224)	(0.16222)	(0.13776)	(0.1212)
$CI^{rbc}_{95\%}$	[0.315366, 0.732895]					
L|R	19609|19153	19609|19153	19609|19153	19609|19153	19609|19153	19609|19153
Obs	38762	38762	38762	38762	38762	38762

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.12: Fuzzy RD estimates - Mental health (CES-D Depression scale) High educated at age 65

	Optimal bandwidth	Bandwidth				
	5.1	1	2	3	4	5
Conventional	0.20229***	0.31867*	0.19431	0.1879*	0.19653**	0.2007***
	(0.07586)	(0.19314)	(0.13101)	(0.10387)	(0.0876)	(0.07725)
$CI^{c}_{95\%}$	[0.053603, 0.350986]					
Bias-corrected	0.18442**	0.71394***	0.29055*	0.21673**	0.18433**	0.18491**
	(0.07586)	(0.19314)	(0.13101)	(0.10387)	(0.0876)	(0.07725)
$CI^{bc}_{95\%}$	[0.035725, 0.333108]					
Robust	0.18442**	0.71394*	0.29055	0.21673	0.18433	0.18491*
	(0.09216)	(0.28403)	(0.18229)	(0.14685)	(0.12596)	(0.11147)
$CI^{rbc}_{95\%}$	[0.003795,0.365037]					
L|R	23657|12108	23657|12108	23657|12108	23657|12108	23657|12108	23657|12108
Obs	35765	35765	35765	35765	35765	35765

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Since educational attainment could affect an individual's skills and their occupation, retirement may have a heterogeneous effect among different educational groups. Tables 1.11 and 1.12 report the RD estimate for low- and high-educated subgroups to investigate if there is heterogeneity. The RD estimates show that there is a noticeable difference in the effect of retirement on low-and high-educated individuals. The positive increase in the CESD depression scale according to the conventional and robust bias-corrected RD estimators are 0.2 and 0.5 points for high and low-educated retirees with 95% confidence intervals [0.003795,0.365037] and [0.315366, 0.732895], respectively. The increase in the CESD depression scale is equivalent to an 18 and 31 percentage point decline in the mental health of high and low-educated retirees relative to non-retirees, respectively. Some studies found that work for many men is a way to manage the depression where the job serves as a defense against depressive anxiety.

Table 1.13 shows the main results of the standard and covariate-adjusted RD estimates, which include covariates. The first row of the table reports the RD treatment effect using the corresponding MSE optimal bandwidth h. The next three rows report the 95% robust bias-corrected confidence intervals, the percentage length change of the covariate-adjusted confidence interval relative to the unadjusted confidence interval, and the P-value associated with the hypothesis that the RD treatment effect equals zero.

Table 1.13: Covariate-adjusted RD estimates – Mental health (CES-D depression scale) at age 65

	Not using covariates	Using covariates
	Standard Fuzzy RD	Covariate-Adjusted Fuzzy RD
Inference with h/p unrestricted		
RD treatment effect	0.4142***	0.3174***
Robust 95% CI	[.26139, .567016]	[.18165, .453147]
CI length change (%)		-11.16
robust p-value	0.000	0.000
rho (h/p)	0.660	0.667
Inference with h/p restricted		
RD treatment effect	0.43498***	0.37921***
Robust 95% CI	[.237609, .632346]	[.188748, .569678]
CI length change (%)		-3.5
robust p-value	0.000	0.000
$h\|p$	3.5|3.5	3.5|3.5
n-|n+	43266|31261	43266|31261
obs	8374|10411	10387|12214
effective obs	74527	74527

Notes: (i) All estimates are computed using a triangular Kernel and nearest neighbor heteroskedasticity-robust variance estimator as suggested by CCT framework. (ii) bandwidth used for h *and* b are data-driven MSE-optimal for either standard RD estimator or covariate-adjusted RD estimator. *,**,*** indicate significance level at 10%, 5%, 1% respectively.

The covariates-adjusted and unadjusted RD inference is estimated twice, once when bandwidth h, for the RD estimator, and b, for the bias-corrected estimator, are chosen separately, and then when they are chosen to be equal, i.e. $h = b$ *or* $(h/b = 1)$. The point estimates of RD range from 0.3 to 0.43 and it is statistically significant at the 1% level in all cases. The results show that including covariates in the RD estimator according to the CCT framework does not dramatically change the point estimate of retirement on the CESD depression scale. That is, including covariates in the RD framework, which are truly predetermined, do not substantially affect the RD point estimate (Calonico et al., 2018). However, augmenting the standard RD estimator with covariates, which is restricted to be equivalent below and above the cut-off, can achieve substantial efficiency gains compared to the unadjusted RD estimator. The rows labeled

CI length change show around 11.1% and 3.5% efficiency gains when bandwidths are unrestricted and restricted, respectively, and optimally chosen using MSE with covariates. That is, including covariates in our RD estimator leads to inference improvements and precise point estimates for the impact of retirement on mental health.

I checked different polynomial orders for the running variable and I found that there is no improvement in the MSE after the second-order polynomial. Therefore, I reported the results of the RD under the linear and quadratic polynomial smoothing function. The last three columns report the RD estimates using a local quadratic estimator ($p = 2$) with different kernel functions. Overall, the RD treatment effect is stable across the three different kernel functions. Retirement has a significant negative effect on retirees' mental health as measured by the CESD depression scale, which indicates a positive increase in the depression scale in the range from 0.38 to 0.42 points. That is, the increase in retirees' depression score maybe around 29% and 32% percentage points according to the nonparametric RD local quadratic estimator under the conventional and the robust bias-corrected variance. Moreover, the value of the RD estimate, using the quadratic estimate, is similar to the linear estimate, which means that the speed of change in the depression scale when individual age is almost stable. The results in table 1.14 indicate that using different weighting schemes for the observations away from the cut-off point have no effect on the RD estimates according to the conventional estimator and the robust bias-corrected estimator. The only difference is related to the optimal bandwidth length.

Table 1.14: RD estimates for different kernel and different polynomial orders - Mental health (CESD depression scale) at age 65

RD estimator	1	2	3	4	5
Conventional	0.41988***	0.42329***	0.42165***	0.40025***	0.41134***
Std. Err.	(0.07113)	(0.05987)	(0.10013)	(0.08032)	(0.09708)
Robust Bias corrected	0.40468***	0.41787***	0.43716***	0.38673***	0.42069***
Std. Err.	(0.08259)	(0.07222)	(0.11356)	(0.08949)	(0.1105)
Robust 95% CI	[0.242813, 0.566548]	[0.276335, 0.559413]	[0.214592, 0.659725]	[0.211321, 0.562132]	[.204122, .637264]
Kernel Type	Uniform	Epanechnikov	Triangular	Uniform	Epanechnikov
Order Loc policy (P)	1	1	2	2	2
BW local poly (h)	2.8	4.2	3.9	4.9	3.9
BW Type	mserd	mserd	mserd	mserd	mserd
observations	74527	74527	74527	74527	74527
L|R	43266|31261	43266|31261	43266|31261	43266|31261	43266|31261
effective no	5062|7360	9366|11359	8374|10411	11446|13209	8374|10411

Notes (i) first column report two different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{fuzzy}$ and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{fuzzy}$. (ii) standard errors are in parentheses. (iii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.15 and Figure 1.13 show the effects of retirement on self-reported health. Self-reported health has consistently been found to be an accurate predictor of future health outcomes and utilization (Idler and Benyamini 1997). The first column in the table indicates retired individuals are about 13 percentage points less likely to report that they are in excellent, very good or good health than non-retired individuals. The results also indicate that the drop in the proportion of retirees who report that they are in good health is not affected by the choice of bandwidth and are consistent across the different RD procedures. However, tables 1.16 and 1.17 indicate that there is gender heterogeneity in the effect of retirement on general health status. Table 1.16 and figure 1.14 indicate that retirement is associated with a negative drop in the proportion of retired females who report that they are in excellent, very good, and good health. However, retirement has a larger negative effect on retired males' self-reported health compared with females and the full sample. That is, retirement is associated with poorer general health among retired male than retired females. The negative drop in the proportion of retired males who report excellent, very good, or good health is about 16 to 24 percentage points against 8 to 12 percentage points for retired females. More specifically, the drop in the proportion of males who report that they are in excellent, very good or good health is almost double the drop in the proportion of females. In other words, females' self-reported general health is less affected by transition into retirement than that of males. In addition, the RD effect of retirement on retired males' general health status is more sensitive to the bandwidth choice relative to retired females, where the discontinuity in health status for compliers at the cut-off point ranges between 16 and 24 for retired males, and between 8 and 12 percentage points for retired females with 95% significant confidence intervals. Also, the three RD inference methods result in the same treatment effect with lower magnitude for the conventional method and all of them are evaluated based on the optimal

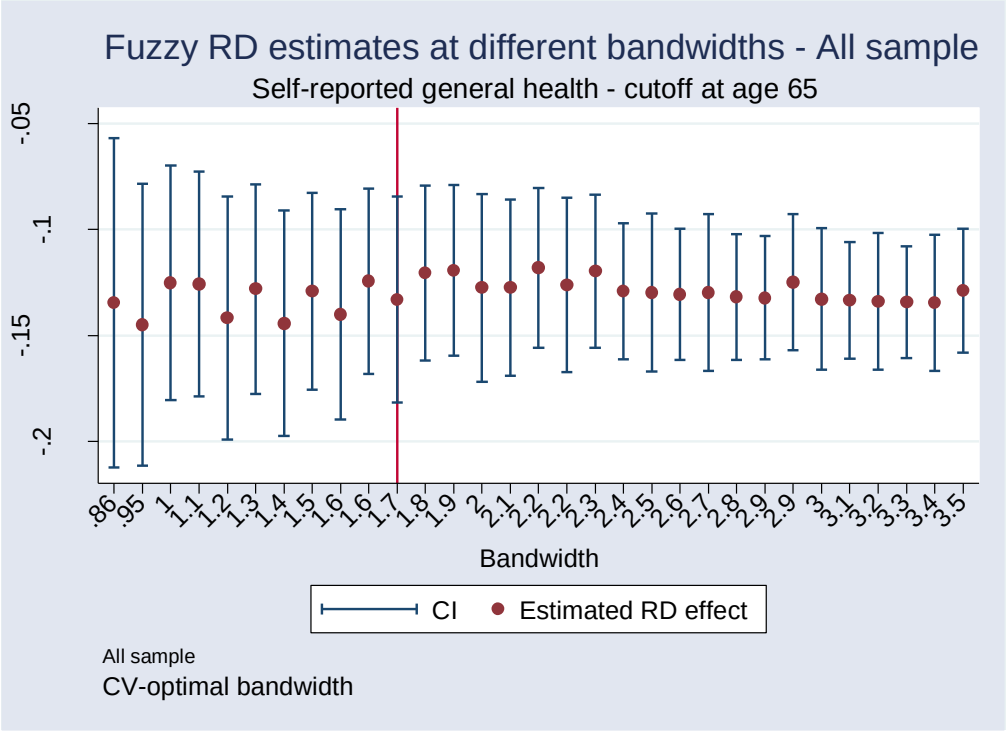
bandwidth. The robustness RD estimates at different bandwidths also indicate that the drop in general health for females is around 10 percentage points while the drop in general health for males begins with an almost 20 percentage point drop in general health and converges to almost 16 percentage points at larger bandwidths. That is, the short-run effect of retirement on general health directly after retirement has a higher magnitude. Comparing to non-retired individuals, a smaller percent of retired individuals reported that they are in excellent, very good, or good health. Since the self-reported general health measures the overall physical and social well-being of the respondent, then retirement, in general, has a negative effect on retired respondents' overall general health status. To check if there is socioeconomic heterogeneity in the effect of retirement on general health, I run the RD estimators for the low-educated and high-educated subgroups separately. Tables 1.18 and 1.19 show the RD estimates for the low- and high-educated subgroups. The results indicate that retirement has a larger negative impact on the self-reported health of low-educated retirees than high-educated retirees. The results in table 1.18 suggest that the low-educated retired individuals are 14 percentage points less likely to report to be in good health relative to low-educated non-retirees. The RD results also are consistent across different bandwidth selections and are highly significant. On the other hand, the RD results in table 1.19 suggest that the high-educated retirees are almost 8 percentage points less likely to report being in good health relative to high-educated non-retirees. These results are expected because education is documented to increase the efficiency of producing health so that highly educated individuals have better health outcomes (Grossman and Kaestner, 1997).

Table 1.15: Fuzzy RD estimates of Self-reported health - All sample at cutoff age 65

	optimal	bandwidth				
	3.2	1	2	3	4	5
Conventional	-0.1346***	-0.14212***	-0.12797***	-0.13259***	-0.13882***	-0.1395***
	(0.01613)	(0.03304)	(0.02173)	(0.01716)	(0.0144)	(0.01266)
$CI^{c}_{95\%}$	[-0.16621, -0.103]					
Bias-corrected	-0.13172	-0.1871	-0.15295	-0.12981	-0.12765	-0.13337
	(0.01613)	(0.03304)	(0.02173)	(0.01716)	(0.0144)	(0.01266)
$CI^{bc}_{95\%}$	[-0.163332, -0.10012]					
Robust	-0.13172	-0.1871	-0.15295	-0.12981	-0.12765	-0.13337
	(0.01891)	(0.05003)	(0.03127)	(0.02473)	(0.02102)	(0.01854)
$CI^{rbc}_{95\%}$	[-0.168785, -0.09466]					
L|R	43266|31261	43266|31261	43266|31261	43266|31261	43266|31261	43266|31261
effective	6395|8707	1454|2722	3336|5212	5435|7820	8374|10411	11446|13209
obs	74527	74527	74527	74527	74527	74527

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.13: Robustness of RD estimates of retirement effect on self-reported health



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for self-reported health. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points increment. Bootstrapped standard errors are based on 120 simulations to construct 95% CI.

Table 1.16: Fuzzy RD estimates of Self-reported Health Status - Females only at Cutoff Age 65

	Optimal bandwidth	Bandwidth				
	2.8	1	2	3	4	5
Conventional	-0.10493***	-0.08127*	-0.08749***	-0.10838***	-0.11773***	-0.12047***
	(0.02481)	(0.04779)	(0.03056)	(0.02385)	(0.01982)	(0.01735)
$CI^{c}_{95\%}$	[-0.153567, 0.0563]					
Bias-corrected	-0.09854***	-0.13322***	-0.08704***	-0.07685***	-0.09389***	-0.10585***
	(0.02481)	(0.04779)	(0.03056)	(0.02385)	(0.01982)	(0.01735)
$CI^{bc}_{95\%}$	[-0.147175, 0.04991]					
Robust	-0.09854***	-0.13322*	-0.08704**	-0.07685**	-0.09389***	-0.10585***
	(0.02889)	(0.07174)	(0.04479)	(0.03505)	(0.02953)	(0.02586)
$CI^{rbc}_{95\%}$	[-0.155173, -0.04191]					
L|R	24328|17488	24328|17488	24328|17488	24328|17488	24328|17488	24328|17488
Obs	41816	41816	41816	41816	41816	41816

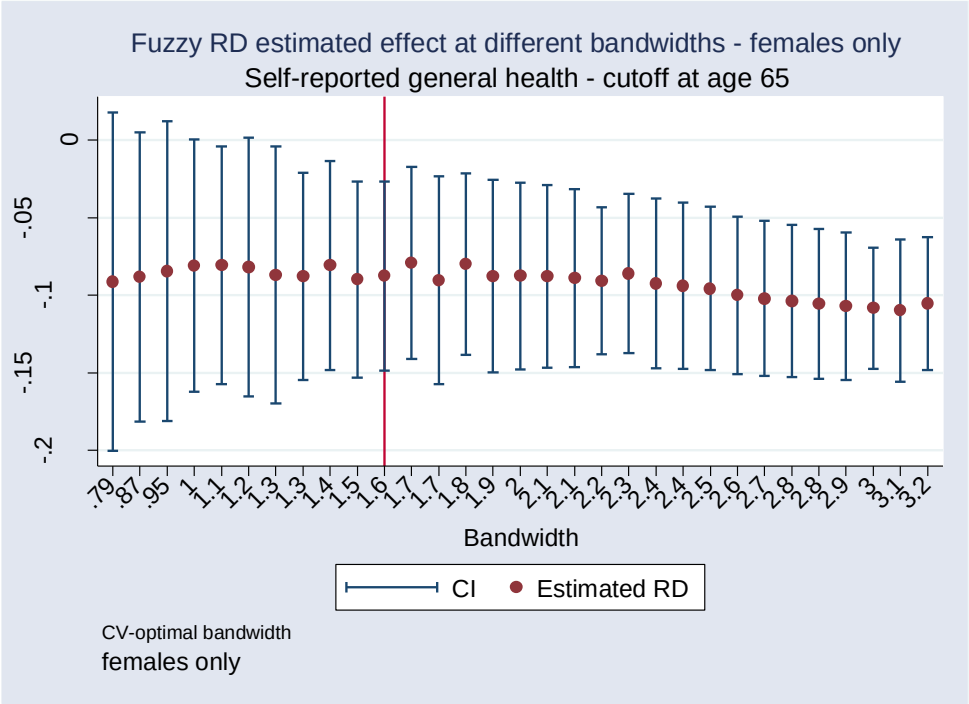
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

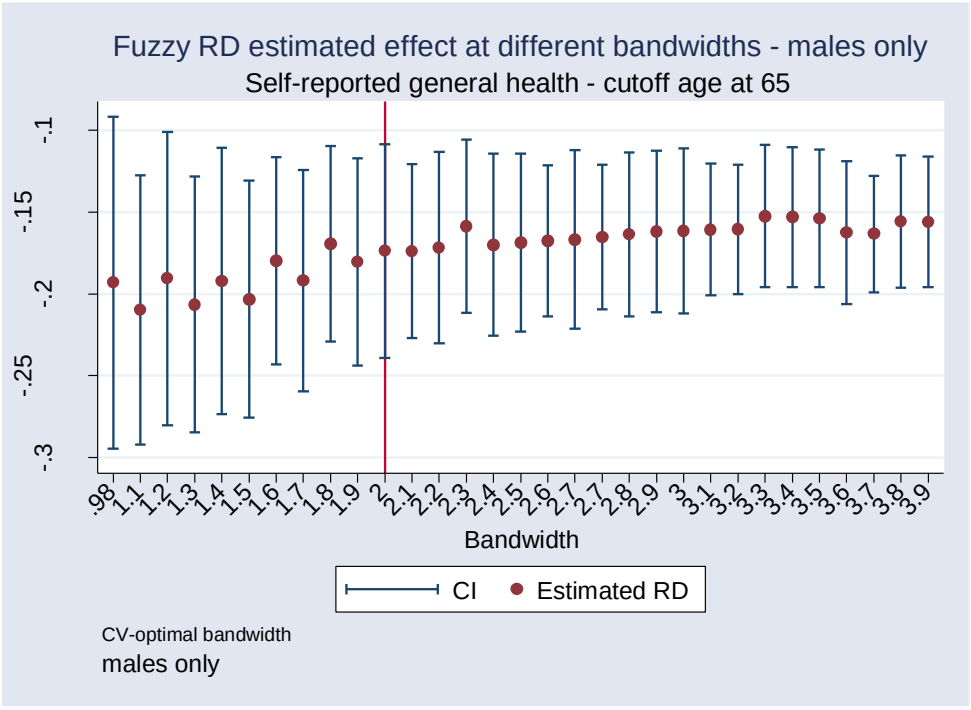
Table 1.17: Fuzzy RD estimates – Self-reported Health Status Males only at Cutoff Age 65

	Optimal bandwidth	Bandwidth				
	3.9	1	2	3	4	5
Conventional	-0.16357***	-0.20825***	-0.1752***	-0.16127***	-0.16352***	-0.16154***
	(0.0212)	(0.04596)	(0.03113)	(0.02491)	(0.02112)	(0.01863)
$CI^{c}_{95\%}$	[-0.205132, -0.12201]					
Bias-corrected	-0.16237***	-0.24354***	-0.22642***	-0.19086***	-0.16731***	-0.16587***
	(0.0212)	(0.04596)	(0.03113)	(0.02491)	(0.02112)	(0.01863)
$CI^{bc}_{95\%}$	[-0.203926, -0.12081]					
Robust	-0.16237***	-0.24354***	-0.22642***	-0.19086***	-0.16731***	-0.16587***
	(0.02536)	(0.06926)	(0.04383)	(0.03507)	(0.03011)	(0.0268)
$CI^{rbc}_{95\%}$	[-0.212075, -0.11266]					
L|R	18938|13773	18938|13773	18938|13773	18938|13773	18938|13773	18938|13773
Obs	32711	32711	32711	32711	32711	32711

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.14: Robustness of RD estimates of retirement effect on Self-reported health at different bandwidths from females and males, separately.





Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for self-reported health. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percentage point of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage point increment. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

Table 1.18: Fuzzy RD estimates of Self-reported general health status for Low- educated at cutoff age 65

	Optimal bandwidth	Bandwidth				
	3.5	1	2	3	4	5
Conventional	-0.14699***	-0.16026***	-0.14982***	-0.14617***	-0.14782***	-0.14604***
	(0.02377)	(0.05164)	(0.03353)	(0.0263)	(0.02198)	(0.01922)
$CI^{c}_{95\%}$	[-0.193573, -0.14004]					
Bias-corrected	-0.1455***	-0.2534***	-0.17393***	-0.15623***	-0.14774***	-0.14913***
	(0.02377)	(0.05164)	(0.03353)	(0.0263)	(0.02198)	(0.01922)
$CI^{bc}_{95\%}$	[-0.192087, -0.09892]					
Robust	-0.1455***	-0.2534***	-0.17393***	-0.15623***	-0.14774***	-0.14913***
	(0.02815)	(0.0797)	(0.04906)	(0.03855)	(0.03252)	(0.02854)
$CI^{rbc}_{95\%}$	[-0.200684, -0.09032]					
L|R	19609|19153	19609|19153	19609|19153	19609|19153	19609|19153	19609|19153
Obs	38762	38762	38762	38762	38762	38762

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.19: Fuzzy RD estimates of Self-reported Health - High-educated at Cutoff Age 65

	Optimal bandwidth	Bandwidth				
	3.9	1	2	3	4	5
Conventional	-0.09095***	-0.07738*	-0.06476**	-0.0778***	-0.09118***	-0.09428***
	(0.01795)	(0.04014)	(0.02661)	(0.02118)	(0.01787)	(0.0158)
$CI^{c}_{95\%}$	[-0.126141, -0.05577]					
Bias-corrected	-0.08521***	-0.05374	-0.08453***	-0.06095***	-0.06368***	-0.07729***
	(0.01795)	(0.04014)	(0.02661)	(0.02118)	(0.01787)	(0.0158)
$CI^{bc}_{95\%}$	[-0.120396, -0.05002]					
Robust	-0.08521***	-0.05374	-0.08453**	-0.06095**	-0.06368**	-0.07729***
	(0.02108)	(0.06098)	(0.03768)	(0.02985)	(0.02564)	(0.02275)
$CI^{rbc}_{95\%}$	[-0.126518, -0.0439]					
L|R	23657|12108	23657|12108	23657|12108	23657|12108	23657|12108	23657|12108
Obs	35765	35765	35765	35765	35765	35765

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.20 reports the RD estimates when the covariates are included in the self-reported general health model. The covariate-adjusted RD estimator indicates a drop in the proportion of retired individuals who report that they are in excellent, very good and good health by 10 percentage points, against 13 percentage points for the standard fuzzy RD estimator using the conventional RD setting. Similarly, the covariate-adjusted RD estimate indicates a drop of 13 percentage points in the self-reported general health of retirees against a 15 percentage points drop in the general health of retirees using the unadjusted RD estimator when the bandwidth is restricted. Consequently, augmenting the standard RD estimator with covariates, which is restricted to be equivalent below and above the cut-off, can achieve substantial efficiency gain compared with the unadjusted RD estimator. The rows labeled CI length change show that approximately 16.5% and 6.5% efficiency gain when bandwidths are unrestricted and restricted, respectively, and optimally chosen using MSE with covariates. That is, including covariates in our RD estimator leads to inference improvements and precise point estimates for the impact of retirement on self-reported general health.

Table 1.21 reports that the RD estimates are homogeneous across the different weight schemes for observations that are far away from the cut-off point. Retirement is associated with a 13 percentage points drop in the proportion of retired people who report that they are in excellent, very good or good health across the different kernel functions. The RD results also show that the retirement impact did not change when the quadratic form of the running variable is introduced in the model. That is, the age-health profile of retired individuals is almost linear. The results are also robust based on the conventional and the robust bias-corrected RD estimators.

Table 1.20: Covariate-adjusted RD estimates – Self-reported health at cutoff age 65

	Not using covariates	Using covariate
	Standard Fuzzy RD	Covariate-Adjusted Fuzzy RD
Inference with h/p unrestricted		
RD treatment effect	-0.13172***	-0.10892***
Robust 95% CI	[-.168785, -.094664]	[-.139835, -.077995]
CI length change (%)		-16.5
robust p-value	0.000	0.000
h/p	0.626	0.633
Inference with h/p restricted		
RD treatment effect	-0.15295***	-0.13207***
Robust 95% CI	[-.214233, -.091661]	[-.189164, -.074974]
CI length change (%)		-6.8
robust p-value	0.000	0.000
$h\|p$	2|2	2|2
$n-\|n+$	43266|31261	43266|31261
obs	74527	74527
effective obs	6395|8707	3336|5212

Notes: (i) All estimates are computed using triangular Kernel and nearest neighbor heteroskedasticity-robust variance estimator as suggested by CCT framework. (ii) bandwidth used for h and b are data-driven MSE-optimal for either standard RD estimator or covariate-adjusted RD estimator. *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.21: RD estimates for different kernel and different polynomial orders – Self-reported health at cutoff age 65

RD estimator	1	2	3	4	5
Conventional	-0.1377***	-0.13543***	-0.13123***	-0.13383***	-0.13214***
Std. Err.	(0.01684)	(0.0156)	(0.02084)	(0.01832)	(0.02023)
Robust Bias corrected	-0.13769***	-0.13235***	-0.12986***	-0.13019***	-0.12956***
Std. Err.	(0.01996)	(0.01829)	(0.02367)	(0.02032)	(0.02286)
Robust 95% CI	[-.176813, -.09857]	[-.168198, -.096492]	[-.17625, -.083468]	[-.170023, -.090364]	[-0.174371, -0.084751]
Kernel Type	Uniform	Epanechnikov	Triangular	Uniform	Epanechnikov
Order Loc policy (ρ)	1	1	2	2	2
BW local poly (h)	2.5	3.2	4.4	4.8	4.4
BW Type	mserd	mserd	mserd	mserd	mserd
observations	74527	74527	74527	74527	74527
L|R	43266|31261	43266|31261	43266|31261	43266|31261	43266|31261
effective no	4543|6699	6395|8707	9880|11800	10914|12720	9880|11800

Notes (i) first column report two different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{fuzzy}$ and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{fuzzy}$. (ii) standard errors are in parentheses. (iii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

1.6.2 Nonparametric RD Results of Health-related Behavior

Table 1.22 shows the non-parametric RD estimates of the effect of retirement on retired individuals' alcoholic beverages consumption behavior. The results suggest that the transition from work to retirement is associated with a significant and negative jump in the discontinuity of drinking participation at the cut-off point by 8.2 to 18 percentage points. That is, retired individuals are less likely to be a drinker relative to non-retirees. The RD treatment effect indicates also that retirement has a strong effect on drinking participation immediately after retirement, which is indicated by the noticeable negative drop in drinking participation when using a one-year bandwidth. However, the effect of retirement on drinking participation is not overly sensitive to changes in the bandwidth after the first year of retirement, i.e., the drop in drinking participation is also robust for the different RD procedures. Figure (1.15) supports the negative jump in drinking participation. It shows estimates of the retirement effect at the cut-off point with 50 to 200 percent of the CV-optimal bandwidth selector, 5 percentage points incremental. Moreover, the RD estimates based on MSE-optimal bandwidth is consistent with the RD estimates based on the CV-optimal bandwidth. This large jump in drinking discontinuity may reflect that drinking participation is highly sensitive to the retirement lifestyle.

Tables 1.24 and 1.25 report the RD estimates for females and males separately. The results suggest that there is a significant gender heterogeneity in the effect of retirement on alcoholic consumption participation. Retired females are almost 6 percentage points less likely to be drinking relative to non-retired females while retired males are almost 12 percentage points less likely to be drinkers relative to non-retired males. In addition, the retirement effect on females' drinking participation is very sensitive to bandwidth.

Table 1.22: Fuzzy RD estimates of drinking participation - all sample at age 65

	Optimal bandwidth	Bandwidth h_{MSE}				
	4.2	1	2	3	4	5
Conventional	-0.08428***	-0.13052***	-0.09603***	-0.08669***	-0.08537***	-0.08212***
	(0.01739)	(0.04348)	(0.02774)	(0.02162)	(0.01801)	(0.01577)
$CI^{c}_{95\%}$	[-.118368 , -.050196]					
Bias-corrected	-0.08899***	-0.18063***	-0.12638***	-0.10256***	-0.09352***	-0.09145***
	(0.01739)	(0.04348)	(0.02774)	(0.02162)	(0.01801)	(0.01577)
$CI^{bc}_{95\%}$	[-0.123074, -0.054901]					
Robust	-0.08899***	-0.18063**	-0.12638***	-0.10256***	-0.09352***	-0.09145***
	(0.02106)	(0.06941)	(0.04104)	(0.03189)	(0.02676)	(0.02344)
$CI^{rbc}_{95\%}$	[-.130274,-.047702]					
obs	74507	74507	74507	74507	74507	74507
L|R	43254|31253	43254|31253	43254|31253	43254|31253	43254|31253	43254|31253
effective	9133|11135	1454|2722	3336|5212	5434 |7820	8373 |10411	11444|13209

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.23: Fuzzy RD estimation of the effect of retirement on the Drinking frequency (Drinking three days per week) all sample 65

	Optimal bandwidth	Bandwidths				
	3.9	1	2	3	4	5
Conventional	-0.0391***	-0.0638*	-0.0491**	-0.04543***	-0.03878***	-0.03362***
	(0.01454)	(0.03608)	(0.02248)	(0.0173)	(0.01431)	(0.01246)
$CI^{c}_{95\%}$	[-0.067593, -0.0106]					
Bias-corrected	-0.04424***	-0.09494***	-0.0571**	-0.05407***	-0.05146***	-0.04838***
	(0.01454)	(0.03608)	(0.02248)	(0.0173)	(0.01431)	(0.01246)
$CI^{bc}_{95\%}$	[-0.072738, -0.01574]					
Robust	-0.04424**	-0.09494	-0.0571*	-0.05407**	-0.05146**	-0.04838**
	(0.01735)	(0.05874)	(0.03389)	(0.0260)	(0.02158)	(0.01879)
$CI^{rbc}_{95\%}$	[-0.078247, -0.01024]					
obs	74527	74527	74527	74527	74527	74527
L|R	43266|31261	43266|31261	43266|31261	43266|31261	43266|31261	43266|31261

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Also, Figure 1.15 indicates that RD results based on the CV-optimal bandwidth choice are not consistent with RD results based on the MSE-optimal bandwidth choice for retired females. Although the RD results based on both bandwidth selectors report a drop in drinking participation for retired females, RD results based on CV-optimal bandwidth indicate confidence intervals that cross the zero line, which means that the retirement effect cannot be distinguished from zero. This difference in results can be partly due to the shorter CV-optimal bandwidth compared to the MSE-optimal bandwidth. Consequently, all ranges of bandwidths in Figure 1.15, which represent a 50 to 200 percent of the CV optimal bandwidth, are less than the MSE-optimal bandwidth. This is noticeable in figure 1.15, where retirement starts to have significant effects on retired females' drinking participation when the bandwidth is larger than 3.6. On the other hand, Figure 1.15 shows that retirement has a significant effect on the drinking participation of retired males. The RD results from the CV-optimal bandwidth are consistent with the RD results from the MSE-optimal bandwidth and both of them report a drop in drinking participation between 9 and 16 percentage points for retired males relative to non-retired males. Moreover, Figure 1.15 indicates that the negative drop in drinking participation for retired males is highly significant at all bandwidths.

The non-parametric RD treatment effect shows that retirement also has a positive side effect on the intensity of drinking for those who continue to drink after retirement. The RD estimator shows a significant negative jump in the probability of having alcohol consumption more than three days per week, which means that retirement reduces drinking frequency for those who continue to drink after retirement. Specifically, I found that retired individuals are 4 percentage points less likely to have alcoholic consumption more than three days per week relative to non-retired individuals. The RD results are consistent for both CV-optimal bandwidth

and MSE-optimal bandwidth and are significant for most of the bandwidth selections. The results also show significant gender heterogeneity in the effect of retirement on the frequency of alcohol consumption for retirees who continue to drink after retirement. The RD estimator indicates that there is a significant, but smaller, negative jump in the frequency of alcohol consumption per week for retired females than retired males. Retired females are around 2 percentage points less likely to have a drink more than 3 days per week with a confidence interval of 95%. The RD results of retired females based on CV-optimal bandwidth and the derived bandwidths in figure 1.15 indicate that the confidence intervals at different bandwidths include the value zero, which makes the RD estimates undistinguishable from zero for retired females relative to non-retired females. In contrast, retired males are 5 percentage points less likely to drink for more than three days per week relative to non-retired individuals and the RD results based on CV-optimal bandwidth are consistent with the RD results based on MSE-optimal bandwidth. The RD results from drinking participation and drinking intensity estimations indicate that retirement has more benefits for males than females regarding alcohol consumption. However, since the proportion of females who drink is, in general, less than the proportion of males who drink, retirement may have a small effect on alcohol consumption, and this effect is sensitive to retirement duration, implied by different bandwidths. In sum, there is gender heterogeneity in the effect of retirement on drinking participation and frequency and the negative drop in drinking is larger for retired males than retired females. The negative effect of retirement on drinking is robust for different bandwidths and across the three RD procedures applied throughout the study.

Table 1.24: Fuzzy RD estimates (Drinking participation) - Females only at age 65

	optimal bandwidth	bandwidths				
	5.6	1	2	3	4	5
Conventional	-0.05883***	-0.10905*	-0.03407	-0.04163	-0.05381**	-0.05958***
	(0.02004)	(0.06079)	(0.03804)	(0.02956)	(0.02456)	(0.02148)
$CI^{c}_{95\%}$	[-0.098119, -0.01955]					
Bias-corrected	-0.05759***	-0.19369***	-0.09003*	-0.03022	-0.03125	-0.03907*
	(0.02004)	(0.06079)	(0.03804)	(0.02956)	(0.02456)	(0.02148)
$CI^{bc}_{95\%}$	[-0.096877, -0.01831]					
Robust	-0.05759**	-0.19369**	-0.09003	-0.03022	-0.03125	-0.03907
	(0.02481)	(0.09534)	(0.0570)	(0.04394)	(0.03673)	(0.03212)
$CI^{rbc}_{95\%}$	[-0.106217, -0.00897]					
Obs	41804	41804	41804	41804	41804	41804
L|R	24323|17481	2432317481	24323|17481	24323|17481	24323|17481	24323|17481

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.25: RD estimates (drinking participation) - Males only at age 65

	Optimal bandwidth	Bandwidths				
	3.5	1	2	3	4	5
Conventional	-0.11471***	-0.14023**	-0.14958***	-0.12203***	-0.10775***	-0.09656***
	(0.02791)	(0.06201)	(0.04019)	(0.03131)	(0.02614)	(0.02291)
$CI^{c}_{95\%}$	[-0.169409, -0.06002]					
Bias-corrected	-0.12589***	-0.13015**	-0.15091***	-0.16672***	-0.14662***	-0.13345***
	(0.02791)	(0.06201)	(0.04019)	(0.03131)	(0.02614)	(0.02291)
$CI^{bc}_{95\%}$	[-0.180584, -0.07119]					
Robust	-0.12589***	-0.13015	-0.15091**	-0.16672***	-0.14662***	-0.13345***
	(0.03277)	(0.10081)	(0.0590)	(0.04598)	(0.03861)	(0.03383)
$CI^{rbc}_{95\%}$	[-0.190112, -0.06167]					
L|R	18931|13772	18931|13772	18931|13772	18931|13772	18931|13772	18931|13772
Obs	32703	32703	32703	32703	32703	32703

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.26: Fuzzy RD estimation on the drinking frequency – cutoff at 65 (Drinking at least three days/week) - Females only

	optimal bandwidth	bandwidths				
	3.9	1	2	3	4	5
Conventional	-0.01684	-0.04541	-0.01555	-0.01611	-0.01714	-0.01874
	(0.01669)	(0.04362)	(0.02591)	(0.01994)	(0.01651)	(0.01441)
$CI^{c}_{95\%}$	[-0.049539, 0.015865]					
Bias-corrected	-0.01409	-0.12772***	-0.03057	-0.017	-0.01387	-0.01358
	(0.01669)	(0.04362)	(0.02591)	(0.01994)	(0.01651)	(0.01441)
$CI^{bc}_{95\%}$	[-0.046788, 0.018616]					
Robust	-0.01409	-0.12772*	-0.03057	-0.017	-0.01387	-0.01358
	(0.02026)	(0.07288)	(0.04045)	(0.03027)	(0.02496)	(0.02172)
$CI^{rbc}_{95\%}$	[-0.053794, 0.025622]					
Obs	41816	41816	41816	41816	41816	41816
L|R	24328|17488	24328|17488	24328|17488	24328|17488	24328|17488	24328|17488

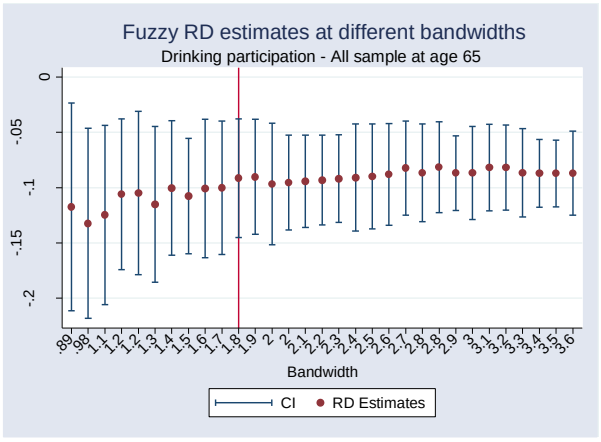
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

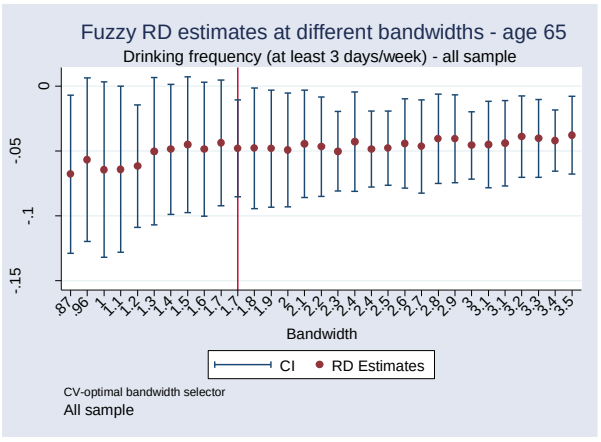
Table 1.27: Fuzzy RD estimation of the effect of retirement on the drinking frequency - Males only at cutoff 65

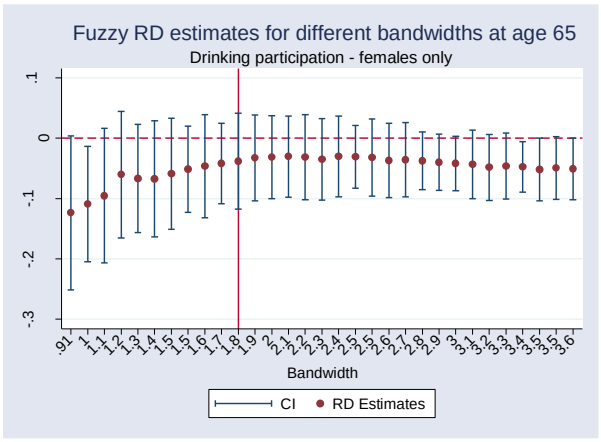
	optimal bandwidth	bandwidths				
		1	2	3	4	5
Conventional	-0.04696*	-0.06378	-0.06674*	-0.06017**	-0.04733*	-0.03661*
	(0.02367)	(0.05757)	(0.03717)	(0.0287)	(0.02378)	(0.02072)
$CI^{c}_{95\%}$	[-0.093364, -0.00056]					
Bias-corrected	-0.05682**	-0.01817	-0.06656*	-0.07416**	-0.07321***	-0.06774***
	(0.02367)	(0.05757)	(0.03717)	(0.0287)	(0.02378)	(0.02072)
$CI^{bc}_{95\%}$	[-0.10322, -0.01042]					
Robust	-0.05682**	-0.01817	-0.06656	-0.07416*	-0.07321**	-0.06774**
	(0.02763)	(0.09418)	(0.05467)	(0.04263)	(0.0356)	(0.03107)
$CI^{rbc}_{95\%}$	[-0.110971, -0.00267]					
Obs	32711	32711	32711	32711	32711	32711
L|R	18938|13773	18938|13773	18938|13773	18938|13773	18938|13773	18938|13773

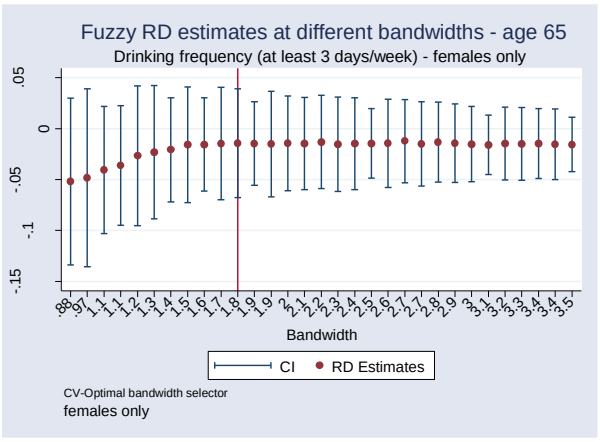
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

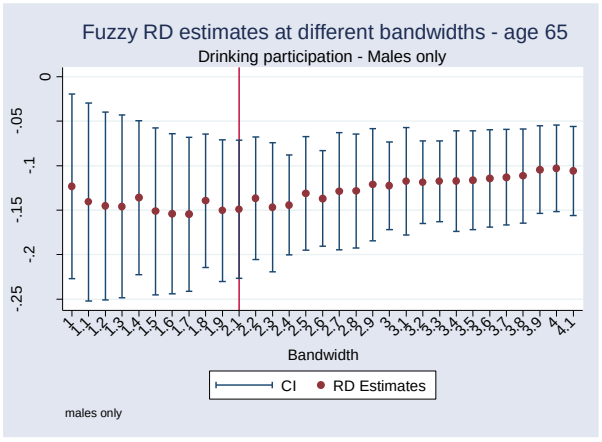
Figure 1.15: Fuzzy RD estimates at different bandwidth for drinking and smoking – age 65

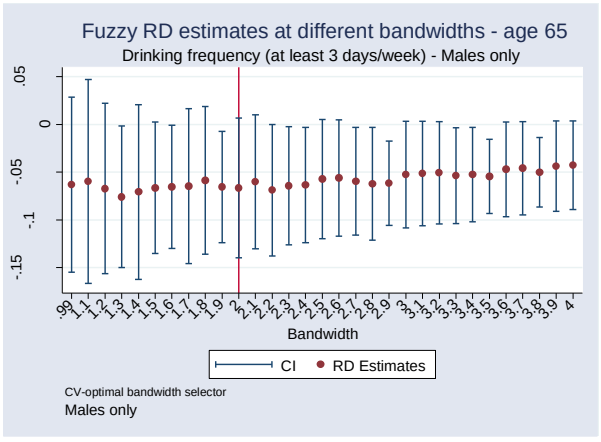












Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for alcoholic consumption. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

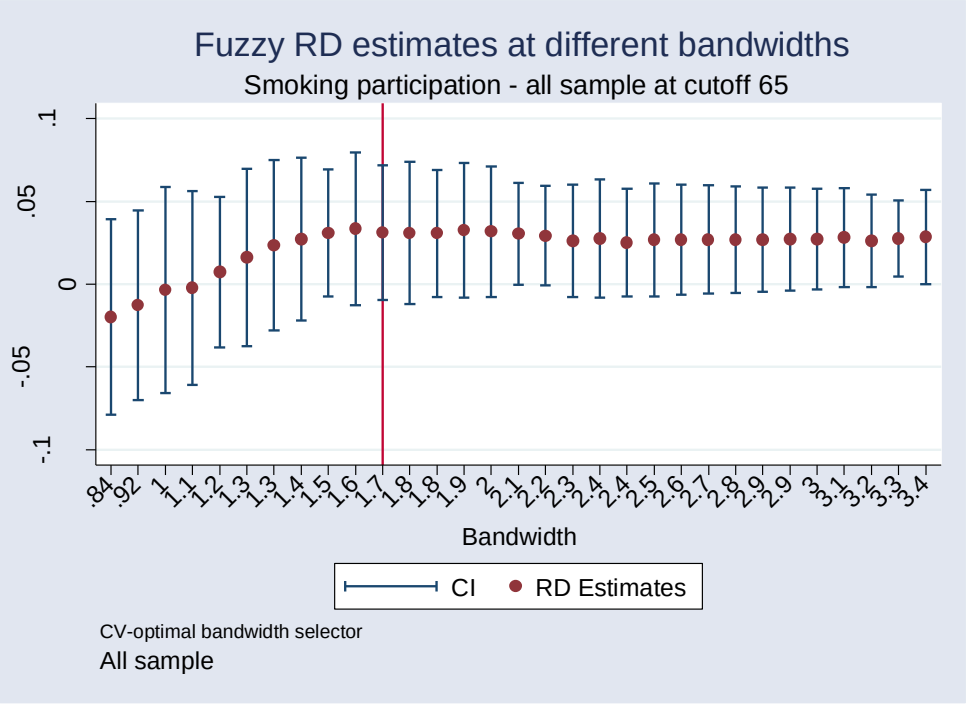
Table 1.28 reports the non-parametric RD treatment effect of retirement on smoking participation. The results indicate that retirement has a positive effect on smoking. Retired individuals are almost 3 percentage points more likely to be a smoker than non-retirees. However, the RD estimates are very sensitive to bandwidth selection, which sheds light on the robustness of the RD effect at the cut-off. Moreover, the RD estimates are not significant at other bandwidths. Figure 1.16 suggests that the RD results based on the CV-optimal bandwidth are not consistent with the RD results based on MSE-optimal bandwidth, and although both show a positive jump in the smoking status of retirees, the result is not significant for the former bandwidth selector. The figure also reports that retirement has a positive effect on smoking status, but this effect is not statistically significant. To check if there is gender heterogeneity in the impact of retirement on smoking status, the RD effect is estimated for males and females separately. Tables 1.29 and 1.30 suggest that retirement has a heterogeneous impact on females and males. Although the RD treatment effect finds a positive impact of retirement on both males and females, the effect of retirement is statistically insignificant for females at all bandwidth selections. In contrast, the impact of retirement on retired males is positive and statistically significant, but the retirement effect is sensitive to bandwidth selection. The RD results based on the CV-optimal bandwidth support the positive impact of retirement for males and the sensitivity of the results to the selected bandwidth. Together, retirement impact on smoking status is positive and has a significant effect on retired males but not on retired females.

Table 1.28: Fuzzy RD treatment effect of retirement on smoking participation -All sample 65

	Optimal bandwidth	Bandwidth				
	3.6	1	2	3	4	5
Conventional	0.0303**	-0.01042	0.03204	0.02718*	0.03075**	0.03247***
	(0.01377)	(0.03184)	(0.02025)	(0.01567)	(0.01308)	(0.01147)
$CI^{c}_{95\%}$	[0.003316, 0.057289]					
Bias-corrected	0.02745**	-0.02936	0.01589	0.02782*	0.02473*	0.02659**
	(0.01377)	(0.03184)	(0.02025)	(0.01567)	(0.01308)	(0.01147)
$CI^{bc}_{95\%}$	[0.000468, 0.054441]					
Robust	0.02745*	-0.02936	0.01589	0.02782	0.02473	0.02659
	(0.01682)	(0.0520)	(0.03011)	(0.02336)	(0.01945)	(0.01701)
$CI^{rbc}_{95\%}$	[0.005521, 0.06043]					
Obs	(74108)	(74108)	(74108)	(74108)	(74108)	(74108)
L|R	43060|31048	43060|31048	43060|31048	43060|31048	43060|31048	43060|31048

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.16: Fuzzy RD estimates at different bandwidth for smoking participation



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for smoking participation. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

Table 1.29: Fuzzy RD treatment effect of retirement on smoking participation -Females only at cutoff 65

	optimal bandwidth	bandwidth				
	4.8	1	2	3	4	5
Conventional	0.02281	0.00558	0.03722	0.0231	0.0224	0.02316
	(0.01539)	(0.0416)	(0.0264)	(0.02051)	(0.01717)	(0.0151)
$CI^{c}_{95\%}$	[-0.007348, 0.052968]					
Bias-corrected	0.02199	0.02561	0.02836	0.03891*	0.02594	0.02285
	(0.01539)	(0.0416)	(0.0264)	(0.02051)	(0.01717)	(0.0151)
$CI^{bc}_{95\%}$	[-0.008168, 0.052148]					
Robust	0.02199	0.02561	0.02836	0.03891	0.02594	0.02285
	(0.01879)	(0.06717)	(0.03925)	(0.03053)	(0.02546)	(0.02229)
$CI^{rbc}_{95\%}$	[-0.014839, 0.058819]					
obs	41581	41581	41581	41581	41581	41581
L|R	24193|17388	24193|17388	24193|17388	24193|17388	24193|17388	24193|17388

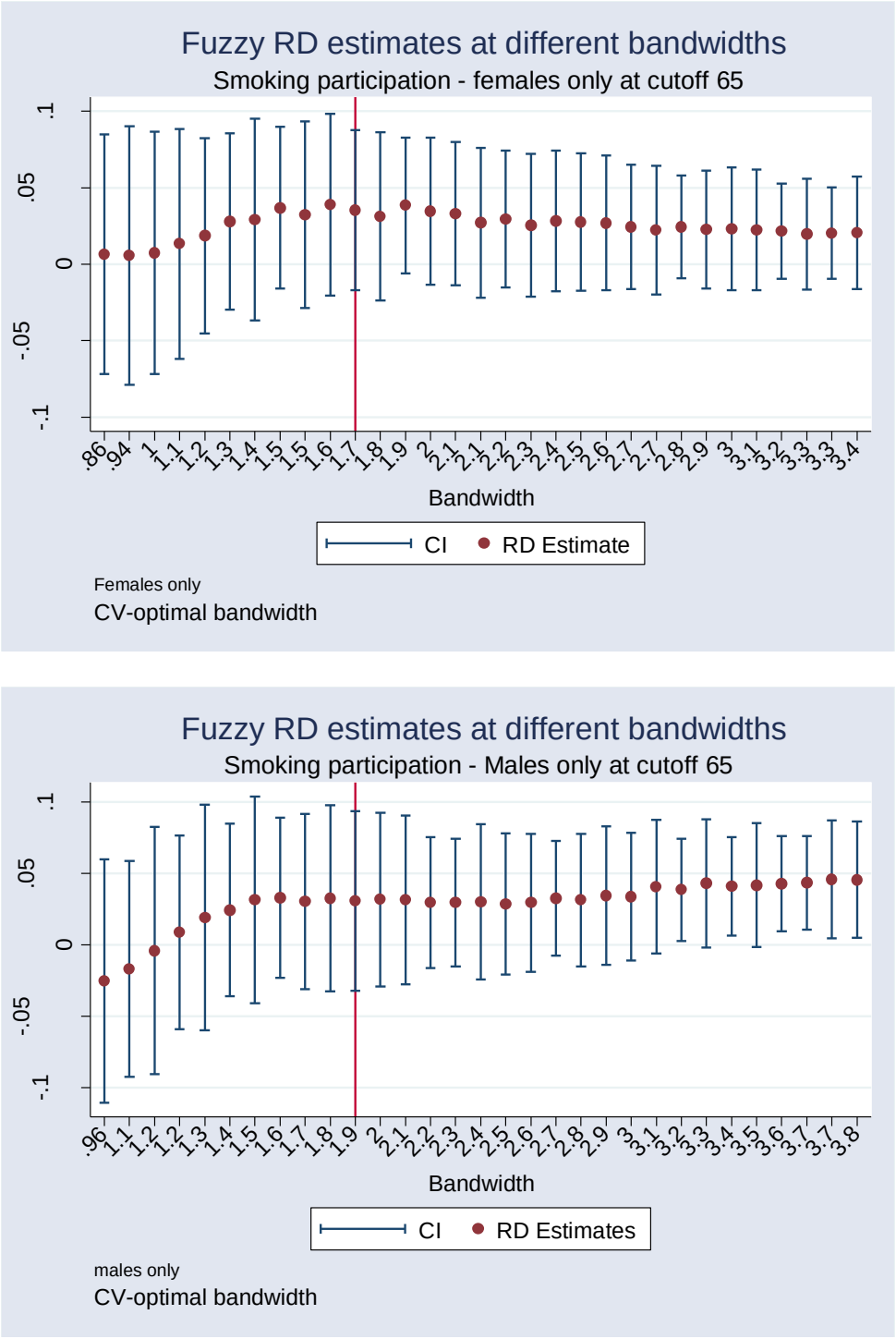
Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.30: Fuzzy RD treatment effect of retirement on smoking participation – Males only at cutoff 65

	optimal bandwidth	bandwidth				
		1	2	3	4	5
Conventional	0.04611**	-0.02481	0.03047	0.03605*	0.04414**	0.04645***
	(0.0183)	(0.04877)	(0.03131)	(0.02419)	(0.02015)	(0.01763)
$CI^{c}_{95\%}$	[0.010243, 0.08198]					
Bias-corrected	0.04588**	-0.09283*	0.00643	0.01964	0.0279	0.03552**
	(0.0183)	(0.04877)	(0.03131)	(0.02419)	(0.02015)	(0.01763)
$CI^{bc}_{95\%}$	[0.010011,0.081748]					
Robust	0.04588**	-0.09283	0.00643	0.01964	0.0279	0.03552
	(0.0224)	(0.08126)	(0.04633)	(0.03597)	(0.02997)	(0.02619)
$CI^{rbc}_{95\%}$	[0. 001973, 0.089786]					
obs	32527	32527	32527	32527	32527	32527
L|R	18867|13660	18867|13660	18867|13660	18867|13660	18867|13660	18867|13660

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.17: Fuzzy RD estimates at different bandwidth for males and females separately



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice for smoking participation. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

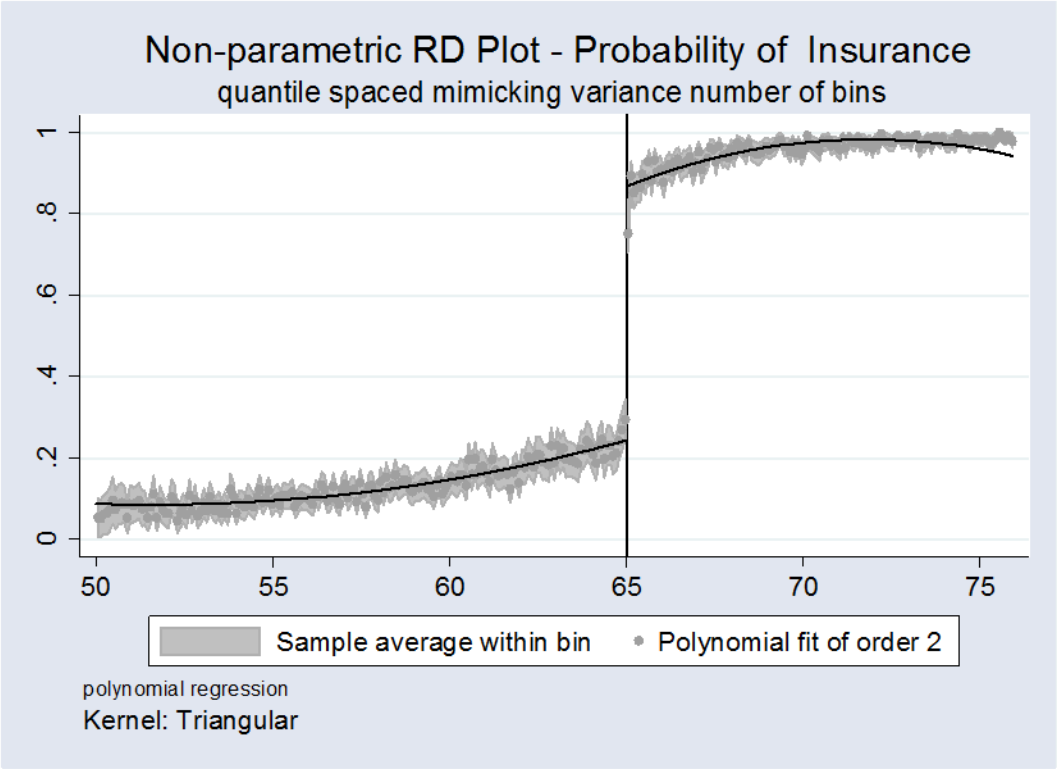
1.7. Non-Parametric Fuzzy RD Robustness Check

I checked the robustness of the RD estimates in the analysis of the effect of retirement on health using a variety of distinct procedures. I proved that the main findings are robust to different model specifications, parametric and nonparametric regression specifications. I also demonstrate that the RD estimates of the effect of retirement do not change across different weighting schemes that give higher weight to the observations in the neighborhood of the cut-off point but away from the cutoff. I also checked the robustness of the results under different polynomial procedures, linear and quadratic, and there was no significant change in the results. The robustness of the regression is verified by estimating the effect of retirement at different bandwidths around the cut-off point that shows increasing increments by 5 percentage points, 50% to 200% of the optimal bandwidth, which firmly supports the validity of the results. Also, the estimations showed that adding covariates to the main model in the RD setting did not change the results dramatically but improved the efficiency of the estimation results. I further assess the robustness of the results to the dramatic increase in health insurance eligibility at age 65. To assess the robustness of the RD results giving the positive jump in the Medicare health eligibility in the U.S., several robustness checks were performed.

Medicare is the major federal program that provides health insurance coverage for nearly all Americans who turn 65. To be eligible for Medicare, individuals must be at least 65 years and have paid Medicare taxes for at least 10 years. However, some persons qualify for Medicare before age 65 if they received social security disability insurance (SSDI). Individuals who meet these eligibility requirements qualify directly for medical hospital insurance (part A) free of charge, but they have to choose whether or not to opt into medical insurance (part B), which is available for a monthly premium (Card et al., 2004). It is well-documented in empirical work

that health insurance coverage increases at age 65, and that access to Medicare is associated with a significant increase in health care utilization (Card et al., 2004, 2008, 2009). Christelis et al. (2014) found that Medicare coverage increased by 73 percentage points at age 65.

Figure 1.18: Non-parametric RD plot – the probability of having insurance at age 65



Note: Based on the quantile spaced mimicking variance, the figure shows that the percentage of individuals who are covered by insurance positively jump at age 65.

The drastic change in health care coverage at age 65 could create a problem for the RD estimators. Specifically, the effect of retirement on health at age 65 could potentially be driven by Medicare eligibility and health care utilization, not necessarily be driven by retirement status. To address this issue, I conducted a falsification test of the RD analysis using several checks. First, I conducted a similar Fuzzy RD analysis at age 62, where a significant jump in retirement is also observed because this is the earliest age at which Americans can receive social security payment. Also, the positive jump in the retirement rate at age 62 is associated with a significant

decline in health care coverage, which can neutralize the effect of health care coverage in the analysis. Second, I run the RD analysis on the low-income subgroup. This group includes individuals who are in the lowest 20% of the income distribution in the dataset or individuals whose income is below the federal poverty line after adjustment for the number of persons in the family (HHS, 2018)[11]. This subgroup will include the persons who may be eligible for the Medicaid health insurance program before age 65. Medicaid is a joint federal and state program that helps low-income adults and people with certain disabilities cover health care costs. Eligibility to Medicaid is based on income and family size. Although the HRS includes a question about whether persons are enrolled in the Medicaid program, there are some cases in which individuals did not report any information about their health insurance coverage before age 65. Therefore, using the previous income criteria may help in identifying persons who may be eligible to have Medicaid insurance coverage to get a reasonable sample size to this subgroup. Third, I run the RD analysis at age 65 for the subgroup of individuals who were eligible for health care coverage before 65 or individuals who have other health care coverage, private or public, for at least five years before they retired. This, then, includes individuals who already had health insurance coverage before reaching age 65, and who continue to have health care coverage after age 65, whether the coverage is public or private.

[11] According to the U.S. federal poverty guidelines used to determine eligibility for certain federal programs, the 2018 poverty guidelines in 48 states are as follow: $12,140 for one person family, $16,460 for two person family, $20,780 for three person family, $25,100 for four person family, $29,420 for five person family, $33,740 for six persons family, $38,060 for seven person family, $42,380 for 8 persons family, and $4,320 for each additional person for families with more than 8 persons.

1.7.1 Regression Discontinuity at Early Retirement Discontinuity – Cut-off at Age 62.

As discussed above, Medicare provides nearly universal health care coverage for Americans over 65. Previous literature suggests that individuals may base their retirement decision on their Medicare eligibility status, which could create an identification issue in the fuzzy RD estimation strategy. Specifically, changes in health outcomes at age 65 could potentially be driven by Medicare eligibility, not necessarily by the retirement decision. To address this issue, I conduct a falsification test of the RD analysis around the early retirement age of 62 in U.S. Table 1.31 indicates the change in health outcomes when individuals retired at age 62. All results are based on the robust bias-corrected RD estimators and use the MSE-optimal bandwidth selector. Panel A shows that retirement at age 62 is associated with a significant decline in the cognitive functioning test score. The fuzzy RD estimates suggest that retired individuals experience a drop in the cognitive score by about 0.52 points with a 95% confidence interval [-0.824629, -0.22317], and this is equivalent to a 5% drop in the cognitive functioning score. Panel B reports the impact of retirement on mental health. The RD estimates suggest that there is an increase in retired individuals' CESD depression scale by almost 0.54 points, which is equivalent to a 37% decrease in mental health status relative to non-retirees. Similarly, retirement has a negative effect on the self-reported general health of retirees. The proportion of retired individuals who report that they are in excellent, very good, or good health dropped by almost 22% relative to non-retirees. In addition, the RD estimates are robust at different bandwidths except for cognitive functioning at one-year bandwidth, which is in line with the estimated effect of retirement at age 65 where retirement does not affect cognitive functioning directly but with a time lag. Based on the RD treatment effect results of retirement at the early retirement age 62, the findings support the results at the official retirement age 65, i.e., the fuzzy RD retirement effects on health

outcomes are robust at the official age of retirement, and the change in health outcomes can be explained by the change in retirement status at this age.

Table 1.31: Fuzzy RD Estimates at Different Bandwidths – All Sample at Cutoff Age 62

	optimal bandwidth	Bandwidths				
Panel A: Cognitive Functioning						
	2.9	1	2	3	4	5
Robust RD	-0.5239***	-0.35244	-0.58106***	-0.5233***	-0.52915***	-0.57133***
	(0.15344)	(0.34066)	(0.20195)	(0.15294)	(0.12643)	(0.1098)
$CI_{95\%}^{rbc}$	[-0.824629, -0.22317]					
Panel B: Mental Health						
	4.5	1	2	3	4	5
Robust RD	0.54417***	0.39717**	0.50182***	0.54199***	0.54591***	0.53852***
	(0.06289)	(0.17284)	(0.10685)	(0.0821)	(0.06842)	(0.05967)
$CI_{95\%}^{rbc}$	[.420911, .667438]					
Panel C: Self-reported General Health						
	3.7	1	2	3	4	5
Robust RD	-0.21693***	-0.31057***	-0.23059***	-0.2210***	-0.21629***	-0.21038***
	(0.01568)	(0.03746)	(0.02339)	(0.01803)	(0.01502)	(0.01311)
$CI_{95\%}^{rbc}$	[-.247664, -.186189]					

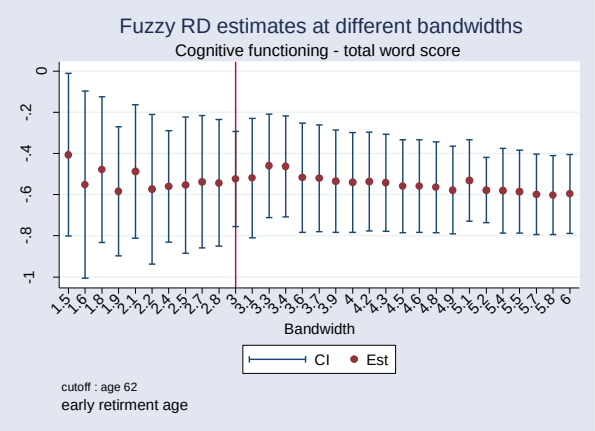
Notes: (i) All estimates are computed using a triangular kernel and MSE-optimal bandwidth selector(ii) first column report the bias-corrected RD estimates with a robust variance estimator $\hat{\tau}_{FRD}^{rbc}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.19 plots the RD results, which are constructed by using the conventional RD procedure and CV-bandwidth selector. The RD estimation is repeated at different bandwidths, with a range between 50% and 200% of the CV-optimal bandwidth. All estimates are based on the triangular kernel, and standard errors are bootstrapped using 120 simulations. These are used to construct a 95% confidence interval for the RD estimates, as shown on each graph. Plot 1 shows that retirement is associated with a negative impact on the cognitive score at different bandwidths and that there is no significant change in the value of the cognitive functioning score when the bandwidth changes. Similarly, plots 2 and 3 show that retirement has a negative effect on both

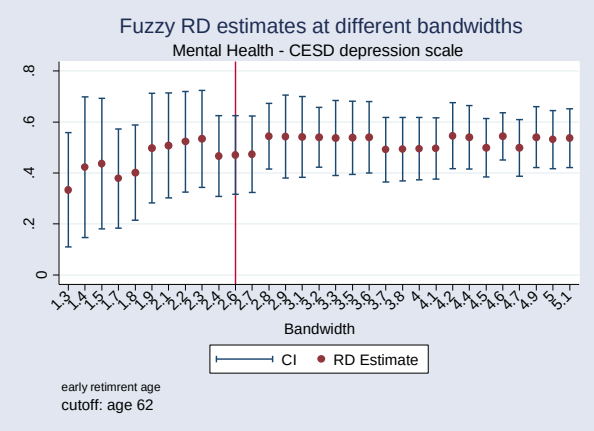
mental health and self-reported general health of retired individuals at different bandwidths. The results in table 1.31 and figure 1.19 show that the RD estimation from the MSE-optimal bandwidth selector is consistent with the results from the CV-optimal bandwidth, and the RD estimates are robust at different bandwidths. Consequently, the RD estimates are robust at the early retirement discontinuity and the official age of retirement discontinuity.

Figure 1.19: Robustness of RD estimates of health outcomes at cut-off age 62

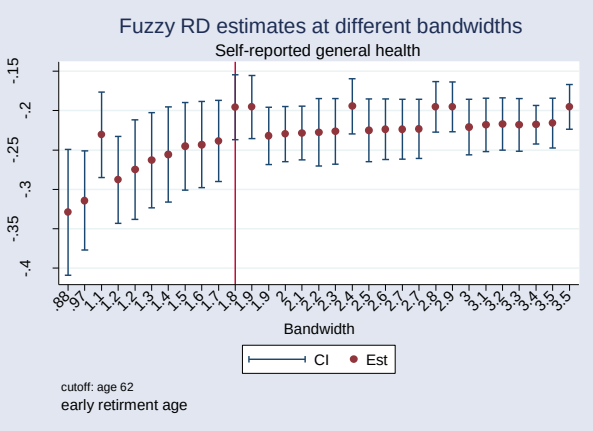
Plot 1



Plot 2



Plot 3



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

Table 1.32 reports the RD results for health behavior at the early retirement age (62). The RD suggests that retirement has a negative and significant impact on drinking. Retired individuals are almost 7 percentage points less likely to be a drinker after retirement relative to non-retirees. The results are not sensitive to different bandwidths. Although the RD results suggest that retirement negatively impacts the drinking intensity of retired individuals, the results are sensitive to bandwidth choices. In contrast, retirement status is associated with a positive impact on smoking, where retired individuals are almost 3 percentage points more likely to smoke relative to non-retirees.

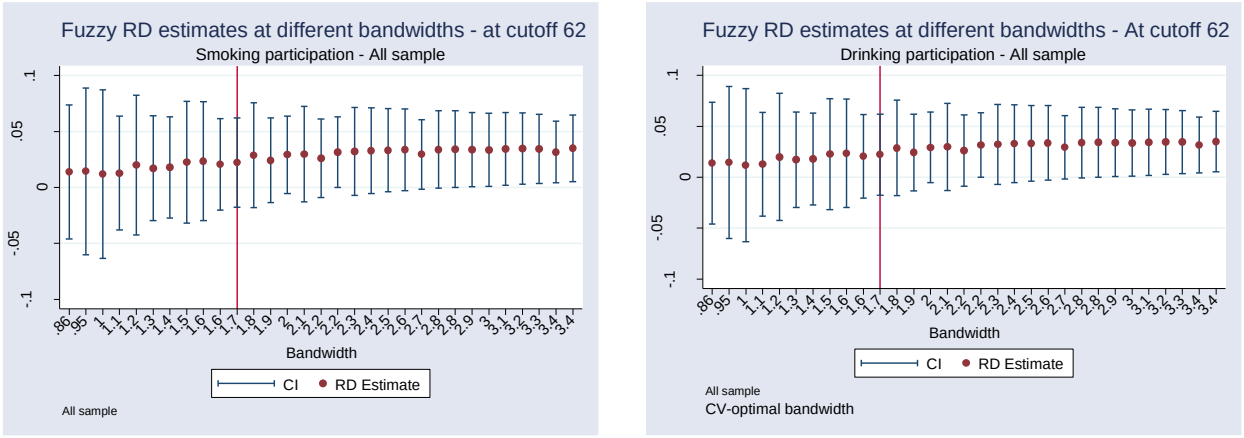
The RD results based on the CV-optimal bandwidth, shown in Figure 1.20, are consistent with the RD results based on the MSE-optimal bandwidth selection. Figure 1.20 shows that the retirement impact on drinking and smoking participation is sensitive to the bandwidth selections and retirement practice a significant effect on them for the bandwidths at least equal to the MSE-optimal bandwidth. As indicated by the figure, the confidence intervals for both drinking participation and smoking participation include the value zero, which makes the RD estimates indistinguishable from the zero. In summary, the RD estimates at the early retirement age, 62, are in line with the RD estimates at the official retirement age, 65, which means the RD of retirement is not affected much by the eligibility to access Medicare at age 65, and the results are robust.

Table 1.32: Fuzzy RD estimates at different bandwidths (health-related behavior) – All sample at cutoff age 62

	Optimal bandwidth	Bandwidth				
		Panel A: Drinking participation				
	3.2	1	2	3	4	5
Robust	-0.0738***	-0.1236**	-0.1164***	-0.0932***	-0.0807***	-0.0744***
	(0.0238)	(0.0699)	(0.0401)	(0.0305)	(0.0253)	(0.0221)
$CI^{rbc}_{95\%}$	[-0.120285, -0.027205]					
		Panel B: Drinking frequency				
	5.2	1	2	3	4	5
Robust	-0.0278**	-0.0456	-0.0266	-0.0241	-0.0256*	-0.0280**
	(0.0111)	(0.0374)	(0.0219)	(0.0164)	(0.0136)	(0.0118)
$CI^{rbc}_{95\%}$	[-.04962, -.006057]					
		Panel C: Smoking participation				
	4.1	1	2	3	4	5
Robust	0.0381***	0.0147	0.0292	0.0340**	0.0373***	0.0397***
	(0.0137)	(0.0367)	(0.0223)	(0.0169)	(0.0140)	(0.0122)
$CI^{rbc}_{95\%}$	[.011164, .064945]					

Notes: (i) All estimates are computed using a triangular kernel and MSE-optimal bandwidth selector(ii) first column report the bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Figure 1.20: Fuzzy RD robustness of drinking and smoking behavior



Notes: Non-parametric Regression discontinuity robustness of bandwidth choice. Each point is a separate non-parametric regression discontinuity point estimate. Fuzzy non-parametric RD treatment effect estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. Bootstrapped standard error are based on 120 simulations to construct 95% CI.

1.7.2 Fuzzy RD Estimates for the Low-Income Subgroup

Table 1.33 reports the RD estimates for individuals in the low-income group. The RD estimate in the first column shows that the jump in retirement probability was associated with a negative jump in the reported word recall test score. That is, the effect of retirement on cognitive function is negative and highly significant. The RD estimator suggests that low income retired individuals experienced a drop in their cognitive test score by about 0.7 points with a 95% confidence interval [-0.946994, -0.410576]. This decline in cognitive functioning is equivalent to an approximately 6.7% decline in the cognitive functioning score of low-income retired individuals. The negative jump in total word recall for the low-income group is relatively lower than the average of the whole sample, 6.7% compared to 8%, respectively. Moreover, the effect of retirement on cognitive functioning is not sensitive to the change in the selected bandwidth for all reported bandwidths.

Table 1.33: Fuzzy RD estimates - Cognitive functioning (Total word recall) low-income group at age 65

	Optimal bandwidth	Bandwidth				
	4	1	2	3	4	5
Conventional	-0.67878***	-0.68736**	-0.68938***	-0.71061***	-0.68166***	-0.67278***
	(0.13684)	(0.33112)	(0.21103)	(0.16575)	(0.13872)	(0.12167)
$CI^{c}_{95\%}$	[-0.946994, -0.410576]					
Bias-corrected	-0.73075***	-1.1844***	-0.68425***	-0.70938***	-0.73223***	-0.70206***
	(0.13684)	(0.33112)	(0.21103)	(0.16575)	(0.13872)	(0.12167)
$CI^{bc}_{95\%}$	[-0.998959, -0.462541]					
Robust	-0.73075***	-1.1844**	-0.68425**	-0.70938***	-0.73223***	-0.70206***
	(0.16234)	(0.53633)	(0.31779)	(0.24587)	(0.20589)	(0.18081)
$CI^{rbc}_{95\%}$	[-1.04894, -0.412564]					
Obs	56314	56314	56314	56314	56314	56314

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.34: Fuzzy RD estimates - Mental health (CES-D Depression scale) Low income at age 65

	Optimal bandwidth	Bandwidth				
	4.4	1	2	3	4	5
Conventional	0.28095***	0.39379**	0.23981**	0.26609***	0.27775***	0.28859***
	(0.07744)	(0.17383)	(0.12243)	(0.09853)	(0.08285)	(0.0729)
$CI^{c}_{95\%}$	[0.129176, 0.432716]					
Bias-corrected	0.27821***	0.89414***	0.37629***	0.26616***	0.25265***	0.25461***
	(0.07744)	(0.17383)	(0.12243)	(0.09853)	(0.08285)	(0.0729)
$CI^{bc}_{95\%}$	[0.126436, 0.429976]					
Robust	0.27821***	0.89414***	0.37629***	0.26616*	0.25265**	0.25461**
	(0.09601)	(0.26053)	(0.17125)	(0.1408)	(0.12143)	(0.1076)
$CI^{rbc}_{95\%}$	[0.090036, 0.466376]					
L|R	21638|30826	21638|30826	21638|30826	21638|30826	21638|30826	21638|30826
Obs	52464	52464	52464	52464	52464	52464

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.35: Fuzzy RD estimates - Self-reported health (Low-income group at age 65)

	Optimal	Bandwidth				
	4.2	1	2	3	4	5
Conventional	-0.11746***	-0.13694***	-0.10793***	-0.11092***	-0.11739***	-0.11757***
	(0.0173)	(0.03978)	(0.02678)	(0.02145)	(0.01802)	(0.01586)
$CI^{c}_{95\%}$	[-0.151365, -0.08355]					
Bias-corrected	-0.11535***	-0.19803***	-0.14124***	-0.1116***	-0.10586***	-0.11222***
	(0.0173)	(0.03978)	(0.02678)	(0.02145)	(0.01802)	(0.01586)
$CI^{bc}_{95\%}$	[-0.149253, -0.08144]					
Robust	-0.11535***	-0.19803***	-0.14124***	-0.1116***	-0.10586***	-0.11222***
	(0.02046)	(0.06215)	(0.03889)	(0.03108)	(0.02655)	(0.02347)
$CI^{rbc}_{95\%}$	[-0.155452, -0.07524]					
L|R	21638|30826	21638|30826	21638|30826	21638|30826	21638|30826	21638|30826
Obs	52464	52464	52464	52464	52464	52464

Notes: (i) All estimates are computed using a triangular kernel. (ii) first column report three different procedures: conventional RD estimates with a conventional variance estimator $\hat{\tau}^{c}_{FRD}$; bias-corrected RD estimates with a conventional variance estimator $\hat{\tau}^{bc}_{FRD}$; and bias-corrected RD estimates with a robust variance estimator $\hat{\tau}^{rbc}_{FRD}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

1.7.3 Fuzzy RD Estimates for the Consistently Insured Subgroup (Before and After Retirement)

Table 1.36 reports the RD results for a subpopulation of individuals who have insurance coverage for at least 5 years before the official age of retirement, and continue to be covered by insurance after retirement. This subpopulation is used to check the robustness of the RD estimates at age 65 and mitigate the effect of the discontinuity in Medicare eligibility at age 65. The non-parametric results suggest that retired individuals who were covered by insurance at least five years before their transition into retirement experienced a significant decline in their cognitive score by about 1.2 points, which is equivalent to around 12 percentage points in the cognitive score relative to non-retires. The RD results also are not sensitive to bandwidth selection, i.e., RD treatment impact on the cognitive score is robust to different bandwidths. Panel B in the table reports the RD estimates for mental health status. The results suggest that retirement is associated with a negative drop in the mental health status of retirees. Retired individuals experience a significant increase in their CESD depression scale by about 0.5 points relative to non-retired individuals. However, the results are not highly significant at all bandwidth selections. Panel C reports the RD results of the impact of retirement on self-reported health. Although the proportion of retired individuals who report they are in good health dropped by almost 6 percentage points, the impact of retirement on self-reported health of retirees is insignificant at the other bandwidths. This finding sheds light on the role of having access to health insurance on self-reported health status for this subpopulation. The RD results on the subpopulation of retirees who have health insurance coverage before the retirement is in line with the RD results from the whole sample at age 65, except the effect on self-reported health,

the results are not robust enough to support the RD treatment effect for self-reported health at age 65.

Table 1.36: RD estimates for retirees with any kind of insurance, at least 5 years before age 65

	MSE-Optimal bandwidth	bandwidth				
		Panel A: Cognitive functioning				
	4.3	1	2	3	4	5
RD	-1.2529***	-1.7269*	-1.2788*	-1.3252**	-1.3108***	-1.1721***
	(0.44329)	(0.95196)	(0.65827)	(0.53933)	(0.46023)	(0.4148)
$CI_{95\%}^{rbc}$	[-2.1217, -0.384025]					
		Panel B: Mental health				
	5	1	2	3	4	5
RD	0.48331**	0.70627*	0.52568*	0.50821*	0.46698**	0.45394**
	(0.20249)	(0.39682)	(0.31564)	(0.27334)	(0.23925)	(0.2153)
$CI_{95\%}^{rbc}$	[.086447, .880176]					
		Panel C: Self- reported health				
	5	1	2	3	4	5
RD	-0.06982*	-0.17014*	-0.03372	-0.021	-0.04444	-0.06224
	(0.04098)	(0.09396)	(0.0703)	(0.05926)	(0.0505)	(0.04642)
$CI_{95\%}^{rbc}$	[-.150138, 0.010496]					

Notes: (i) All estimates are computed using a triangular kernel and MSE-optimal bandwidth selector (ii) first column report the bias-corrected RD estimates with a robust variance estimator $\hat{\tau}_{FRD}^{rbc}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.37 reports the RD results for health-related behavior for a subgroup of individuals who have health insurance coverage at least 5 years before age 65 and continue to have health coverage after retirement. The results suggest that retirement has a negative impact on drinking, with retirees 9 percentage points less likely to be drinkers compared to non-retirees. However, the RD results are sensitive to bandwidth selection. In contrast, the RD results show an insignificant effect of retirement on drinking frequency and smoking. These findings are inconsistent with the RD results at the official age of retirement.

Table 1.37: RD estimates for retirees with any kind of insurance, at least 5 years before age 65

	MSE-Optimal bandwidth	Bandwidth				
Panel A: Drinking participation						
	5.07	1	2	3	4	5
Robust RD	-0.0961**	-0.1567	-0.1004	-0.1130*	-0.0962*	-0.0965**
	(0.0485)	(0.1187)	(0.0832)	(0.0672)	(0.0557)	(0.0489)
$CI_{95\%}^{rbc}$	[-.191238, -.001014]					
Panel B: Drinking frequency						
	5.3	1	2	3	4	5
Robust RD	-0.0269	0.0462	-0.0182	-0.0254	-0.0100	-0.0191
	(0.0293)	(0.0734)	(0.0610)	(0.0517)	(0.0439)	(0.0386)
$CI_{95\%}^{rbc}$	[-.084332, .030638]					
Panel C: Smoking participation						
	6	1	2	3	4	5
Robust RD	-0.0051	-0.0166	-0.0362	-0.0342	-0.0254	-0.0147
	(0.0335)	(0.0977)	(0.0648)	(0.0523)	(0.0437)	(0.0384)
$CI_{95\%}^{rbc}$	[-.07075, .060649]					

Notes: (i) All estimates are computed using a triangular kernel and MSE-optimal bandwidth selector(ii) first column report the bias-corrected RD estimates with a robust variance estimator $\hat{\tau}_{FRD}^{rbc}$. (iii) standard errors are in parentheses. (iv) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

1.8. The Validity of the Regression Discontinuity Design

Although the mechanism through which individuals are assigned to treatment is based on observable features, and the quantitative effect of the treatment can be easily identified, the assignment rule can be manipulated and the continuity assumption, the spirit of the RD setting, violated and hence estimates of the RD estimator invalidated. The validation of the RD is based on the continuity assumption, which is an unobservable assumption. However, the RD design has a variety of empirical methods that can provide plausible and indirect evidence about the validity of its assumptions (Cattaneo, 2018).

1.8.1 Predetermined Covariates

One of the most important falsification tests in the RD literature is to test whether the treated and untreated units near the cut-off point (threshold neighborhood) are similar in terms of observable features. That is, units above and below the cut-off point are similar in all variables that should not be affected by the treatment. The idea is simply that if covariates are strongly correlated to the outcome variable and they are discontinuous at the cut-off, the continuity of the potential expected outcome functions is unlikely to hold, and thus the validity of the design is violated. To implement this formal falsification test, I run the "rdrobust codes of CCT" on each of the covariates in the study. The null hypothesis is that predetermined covariates are not affected by the treatment at the cut-off. Hence, the null hypothesis should not be rejected if the RD design is valid.

Although the graphical visualization for each covariate against the running variable can reveal if there are noticeable discontinuities at the cut-off point, quantitative statistical analysis is required before a more objective and formal conclusion regarding the continuity of the predetermined covariates can be reached. To implement the falsification tests analysis, the

optimal bandwidth must be chosen for each test separately. The chosen bandwidths will generally be different from the bandwidth that is used to run the RD estimation on the original outcomes. Since each covariate may have a different conditional mean function and different curvature, the optimal bandwidth for local polynomial estimation and inference will also be different for every variable; i.e., the statistical analysis is conducted separately for each covariate using the local polynomial estimators. Table 1.38 reports the local polynomial estimation and inference results for several different pre-determined covariates available in the study. The results show that all RD estimates are small, most of them close to zero, and all 95% confidence intervals contain zero, with p-values ranging from 0.13 to 0.867. Consequently, the null hypothesis that predetermined covariates are not affected by treatment (continuity assumption) is not rejected for all variables. As a result, the assumption of continuity is empirically validated. In other words, there is no empirical evidence that these predetermined covariates are discontinuous at the cut-off. Note that the number of observations used in the analysis varies for each covariate, which occurs because the MSE-optimal bandwidth is different for every covariate tested for continuity. Some studies in the RD setting argue that CER-optimal bandwidth is more appropriate than MSE-optimal bandwidth when running the falsification test. According to Cattaneo et al. (2018), the choice between these two alternative bandwidth selectors gives a natural trade-off between the size and power of the falsification tests. They argue that the MSE-optimal bandwidth leads to more powerful hypothesis testing with possibly larger size distortions than tests that are run using the CER-optimal bandwidth. However, when I implemented the falsification test using the CER-optimal bandwidth selector, there were no changes in the empirical conclusion for all predetermined covariates.

Table 1.38: Formal falsification test for the validity of the Continuity-Based Analysis for predetermined Covariates

			robust inference					
variables	MSE Optimal Bandwidth	RD Estimator	P-value	[Confidence	Interval]	Effective Observation	L|R	Observations
Females	2.8	0.01302	0.379	-0.01602	0.042056	10931|10109	56780|37351	94131
Married	3.7	0.02082	0.186	-0.00295	0.044593	15221|13381	5678037351	94131
Black	4.4	-0.01017	0.243	-0.02723	0.006897	18073 |15604	56780|37351	94131
Other races	3	-0.00328	0.574	-0.0147	0.008141	11926|10981	56780|37351	94131
Divorced, S, W	4.8	-0.01142	0.275	-0.03192	0.009079	19485|16712	56780|37351	94131
high school	4.4	-0.01533	0.192	-0.03835	0.007684	18073|15604	56780|37351	94131
university degree	3.5	0.01874	0.195	-0.00323	0.040706	13423 |12074	56780|37351	94131
Middle income	3	-0.0042	0.567	-0.01858	0.010182	11926|10981	56780|37351	94131
High income	5	0.00061	0.867	-0.00651	0.007725	20551|17539	56780|37351	94131
Spouse retired	3.6	0.02155	0.192	-0.0035	0.04659	14518|12846	56780|37351	94131
U.S birth	4.5	0.0100	0.379	-0.01229	0.032291	18440|15848	56780|37351	94131
Northwest	4.7	-0.00388	0.649	-0.02059	0.012836	19122|16390	56780|37351	94131
Midwest	3.7	-0.01146	0.29	-0.0327	0.009775	15221|13381	56780|37351	94131
South	3.8	0.01899	0.133	-0.00581	0.043792	15221|13381	56780|37351	94131
West	4.6	-0.0024	0.788	-0.01983	0.015034	18781|16098	56780|37351	94131
Others regions	4.5	-0.00096	0.271	-0.00268	0.000752	18781|16098	56780|37351	94131

Notes: RD estimated using local polynomial estimator with triangular kernel weights and the MSE-optimal bandwidth selector

1.8.2 Density of Running Variable Test

The continuity assumption of the outcome, as a function of the running variable x, rules out any discrete jump in the running variable at the cut-off point, so that any observed jump in outcomes can be attributed to the treatment effect. This assumption is more credible when the running variable cannot be manipulated by individuals near the threshold value, which, in this study, is age. That is, if there is a discontinuity in the density of the running variable at the cut-off point, it can be an indicator that individuals can precisely manipulate the running variable. Consequently, if the running variable is discontinuous at the cut-off, this may invalidate the continuity assumption in the RD setting.

To implement the manipulation test, the density of units in the neighborhood near the cutoff must be estimated to test the hypothesis that the density is discontinuous at the cut-off. There have been three widely documented manipulation tests in the literature. First, McCrary (2008) developed a test based on the nonparametric local-polynomial density estimator, which is an extension of the local linear density estimator of Cheng, Jianqing, and Marron (1997). The McCrary test is based on obtaining a histogram and then smoothing the histogram using local linear regression separately on either side of the cut-off. Second, Otsu, Xu, and Matsushita (2014) developed an empirical likelihood method using boundary-corrected kernels. Third, Cattaneo, Jansson, and Ma (2017b) developed a set of manipulation tests based on a novel local-polynomial density estimator. Their methods were shown to provide demonstrable improvements in both size and power under appropriate assumptions, relative to other approaches currently available in the literature (see Cattaneo and Escanciano, 2017; Cattaneo, Titiunik, and Vazquez-Bare, 2017c; Calonico, Cattaneo, and Titiunik, 2015a)

The hypothesis test of the continuity of the density of the running variable, $f(x)$, at the cut-off is based on estimating the density of observations near the cut-off separately for observations above and below the cut-off, using the local polynomial density estimator. The null hypothesis can be stated as

$$H_0: \lim_{x\uparrow\bar{x}} f(x) = \lim_{x\downarrow\bar{x}} f(x) \quad vs \quad H_1: \lim_{x\uparrow\bar{x}} f(x) \neq \lim_{x\downarrow\bar{x}} f(x)$$

Therefore, failing to reject the null hypothesis implies that there is no statistical evidence of manipulation at the cut-off point, and thus provides evidence supporting the validity of the RD design. The empirical test is implemented using the "rddensity" statistical package developed by CCT.[12] Table 1.39 reports the results from implementing the manipulation test. I run the test twice using the conventional RD estimator and the robust bias-corrected RD estimator. The test statistic is constructed using polynomial order $p = 2$ in both cases. The test statistic of the manipulation test according to the conventional RD setting is $T = -1.057$, with a $p - value$ of 0.2905, and the test statistics according to the robust RD estimator is $T = -0.0102$, with a $p - value$ of 0.9918. Therefore, in this application, the graphical visualisation[13] and the formal manipulation test indicate that there is no statistical evidence of systematic manipulation of the running variable.

Table 1.39: Manipulation test of the running variable

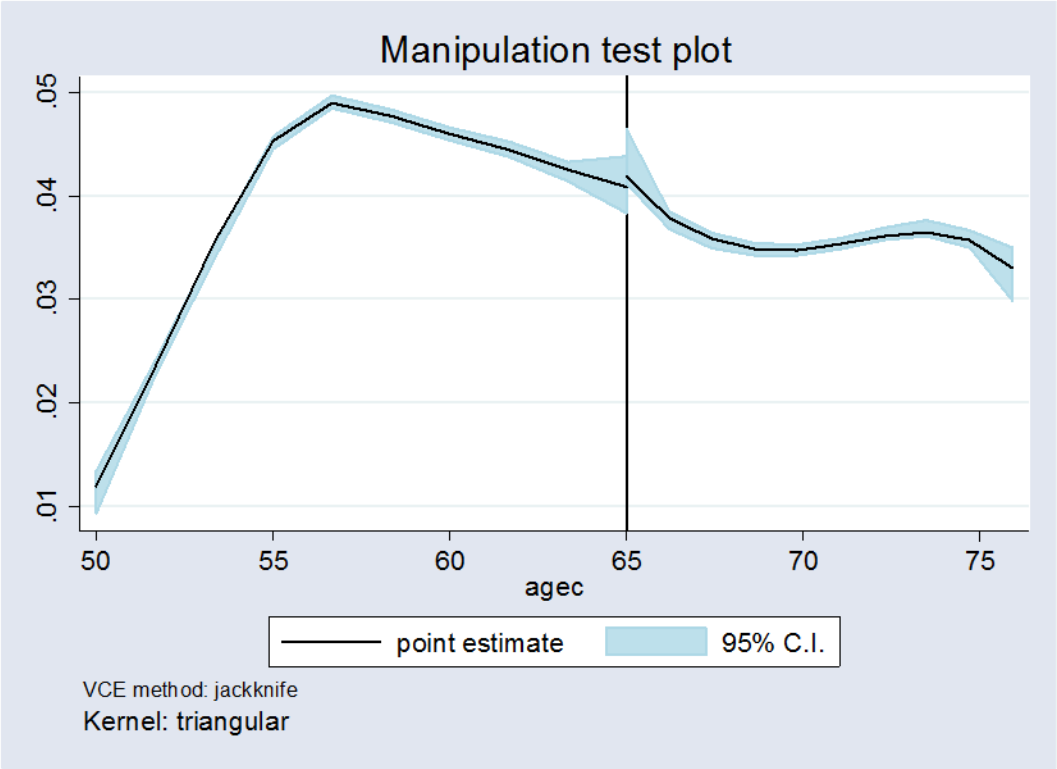
Method	test statistic	p-value
Conventional	-1.057	0.2905
Robust	-0.0102	0.9918

Notes: The RD estimates are based on the conventional RD estimator and the Robust biased-corrected RD estimator. (ii) the test is constructed by using "lpdensity" statistical package.

[12] https://sites.google.com/site/rdpackages/rddensity

[13] This plot is constructed using the statistical package "lpdensity" for local polynomial-based density estimation for manipulation test. The package is developed by CCT at https://sites.google.com/site/rdpackages/rddensity

Figure 1.21: Running variable density test



1.8.3 Sensitivity to Observations Near the Cut-off

The idea behind this falsification test is to exclude observations that are very close to the cut-off point to investigate how sensitive the results are to this exclusion. To implement this test, after excluding the units close to the cut-off, the RD estimation and inference analysis are repeated using the remaining sample. This procedure, called a “donut hole" approach, is very useful to assess the sensitivity of the RD results to observations closest to the cut-off point, which may have the most influence on the estimated RD results when estimating the local polynomials, the unavoidable extrapolation[14] (see Cattaneo, Idrobo, and Titiunik, 2018).

[14] when the few observations in the near neighborhood of the cutoff are likely to have the most influential on the fitted local polynomials.

Table 1.40 reports the RD estimates after excluding some observations around the cut-off. The results show that excluding observations in the 0.10 radius of the cut-off lead to similar results as the original analysis, zero radius exclusion. The exclusion of these observations changes the RD estimate for cognitive functioning from -0.86065 to -0.81731, with the robust confidence interval changing from [-1.12260, -0.598694] to [-1.16172, - 0.47289]. Similarly, excluding 0.10 of the observations changed the RD estimates for mental health and self-reported health from 0.4142 to 0.36392 and from -0.13172 to -0.12299, respectively.

Table 1.40 shows the results from the Donut hole approach by repeating this procedure several times to assess the actual sensitivity for different radiuses of excluded units in the neighborhood of the cut-off, and the figure helps visualize the results graphically. The results from the RD estimates at different exclusion radiuses are unchanged. Moreover, all the new RD estimates and the original estimates are still significant at 1%. Consequently, the fuzzy RD estimates are robust for the exclusion of some observations around the cut-off point for all health outcome variables at a 1% significance level.

1.8.4 Sensitivity to Bandwidth Choice

This falsification test is related to the sensitivity of the RD estimates to bandwidth choice. Intuitively, the Donut hole test assesses the sensitivity of the RD estimates to removing units from the center of the neighborhood of the cut-off point, while the bandwidth sensitivity test assesses the sensitivity of the results to removing or adding units to the endpoints, window, of the neighborhood. To implement this approach empirically, I reported the RD estimates at different bandwidths around the optimal-MSE bandwidth for all outcome variables in the study. When reviewing the RD estimates for all outcome variables, the results were broadly consistent

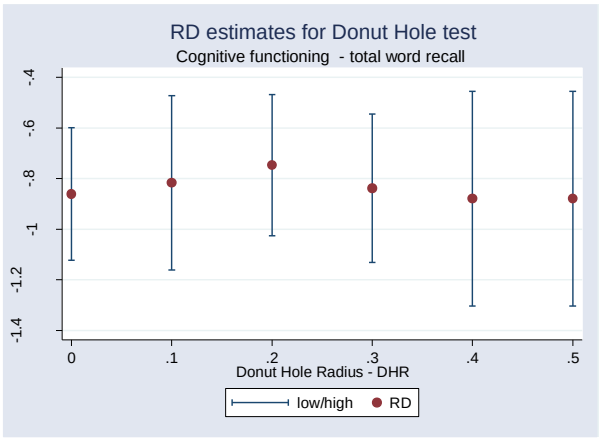
with the empirical findings obtained with either the MSE-optimal bandwidth selector and CV-optimal bandwidth selector.

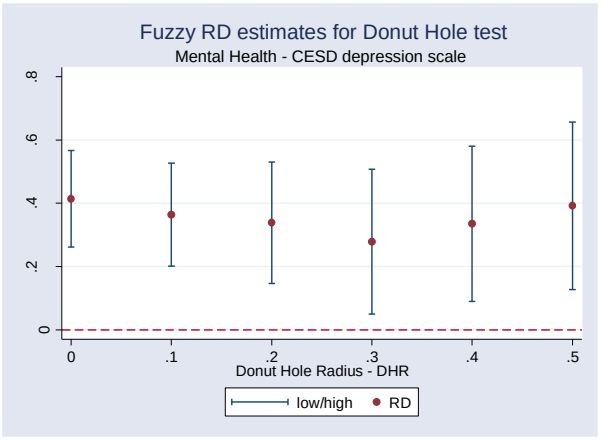
Table 1.40: Fuzzy RD sensitivity test to observations near cutoff: The Donut-Hole Approach

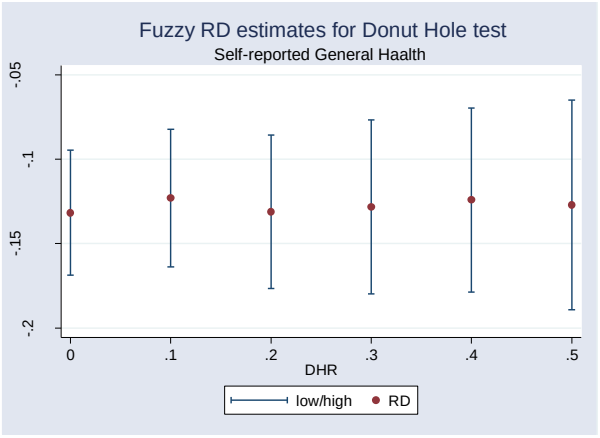
Donut-Hole Radius	MSE-Optimal bandwidth	RD estimates	Robust inference		Excluded Observation	
			P-value	[95% Conf. intervals]	Left	Right
Panel A: Cognitive Functioning						
0.0	4.8	-0.86065	0.000	[-1.12260, -0.598694]	0	0
0.1	3.5	-0.81731	0.000	[-1.16172, - 0.47289]	117	215
0.2	5.2	-0.74733	0.000	[-1.02606, -0.46861]	145	221
0.3	5	-0.83853	0.000	[-1.13215, -0.54492]	140	221
0.4	3.5	-0.87952	0.000	[-1.30431, -0.45473]	117	229
0.5	3.5	-0.87952	0.000	[-1.30431, -0.45473]	358	399
Panel B: Mental Health						
0.0	3.9	0.41420	0.000	[0.26139, 0.567016]	0	0
0.1	3.8	0.36392	0.000	[0.201555, 0.526289]	117	215
0.2	3.4	0.33869	0.001	[0.146818, 0.530553]	145	221
0.3	3.1	0.27865	0.007	[0.049373, 0.507931]	140	221
0.4	3.1	0.33535	0.007	[0.089778, 0.580916]	117	229
0.5	3.3	0.39211	0.004	[0.127380, 0.656836]	358	399
Panel C: Self-reported General Health						
0.0	3.2	-0.13172	0.000	[-0.168785, -0.09466]	0	0
0.1	3.1	-0.12299	0.000	[-0.163779, -0.08220]	117	215
0.2	3	-0.13106	0.000	[-0.176604, -0.08552]	145	221
0.3	2.9	-0.12828	0.000	[-0.179902, -0.07667]	140	221
0.4	3	-0.12414	0.000	[-0.178681, -0.06961]	117	229
0.5	3.1	-0.12707	0.000	[-0.189147, -0.06499]	358	399

Notes: (i) All RD estimates are computed using a triangular kernel and MSE-optimal bandwidth selector, (ii) first column report the radius around the cutoff point.

Figure 1.22: RD Estimation for the Donut-Hole Approach







1.8.5 Treatment Effect Derivative (TED) and Complier Probability Derivative (CPD) Test

Recent work on RD design has developed a new sophisticated test of the Stability of the Fuzzy Regression Discontinuity Model. Although the treatment effect estimated from RDD has been widely documented to have high internal validity (Wing and Bello- Gomez, 2018), there are two issues in RD design that can affect the external validity of the estimates, which require more assessment. The first issue of the external validity of RD arises from estimating the treatment effect from the subpopulation of individuals close to the cut-off point. Since RDD estimates the treatment effect at the cut-off point, the LATE at the cut-off may be different from the treatment effect at other values of the running variables, which means that the estimated RD treatment effect may not be precisely estimated to reflect the overall effect of the treatment. The second problem of the external validity of the fuzzy RD arises from estimating treatment exposure among a subpopulation of compliers.

Therefore, the logical question that should be addressed is: if the cutoff value is given by c, would individuals with $x \neq c$ but have x near c experience a similar treatment effect to those having $x = c$? I perform two tests to see what will happen to the estimated RD treatment effect when the value of the running value is a little bit further away from the cut-off point, i.e., $\frac{\partial \tau}{\partial x}$. The theoretical issue when empirically trying to estimate the value of $\frac{\partial \tau}{\partial x}$ is that the conditional mean function, $\lim_{x \downarrow \bar{x}} E[Y(1) \mid X_i = x]$ can be empirically estimated but its counterfactual segment, $\lim_{x \uparrow \bar{x}} E[Y(1) \mid X_i = x]$, is always missing. By the same logic, the conditional mean function, $\lim_{x \uparrow \bar{x}} E[Y(0) \mid X_i = x]$, can be estimated but its counterfactual segment, $\lim_{x \downarrow \bar{x}} E[Y(0) \mid X_i = x]$, is always missing. Therefore, the difference between $lim\, E[Y(1) \mid X_i = x]$ and $\lim E[Y(0) \mid X_i = x]$ is unknown at any point away from the cut-off

point, c , because one of the counterfactual segments of the conditional mean functions is always missing. Fortunately, Dong and Lewbel (2015) showed that derivatives of the treatment effect of the RD model with respect to the running variable, $\frac{\partial \tau}{\partial x}$, at the cut-off point, known as the Treatment Effect Derivative (TED), can be non-parametrically identified and can be used to test the external validity of the RD estimates and hence extrapolating the estimated LATE for subpopulation away from the cut-off point.

The main intuition behind the TED test is to see how the treatment effect would change in the neighborhood of the cut-off point. That is, if the TED is significantly equal to zero, then the treatment effect is locally constant, and the RD estimates have strong external validity (Dong and Lewbel, 2015). On the other hand, if $\tau'(c) = \frac{\partial \tau(c)}{\partial x}$ is large and significantly different from zero, then a slight change in the running variable would be associated with a significant change in the treatment effect and the external validity of the RD estimates should be addressed.

Following Dong and Lewbel (2015) and Giovanni et al., (2016), recall $\tau(c)$ to be the RD local average treatment effect (LATE) at $x = c$ among compliers, i.e., the average difference in health outcomes across individuals who randomly switched from working to retirement at the official age.

Definition 1: let $g(x) = \mathbb{E}[Y|X = x]$. for small $\varepsilon > 0$, define right and left limits of the conditional mean function $g(\cdot)$, as $g_+(x) = \lim_{\varepsilon \to 0} g(x + \varepsilon)$ and $g_-(x) = \lim_{\varepsilon \to 0} g(x - \varepsilon)$, respectively.

Definition 2: let $f\ (x) = E\ [\ T|X = x\ ]$ is the probability of treatment (retired) given $X = x$. By analogy to 1, for small $\varepsilon > 0$ define right and left limits of the conditional probability function of being treated $p(\cdot)$ as $f_+(x) = \lim_{\varepsilon \to 0} f(x + \varepsilon)$ and $f_-(x) = \lim_{\varepsilon \to 0} f(x - \varepsilon)$, respectively.

According to TED identification, the derivative $\tau'(x) = \frac{\partial \tau(x)}{\partial x}$ can be empirically estimated given one-sided derivatives. Holding the threshold fixed at $x = c$, for small $\varepsilon > 0$, define the right and left derivatives of functions $g(x)$ and $f(x)$ at point x as

$$g'_+(x) = \lim_{\varepsilon \to 0} \frac{g(x+\varepsilon) - g(x)}{\varepsilon} \quad and \; g'_-(x) = \lim_{\varepsilon \to 0} \frac{g(x) - g(x-\varepsilon)}{\varepsilon}$$

$$f'_+(x) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon) - f(x)}{\varepsilon} \quad and \; f'_-(x) = \lim_{\varepsilon \to 0} \frac{f(x) - f(x-\varepsilon)}{\varepsilon}$$

Local polynomial regressions can be used to estimate both $g(\cdot)$ and $f(\cdot)$ and their derivatives separately on each side of the threshold.

Let $q(x) = E[Y(1)|X = x] - E[Y(0)|X = x]$, then $q(c) = g_+(c) - g_-(c)$. Analogously, let $p(x)$ denote the conditional probability that someone is a complier conditional on $x = c$. $p(c) = f_+(c) - f_-(c)$. Recall the standard fuzzy RD estimator, τ at point $x = c$ is identified by

$$\tau(c) = \frac{g_+(c) - g_-(c)}{f_+(c) - f_-(c)} = \frac{q(c)}{p(c)} \quad (21)$$

if the previous assumptions of (HTV, 2001) are held, then the fuzzy treatment effect, $\tau(c)$, is identified by equation 21 and fuzzy RD TED, $\tau'(c)$, is identified by

$$\tau'(c) = \frac{g'_+(c) - g'_-(c) - [f'_+(c) - f'_-(c)] \cdot \tau(c)}{f_+(c) - f_-(c)} \quad (22)$$

$$\tau'(c) = \frac{g'_+(c) - g'_-(c)}{p(c)} - \frac{p'(c) \cdot \tau(c)}{p(c)} \quad (23)^{15}$$

This equation implies that if the compliance rate is locally constant, $p'(c) = 0$, the fuzzy RD TED will equal the first term only. However, in fuzzy RD, the compliance rate $p(c)$ is not

[15] Given that $\tau(c) = \frac{q(c)}{p(c)}$, then $\tau'(x) = \frac{\partial \tau(x)}{\partial x} = \frac{q'(x)p(x) - q(x)p'(x)}{p(x)^2}$ or $\tau'(x) = \frac{q'(x)}{p(x)} - \frac{q(x)\,p'(x)}{p(x)^2} = \frac{q'(x)}{p(x)} - \frac{p'(x)\,\tau(x)}{p(x)}$

constant and hence the second term in equation (23) will be proportional to $p'(c)$ and account for the effect of a change in the fraction of the population who are compliers at $x = c$ when we move away from the threshold point (c).

The equation also shows that fuzzy RD has two potential sources of instability. First, fuzzy RD treatment effect can be unstable because $q'(c)$ could be far from zero which means that the treatment effect for the average compliers will significantly change when the running variable (x) moves a little bit away from the threshold (c). The second source of fuzzy instability may arise because $p'(c)$ is far from zero, which means that the population of compliers will change dramatically as the running variable x moves away from the threshold (c).

Tables 1.41 and 1.42 report both the CPD and TED estimates. I compare estimates from CPD and TED based on the common two bandwidth selectors: the CCT (Calonico, Cattaneo, and Titiunik, 2014), and the CV bandwidth selector (Ludwig and Miller, 2007). Also, the Kernel triangular function is used to estimate CPD and TED under the two bandwidth selectors where it is widely documented to be optimal for estimating the conditional mean function at a boundary point (Fan and Gijbels, 1996). The local quadratic regression is estimated separately for the reduced form outcome and treatment conditional mean functions and then constructed TED from the estimated intercepts and slopes in the four conditional mean functions using 200 simulations to bootstrap the standard errors.

Table 1.41: TED and CPD of Fuzzy RD Treatment Effects of Retirement on Health

		CCT			CV		
		All sample	Females	Males	All Sample	Females	Males
		Panel A: Cognitive functioning					
Cognitive functioning	CPD	-0.0084	-0.0185	0.0045	-0.0125	-0.0199	-0.0009
		(0.0093)	(0.0124)	(0.0141)	(0.0085)	(0.0118)	(0.0128)
	TED	-0.00883	-0.00226	-0.00449	-0.00692	-0.00945	-0.00486
		(0.0232)	(0.085)	(0.0212)	(0.0241)	(0.1696)	(0.0205)
		Panel B: Mental health					
Mental Health	CPD	-0.0093	-0.0165	0.0004	-0.0055	-0.0147	0.0061
		(0.0088)	(0.0117)	(0.0134)	(0.0108)	(0.0130)	(0.0164)
	TED	0.0069	-0.0433	0.0114	0.0198	-0.0508	0.0183
		(0.0227)	(0.2257)	(0.0171)	(0.0309)	(0.3433)	(0.0231)
		Panel C: Self-reported health					
Self-Reported health	CPD	-0.0078	-0.0156	0.0026	-0.0035	0.0239	-0.0045
		(0.0093)	(0.0123)	(0.0141)	(0.0122)	(0.0291)	(0.0397)
	TED	0.0322	0.0600	-0.0161	0.0004	-0.0146	-0.0115
		(0.0945)	(0.2078)	(0.0458)	(0.1510)	(0.0822)	(0.1409)

Note: All estimates are based on local quadratic regressions; CCT refers to the optimal bandwidth for local polynomial regression discontinuity (RD) point estimators and inference procedures developed in proposed by Calonico, Cattaneo, and Titiunik (2014) and Calonico et al. (2016); CV refers to the cross-validation optimal bandwidth proposed by Ludwig and Miller (2007). CPD and TED estimates are based on the triangular kernel, and Bootstrapped Standard errors based on 200 simulations are in parentheses*,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.42: TED and CPD of Fuzzy RD Treatment Effects of Retirement on Health-related behavior

		CCT			CV		
		All sample	Females	Males	All sample	Females	Males
		Panel A: Drinking participation					
Drinking participation	CPD	-0.0091	-0.0164	0.0006	-0.0034	-0.0178	-0.0009
		(0.0088)	(0.0117)	(0.0134)	(0.0122)	(0.0111)	(0.0128)
	TED	-0.0603	0.0141	0.0112	-0.1018	0.0098	-0.0049
		(0.6104)	(0.1072)	(0.6956)	(0.9752)	(0.0318)	(0.0550)
		Panel B: Drinking intensity					
Drinking intensity	CPD	-0.0093	-0.0165	0.0004	-0.0055	-0.0147	0.0061
		(0.0088)	(0.0117)	(0.0134)	(0.0108)	(0.0130)	(0.0164)
	TED	-0.0221	0.0094	-0.0463	-0.0363	0.0095	-0.0464
		(0.0427)	(0.0240)	(0.0517)	(0.0560)	(0.0283)	(0.0634)
		Smoking participation					
Smoking intensity	CPD	-0.0202***	-0.0149	0.0019	0.0120	0.0112	-0.0058
		(0.0070)	(0.0124)	(0.0141)	(0.0219)	(0.0248)	(0.0331)
	TED	-0.0622	0.0047	-0.5050	0.0255	-0.0807	-0.0075
		(0.0680)	(0.0322)	(0.5727)	(0.1800)	(0.0893)	(0.0606)

Note: All estimates are based on local quadratic regressions; CCT refers to the optimal bandwidth for local polynomial regression discontinuity (RD) point estimators and inference procedures developed in proposed by Calonico, Cattaneo, and Titiunik (2014) and Calonico et al. (2016); CV refers to the cross-validation optimal bandwidth proposed by Ludwig and Miller (2007). CPD and TED estimates are based on the triangular kernel, and Bootstrapped Standard errors based on 200 simulations are in parentheses. *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Panel A in Table 1.41 reports the CPD and TED estimates for the cognitive functioning of retirees. CPD ranges from -0.0009 to -0.019, which are small and statistically insignificant. The estimates show that moving away from the threshold point by one point, the percent of individuals who are compliers will decline by somewhere between 0.09% and 2%. In addition, TED estimates are small and not statistically significant.

Put together, these results indicate that the set of compliers is stable and the TED is numerically small and statistically insignificant. The near-zero TED effect suggests that the effect of retirement on cognitive functioning would likely remain valid at lower or higher values of the running variable. In other words, the TED estimates suggest that marginally raising or

lowering the running variable would not result in a significant change in the estimated effects of retirement on cognitive function. To check how the CPD and TED are sensitive to the optimal bandwidth choice selectors, I estimated the CPD and TED using CCT and CV and the results show that CPD and TED at varying bandwidth selector are almost all near zero, i.e., they are statistically insignificant.

Similarly, the estimates of CPD and TED for mental health and self-reported health show that they are small and statistically insignificant. The results of CPD and TED for drinking participation and intensity are insignificant. However, the CPD for smoking participation is significant and equals -0.020. This estimate suggests that moving away from the cut-off point by one unit decreases the percent of individuals who are compliers by 2 percentage points. In contrast, the TED estimate is statistically insignificant. Together, the results indicate that although the set of compliers is not stable for RD estimation of smoking participation, the estimated RD retirement effect on smoking participation is stable.

Finally, having a relatively small and insignificant TED in magnitude, or TED equal zero, increases the power of the external validity of the RD LATE estimate, since small changes in the running variable are associated with no significant change in the average treatment effect of retirement on health, as measured by cognitive functioning, mental health, and self-reported health, respectively, and health-related behavior as measured by drinking and smoking.

1.9. Parametric Estimation of the Fuzzy Regression Discontinuity Design (FRD)

The non-parametric procedure tries to find the optimal data range around the cut-off point to estimate the local polynomial function that generates consistent estimates, while the parametric approach tries to pick up the right information from observations that are far away from the cut-off point to estimate the average outcome for observations near the cut-off (Jacob and Zhu, 2012). That is, non-parametric methods limit the analysis to a subpopulation of observations that are close to the cut-off and reduce the bias in the estimation. Therefore, the parametric estimation can be introduced in this part as a sensitivity check for using more information from the observation far away from the neighborhood of the cut-off point.

The validity of the IV in the RD setting depends heavily on whether there is a significant jump in the chance of retirement at age 65. Consequently, the validity of the instrument variable will be verified by checking whether there is a discontinuity in the treatment variable at the cut-off in the graphs and first-stage F statistics in our IV/2SLS regressions. Figure 1.1 uses the RD non-parametric local polynomial estimator for the age-retirement profile with 95% confidence intervals for the whole sample. The graph confirms a positive and significant jump in the retirement decision at the official age of eligibility for the social security fund. The magnitude of the discontinuity is noticeable at ages 65 and 62, indicating individuals are far more likely to retire when they reach the ages of 62 and 65. However, whether the jump is due to a social norm in the economy or due to financial incentives by the eligibility to access a pension fund will not affect the analysis. To test the validity of the instrument more formally, the F-statistics at the first stage of the 2SLS analysis is investigated and the results show that the instrument is not weak as all the first stage F statistics are bigger than 10 for the full sample (Stock and Yogo, 2002).

Table 1.43: Fuzzy parametric RD estimates – Health outcomes at cutoff age 65 (All sample)

	Cognitive Functioning Total word recall		Mental health CESD scale		General Health Self-reported health	
Retired	-0.5090***	(0.0288)	0.5661***	(0.0164)	-0.1956***	(0.0036)
(Age-65)	-0.0435***	(0.0110)	-0.0775***	(0.0063)	0.0153***	(0.0014)
(Age-65)^sq	-0.0024***	(0.0008)	-0.0030***	(0.0004)	0.0008***	(0.0001)
T*(age-65)	-0.0632***	(0.0235)	0.0369***	(0.0134)	-0.0074**	(0.0030)
T*(age-65)^sq	-0.0004	(0.0014)	0.0051***	(0.0008)	-0.0014***	(0.0002)
Female	1.3559***	(0.0219)	0.1638***	(0.0125)	0.0212***	(0.0027)
Marital Status (reference single)						
Married	0.3729***	(0.0591)	-0.3026***	(0.0331)	0.0303***	(0.0073)
Others (S,D,W)	0.2562***	(0.0547)	0.1225***	(0.0306)	0.0045	(0.0067)
Races (reference White)						
Black	-1.2200***	(0.0300)	0.0953***	(0.0169)	-0.0552***	(0.0037)
Others	-0.9818***	(0.0482)	0.2738***	(0.0267)	-0.0825***	(0.0059)
Education (reference less than high school)						
High school	1.4479***	(0.0272)	-0.5450***	(0.0156)	0.1725***	(0.0034)
University	2.5537***	(0.0335)	-0.8448***	(0.0191)	0.2474***	(0.0042)
Middle income	0.2756***	(0.0367)	-0.2263***	(0.0208)	0.0507***	(0.0046)
High income	0.4921***	(0.0600)	-0.2267***	(0.0332)	0.0434***	(0.0073)
Spouse retired	-0.0496*	(0.0269)	-0.0405***	(0.0154)	0.0056*	(0.0034)
Census (Northwest reference)						
Midwest	-0.1241***	(0.0338)	-0.0546***	(0.0194)	0.0091**	(0.0043)
South	-0.1545***	(0.0308)	0.0814***	(0.0176)	-0.0239***	(0.0039)
West	0.0099	(0.0358)	-0.0116	(0.0204)	-0.0097**	(0.0045)
Others	0.2457	(0.3349)	0.0868	(0.1844)	-0.0661	(0.0406)
US birth	0.1950***	(0.0342)	-0.2086***	(0.0193)	0.0554***	(0.0042)
Children no	-0.0129**	(0.0054)	0.0157***	(0.0031)	-0.0049***	(0.0007)
Constant	8.7655***	(0.0705)	1.5846***	(0.0399)		

Notes: (i) standard errors are in parentheses. (ii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.43 reports the estimated treatment effect of retirement for the three health outcomes using the Fuzzy RD parametric analysis and the cross-validation optimal bandwidth. After controlling for other covariates and confounders, the table shows that retirement has a strong negative impact on health outcomes.

The first column reports the RD estimates of the impact of retirement on the cognitive function. The results suggest that retirement has a negative effect on cognitive functioning by an approximately 0.5 point decrease in total word recall, which is equivalent to almost 5 percentage points decrease in cognitive function of retirees relative to non-retirees. The impact of other covariates is consistent with prior studies. Black and other races have significantly lower cognitive functioning scores compared to whites. Females experience an approximately 1.3 point increase in their cognitive scale compared to males, i.e., females, on average, have higher cognitive functioning than males in the HRS dataset. The socioeconomic status has a significant effect on cognitive functioning. Individuals who have at least a high school degree or a university degree or above have 1.4 and 2.5 points higher cognitive score than the cognitive score of individuals with less than a high school degree, which supports the causal relationship between cognition and education. Income level also has a significant impact on cognitive functioning. The cognitive scores of individuals in the middle- and high-income groups are almost 0.4 and 0.7 points higher, respectively than individuals in the low-income group.

The second column in table 1.43 shows that retirement is associated with a 0.56 point increase in the CESD depression scale, i.e., retirement has a negative impact on retirees' mental health. The results also show that females' CESD depression scale is 0.16 points higher than that of males. That is, females are more likely to experience negative mental health than males. Socioeconomic status has a significant effect on the depression scale. Married individuals'

CESD depression scale is 0.3 points less than non-married individuals and this result is in line with findings from empirical research in the happiness literature that shows that married individuals are happier than unmarried individuals (Helliwell, 2003; Blanchflower and Oswald, 2008). The depression scale of individuals with at least high school or university education is almost 0.5 to 0.8 points less, respectively, than that of individuals who have less than a high school degree. Individuals in the medium and high-income groups score about 0.2 points lower in the depression scale than those in the low-income group. Also, the results show that the partner's retirement status has a small significant effect on the depression score of individuals. Persons whose partners are retired experience an approximately 0.04 point lower score in their CESD depression scale relative to persons whose partners are still in the labor force.

The third column in table 1.43 shows that retirees are 19 percentage points less likely to report excellent, very good, and good health than non-retirees. Controlling for demographic characteristics, the results show that married individuals are 3 percentage points more likely to report good health than unmarried individuals. Females are 2 percentage points more likely to be in good health relative to males. Blacks and other races are 5 and 8 percentage points less likely to be in good health relative to whites. Controlling for the socioeconomic status of individuals, we can see that individuals with high educational attainment have good health status. For instance, individuals with at least high school and university and postgraduate degrees are 17 and 24 percentage points more likely to be in good health relative to individuals with less than high school. This result is consistent with the empirical literature that shows that well-educated individuals invest more in their health (Grossman and Kaestner, 1997). In addition, individuals in middle and high-income groups are, respectively, 5 and 4 percentage points more likely to report higher levels of health satisfaction relative to individuals in low-income groups. Ettner (1996)

argues that income is one of the channels through which retirement can impact health, where health is a normal good that affects health. The results show that the partner's retirement status has a small positive effect on an individual's own physical health status and this is consistent with existing literature that provides evidence on intra-household retirement externalities.

Table 1.44: Fuzzy parametric RD estimates – Health outcomes at official cutoff age 65 (females only)

	Cognitive Functioning Total word recall		Mental health CESD scale		General Health Self reported health	
Retired	-0.5359***	(0.0389)	0.6050***	(0.0231)	-0.2017***	(0.0048)
Age-65	-0.0247*	(0.0148)	-0.0789***	(0.0088)	0.0175***	(0.0018)
(Age-65)^sq	-0.0016	(0.0010)	-0.0028***	(0.0006)	0.0009***	(0.0001)
T*(age-65)	-0.0890***	(0.0319)	0.0341*	(0.0190)	-0.0082**	(0.0039)
T*(age-65)^sq	-0.0014	(0.0020)	0.0055***	(0.0012)	-0.0016***	(0.0002)
Marital Status (reference single)						
Married	0.2411***	(0.0821)	-0.2897***	(0.0477)	0.0196**	(0.0099)
Others (S,D,W)	0.2129***	(0.0722)	0.0928**	(0.0420)	-0.0021	(0.0087)
Races (reference White)						
Black	-1.2689***	(0.0391)	0.0797***	(0.0230)	-0.0620***	(0.0048)
Others	-1.1395***	(0.0673)	0.3264***	(0.0387)	-0.1009***	(0.0080)
Education (reference less than high school)						
High school	1.4472***	(0.0372)	-0.6289***	(0.0222)	0.1956***	(0.0046)
University	2.5010***	(0.0475)	-0.9689***	(0.0283)	0.2723***	(0.0059)
Middle income	0.1398**	(0.0559)	-0.2386***	(0.0327)	0.0446***	(0.0068)
High income	0.0984	(0.1070)	-0.2364***	(0.0610)	0.0328***	(0.0126)
Spouse retired	-0.1208***	(0.0390)	0.0397*	(0.0232)	0.0055	(0.0048)
Census (Northwest reference)						
Midwest	0.0019	(0.0456)	-0.0815***	(0.0273)	0.0140**	(0.0056)
South	-0.0491	(0.0415)	0.0382	(0.0247)	-0.0145***	(0.0051)
West	0.1219**	(0.0487)	-0.0660**	(0.0289)	-0.0036	(0.0060)
Others	-0.3822	(0.5057)	0.0765	(0.2819)	-0.1263**	(0.0583)
US birth	0.3063***	(0.0513)	-0.2787***	(0.0300)	0.0720***	(0.0062)
Children no	-0.0100	(0.0073)	0.0115***	(0.0044)	-0.0035***	(0.0009)
Constant	10.2017***	(0.0946)	1.8605***	(0.0557)		
Obs	52,083		52,083		52,083	

Notes: (i) standard errors are in parentheses. (ii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.45: Fuzzy parametric RD estimates – Health outcomes at cutoff age 65 (males only)

	Cognitive Functioning Total word recall		Mental health CESD scale		General Health Self-reported health	
retired	-0.4586***	(0.0429)	0.5074***	(0.0230)	-0.1885***	(0.0055)
Age-65	-0.0642***	(0.0166)	-0.0765***	(0.0089)	0.0132***	(0.0021)
(Age-65)^sq	-0.0029**	(0.0012)	-0.0031***	(0.0007)	0.0007***	(0.0002)
T*(age-65)	-0.0352	(0.0348)	0.0435**	(0.0187)	-0.0071	(0.0045)
T*(age-65)^sq	0.0006	(0.0021)	0.0044***	(0.0011)	-0.0011***	(0.0003)
Marital Status (reference single)						
Married	0.4703***	(0.0861)	-0.2940***	(0.0453)	0.0387***	(0.0109)
Others (S,D,W)	0.2828***	(0.0846)	0.1703***	(0.0445)	0.0197*	(0.0107)
Races (reference White)						
Black	-1.1343***	(0.0471)	0.1195***	(0.0249)	-0.0432***	(0.0060)
Others	-0.8050***	(0.0689)	0.2179***	(0.0358)	-0.0625***	(0.0086)
Education (reference less than high school)						
High school	1.4349***	(0.0398)	-0.4467***	(0.0214)	0.1429***	(0.0052)
University	2.6173***	(0.0471)	-0.7062***	(0.0253)	0.2212***	(0.0061)
Middle income	0.3972***	(0.0489)	-0.2351***	(0.0262)	0.0584***	(0.0063)
High income	0.6741***	(0.0725)	-0.2632***	(0.0379)	0.0544***	(0.0091)
Spouse retired	0.0109	(0.0371)	-0.1092***	(0.0199)	0.0039	(0.0048)
Census (Northwest reference)						
Midwest	-0.2985***	(0.0502)	-0.0170	(0.0271)	0.0024	(0.0065)
South	-0.2991***	(0.0461)	0.1381***	(0.0248)	-0.0361***	(0.0060)
West	-0.1361***	(0.0527)	0.0570**	(0.0282)	-0.0178***	(0.0068)
Others	0.6278	(0.4434)	0.1135	(0.2346)	-0.0161	(0.0565)
US birth	0.1244***	(0.0460)	-0.1582***	(0.0244)	0.0391***	(0.0059)
Children no	-0.0187**	(0.0079)	0.0191***	(0.0042)	-0.0065***	(0.0010)
Constant	8.6911***	(0.1024)	1.4532***	(0.0544)		
Obs	36,727		40,592		40,592	

Notes: (i) standard errors are in parentheses. (ii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

To check if there is gender heterogeneity in the effect of retirement on health, the fuzzy parametric RD regression is estimated for each gender separately. Tables 1.44 and 1.45 indicate that there is a small difference in the cognitive score between males and females. Retired females and retired males have an almost 0.53 and 0.45 point lower cognitive score, respectively, relative

to non-retired females and non-retired males. The results show that retired females have a higher increase in the depression score compared to retired males. Retired females experience a 0.6 point increase in the CESD depression score relative to non-retired females, while retired males experience a 0.5 point increase in the depression scale relative to non-retired males. Similarly, the RD results indicate a small difference between retired male and females in self-reported general health. Retired females are 20 percentage points less likely to report being in good health relative to non-retired females, while retired males are 18 percentage points less likely to report being in good health relative to non-retired males.

Table 1.46 shows the results of the causal effect of retirement on health-related behavior. Smoking and drinking are widely known to be the causes of many chronic health problems (Sturm, 2002). The parametric RD results indicate that transition from working to retirement makes retired individuals less likely to drink but more likely to smoke. However, retirement has a quantitative effect on drinking intensity less than drinking participation. For instance, retired individuals are 3 percentage points less likely to be drinkers relative to non-retirees while they are almost 3 percentage point more likely to be a smoker relative to non-retirees.

The estimation results show that retirement has a negative effect on drinking participation but a small negative impact on the frequency of drinking during the week for those who continue to drink. For instance, retired individuals are 3 percentage points less likely to drink relative to non-retirees and they are 0.7 percentage points less likely to have a drink more than three times per week, for those who continue to drink. Females are not only less likely to drink but they are less likely to have more than three drinks per day for those who continue to drink. For instance, females are 11 percentage points less likely to be a drinker after retirement relative to males and they are 12 percentage points less likely to have more than three drinks per day. Educational

attainment has a positive effect on both drink participation and frequency. Individuals who have university and postgraduate degrees are 11 percentage points less likely to be a drinker, and 20 percentage points less likely to have more than three drinks per week. Therefore, educational attainment has a significant quantitative impact on drinking and hence health. The results also show that individuals with a retired partner are 0.8 percentage points less likely to be a drinker but this has no significant effect on their drinking habits or intensity. The positive impact of retirement on health-related behavior can be interpreted as a change in lifestyle after retirement where release from peer effects (such as drinking with coworkers or the boss as a part of work) and job-related stress may decrease drinking.

The parametric RD results indicate that retired individuals are 3 percentage points less likely to smoke compared to non-retirees. Controlling for demographic characteristics, we can see that females' smoking participation is lower than males. For instance, females are almost 5 percentage points less likely to be smokers relative to males. Married individuals are 12 percentage points less likely to smoke than unmarried individuals. Controlling for socioeconomic status, we can see that educational attainment has a significant negative effect on smoking participation. Individuals who have at least high school and university and postgraduate degrees are 6 and 14 percentage points, respectively, less likely to be smokers relative to individuals who have less than a high school degree.

Tables 1.46 and 1.47 show the results for males and females separately. The results indicate a small difference between retired females and retired males in drinking and smoking participation: retired females are 4 and 3 percentage points less likely to drinkers or smokers, respectively, relative to non-retired females, while retired males are 2 percentage points less likely to be drinkers or smokers than non-retired males.

Table 1.46: Fuzzy parametric RD estimates – Health behavior at cutoff 65 (All sample)

	Drinking participation		Drinking at least 3 time per week		Smoking participation	
Retired	-0.0356***	(0.0043)	0.0069***	(0.0026)	0.0317***	(0.0033)
Age-65	-0.0031*	(0.0016)	-0.0051***	(0.0010)	-0.0132***	(0.0013)
(Age-65)^sq	0.0004***	(0.0001)	0.0001	(0.0001)	-0.0004***	(0.0001)
T* (age-65)	0.0050	(0.0035)	0.0014	(0.0021)	0.0027	(0.0027)
T*(age-65)^sq	-0.0009***	(0.0002)	-0.0001	(0.0001)	0.0004**	(0.0002)
Female	-0.1149***	(0.0032)	-0.1293***	(0.0020)	-0.0511***	(0.0025)
Marital Status (reference single)						
Married	-0.0181**	(0.0086)	-0.0439***	(0.0052)	-0.1239***	(0.0066)
Others (S,D,W)	0.0084	(0.0080)	0.0119**	(0.0048)	0.0256***	(0.0061)
Races (reference White)						
Black	-0.1040***	(0.0044)	-0.0206***	(0.0027)	-0.0143***	(0.0034)
Others	-0.0860***	(0.0069)	0.0087**	(0.0042)	-0.0311***	(0.0054)
Education (reference less than high school)						
High school	0.1189***	(0.0041)	-0.0079***	(0.0025)	-0.0628***	(0.0031)
University	0.2049***	(0.0050)	-0.0427***	(0.0030)	-0.1416***	(0.0039)
Middle income	0.0741***	(0.0054)	0.0184***	(0.0033)	-0.0447***	(0.0042)
High income	0.0958***	(0.0087)	0.0037	(0.0052)	-0.0769***	(0.0067)
Spouse retired	-0.0079**	(0.0040)	0.0027	(0.0024)	0.0091***	(0.0031)
Census (Northwest reference)						
Midwest	-0.0813***	(0.0051)	-0.0171***	(0.0031)	-0.0024	(0.0039)
South	-0.1387***	(0.0046)	-0.0187***	(0.0028)	-0.0037	(0.0035)
West	-0.0310***	(0.0053)	-0.0124***	(0.0032)	-0.0313***	(0.0041)
Others	-0.0179	(0.0480)	-0.0212	(0.0291)	0.0122	(0.0369)
US birth	0.0245***	(0.0050)	0.0130***	(0.0030)	0.0423***	(0.0039)
Children no	-0.0050***	(0.0008)	0.0002	(0.0005)	0.0022***	(0.0006)
Observation	92,651		92,675		92,095	

Notes: (i) standard errors are in parentheses. (ii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.47: Fuzzy parametric RD estimates – Health behavior at cutoff 65 (Females only)

	Drinking participation		Drinking at least 3 time per week		Smoking participation	
Retired	-0.0429***	(0.0057)	-0.0028	(0.0025)	0.0362***	(0.0043)
Age-65	-0.0033	(0.0022)	-0.0026***	(0.0009)	-0.0151***	(0.0016)
(Age-65)^sq	0.0003**	(0.0002)	0.0001	(0.0001)	-0.0005***	(0.0001)
T*(age-65)	0.0055	(0.0047)	0.0001	(0.0020)	0.0075**	(0.0035)
T*(age-65)^sq	-0.0009***	(0.0003)	0.0000	(0.0001)	0.0004	(0.0002)
Marital Status (reference single)						
Married	-0.0750***	(0.0118)	-0.0461***	(0.0051)	-0.1368***	(0.0089)
Others (S,D,W)	-0.0195*	(0.0104)	-0.0028	(0.0045)	0.0184**	(0.0078)
Races (reference White)						
Black	-0.1238***	(0.0057)	-0.0104***	(0.0024)	-0.0341***	(0.0043)
Others	-0.1274***	(0.0096)	-0.0052	(0.0041)	-0.0456***	(0.0072)
Education (reference less than high school)						
High school	0.1538***	(0.0055)	-0.0039	(0.0024)	-0.0624***	(0.0042)
University	0.2454***	(0.0070)	-0.0147***	(0.0030)	-0.1421***	(0.0053)
Middle income	0.0814***	(0.0081)	0.0099***	(0.0035)	-0.0380***	(0.0061)
High income	0.1259***	(0.0151)	0.0096	(0.0065)	-0.0533***	(0.0114)
Spouse retired	-0.0112**	(0.0057)	0.0066***	(0.0025)	0.0150***	(0.0043)
Census (Northwest reference)						
Midwest	-0.0836***	(0.0067)	-0.0102***	(0.0029)	-0.0045	(0.0051)
South	-0.1444***	(0.0061)	-0.0121***	(0.0026)	-0.0143***	(0.0046)
West	-0.0188***	(0.0071)	-0.0079***	(0.0031)	-0.0393***	(0.0054)
Others	-0.0132	(0.0696)	-0.0209	(0.0300)	-0.0959*	(0.0524)
US birth	0.0783***	(0.0074)	0.0252***	(0.0032)	0.0527***	(0.0056)
Children no	-0.0069***	(0.0011)	-0.0003	(0.0005)	-0.0007	(0.0008)
Obs	52069.0000		52083.0000		51755.0000	

Notes: (i) standard errors are in parentheses. (ii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

Table 1.48: Fuzzy parametric RD estimates – Health behavior at cutoff 65 (males only)

	Drinking participation		Drinking at least 3 time per week		Smoking participation	
retired	-0.0247***	(0.0065)	0.0131***	(0.0050)	0.0249***	(0.0051)
Age-65	-0.0022	(0.0025)	-0.0060***	(0.0019)	-0.0106***	(0.0020)
(Age-65)^sq	0.0005***	(0.0002)	0.0002	(0.0001)	-0.0002	(0.0001)
T*(age-65)	0.0036	(0.0052)	0.0003	(0.0041)	-0.0035	(0.0042)
T*(age-65)^sq	-0.0008***	(0.0003)	-0.0003	(0.0003)	0.0004	(0.0003)
Marital Status (reference single)						
married	0.0375***	(0.0127)	-0.0379***	(0.0099)	-0.1124***	(0.0101)
Others (S,D,W)	0.0606***	(0.0125)	0.0445***	(0.0097)	0.0388***	(0.0099)
Races (reference White)						
Black	-0.0730***	(0.0070	-0.0359***	(0.0054)	0.0162***	(0.0056)
Others	-0.0433***	(0.0101	0.0209***	(0.0078)	-0.0168**	(0.0080)
Education (reference less than high school)						
High school	0.0750***	0.0060	-0.0137***	(0.0047)	-0.0639***	(0.0048)
University	0.1599***	0.0071	-0.0703***	(0.0055)	-0.1406***	(0.0057)
Middle income	0.0764***	0.0074	0.0197***	(0.0057)	-0.0515***	(0.0058)
High income	0.0994***	0.0107	0.0028	(0.0083)	-0.0885***	(0.0085)
Spouse retired	-0.0109*	0.0056	-0.0011	(0.0044)	0.0042	(0.0044)
Census (Northwest reference)						
Midwest	-0.0769***	0.0076	-0.0271***	(0.0059)	0.0017	(0.0060)
South	-0.1306***	0.0070	-0.0268***	(0.0054)	0.0104*	(0.0055)
West	-0.0454***	0.0079	-0.0193***	(0.0062)	-0.0214***	(0.0063)
Others	-0.0336	0.0659	-0.0119	(0.0513)	0.1100**	(0.0522)
US birth	-0.0272***	0.0069	-0.0015	(0.0053)	0.0317***	(0.0054)
Children no	-0.0026**	0.0012	0.0013	(0.0009)	0.0056***	(0.0009)
Obs	40,582		40,592		40,340	

Notes: (i) standard errors are in parentheses. (ii) *,**,*** indicate significance level at 10%, 5%, 1% respectively.

1.10. Conclusion and Discussion

This paper estimates the causal effect of retirement on health status and health-related behavior using ten waves (1996-2014) of the U.S. Health and Retirement Study (HRS) survey. Non-parametric and parametric Fuzzy Regression Discontinuity Design (RDD) techniques are applied to address the potential reverse causality from health to retirement (endogeneity issue) by exploiting the exogenous variation in retirement decisions induced by U.S. pension eligibility ages. Using non-parametric and parametric methods is a useful way to argue that the estimation results do not rely on the chosen strategy. The retirement-age distribution in HRS dataset shows that there are two discontinuities in the retirement decision of individuals at age 62 and 65, where the former is the earliest age of retirement to receive partial benefits and the latter is the official retirement age (full claim for pension benefits) in the U.S. However, the regression discontinuity estimates and graphical investigation suggested that the discontinuity in retirement is significant and important at age 65, when there is an increasing probability of individuals retiring as they approach the cut-off point at age 65.

In general, retirement has a significant negative impact on health status. The non-parametric fuzzy RD results show that the significant jump in retirement probability is associated with a negative and highly significant jump in the reported total word recall test score. Using the MSE-optimal bandwidth, the non-parametric fuzzy RD estimator suggests that retired individuals experience a drop in their cognitive test score by about 0.8 points, which is equivalent to an approximately 8% decline in the cognitive functioning score of retired individuals. Retirement has a significant negative impact on mental health, as measured by the CESD depression scale. The non-parametric fuzzy RD estimator suggests that retired individuals experience a dramatic increase in the CESD scale by about 0.42 points with a 95% confidence

interval [0.294554, 0.550295]. In addition, the RD estimation is not sensitive to changes in the bandwidth choice. The RD effect of retirement on mental health is stable across the three RD inference methods with a small difference in standard errors in the robust-bias corrected approach. Retirement is associated with a significant negative impact on self-reported general health status. Retired individuals are 13 percentage points less likely to report that they are in excellent, very good or good health than non-retired individuals. The effects of other covariates used in the study are consistent with previous literature. Moreover, the effect of retirement on the three health status measures, cognitive functioning, mental health, and self-reported health, are not sensitive to different bandwidth selections.

The non-parametric fuzzy RD results based on the CV-optimal bandwidth selection are consistent with the RD results based on the MSE-optimal bandwidth selection. However, the MSE-optimal bandwidth length, for all estimations, is larger than the CV-optimal bandwidth. These findings are expected because the adjusted MSE-optimal bandwidth by CCT gives more weight to the variance, and the chosen bandwidth could be larger to minimize both the bias and the variance under the given bandwidth.

Regarding health-related behavior, retirement has a negative effect on alcohol consumption participation and intensity but has a positive impact on smoking participation. The non-parametric fuzzy RD estimates indicate that retirees are 8.8 percentage points less likely to be drinkers relative to non-retirees and they are 4 percentage points less likely to consume alcohol more than three times per week, for those who continue drinking after retirement, relative to non-retirees. However, the effect of retirement on alcohol consumption intensity per week is small and not significant for all bandwidth choices and across the three different non-parametric fuzzy RD estimators. Also, retired individuals are 3 percentage points more likely to

be smokers relative to non-retired individuals. However, the impact of retirement on smoking is small in magnitude and very sensitive to bandwidth choice.

The findings suggest that there is significant heterogeneity across gender and socioeconomic groups. In general, the transition from working to retirement status has a larger negative impact on retired males than retired females. Retired males experience a 1.08 point decline in their cognitive scale relative to non-retired males, while retired females experience a 0.84 decline in their cognitive score relative to non-retired females. Similarly, retired males are 16 percentage points less likely to be in excellent, very good, or good health, while retired females are 9 percentage points less likely to be in good health. For socioeconomic status, retirement has a larger negative impact on low-educated retired individuals than high-educated retired individuals. Low and high educated retirees have a 1.9 point and 0.8 point decline in their cognitive scale compared to low and high educated non-retirees, respectively. The low-educated retirees experience a 0.52 point increase in the CESD depression scale relative to low-educated non-retirees, while the highly-educated retirees experience a 0.18 point increase in the CESD depression scale relative to highly-educated non-retirees. That is, highly educated retirees experience a lower negative impact on their mental health. Also, low-educated and high educated retirees are 14 and 8 percentage points less likely to report that they are in good health relative to low and high-educated non-retirees, respectively. Previous literature has shown that well-educated individuals have better health outcomes because they are more efficient in their health investment (Grossman and Kaestner, 1997). For health-related behavior, retirement has a higher negative impact on drinking participation and drinking intensity for retired males compared to retired females. Retired females and retired males are 5 and 12 percentage points less likely to be drinkers, respectively, and 1 and 5 percentage points less likely to drink more than three times

per week among retirees who decided to continue drinking after retirement, respectively. In contrast, retired females and retired males are 2 and 4 percentage points more likely to be smokers relative to non-retired females and non-retired males, respectively.

To address the issue of the dramatic increase in the Medicare health insurance eligibility at age 65, I assessed the robustness of the RD results by performing several checks. First, I conducted a similar RD analysis at age 62 when there is a significant jump in retirement at age 62, the earliest age at which Americans can receive their social security payment. The fuzzy RD results at cut-off age 62 are consistent with the RD results at cut-off age 65 for health and health-related behavior. Second, I run the RD analysis on subgroups of low-income individuals who are in the lowest 20% of the income distribution in the dataset or individuals whose income is below the federal poverty line after adjustment for the number of persons in the family. This subgroup includes persons who may be eligible for Medicaid before age 65. The RD results from this subgroup are consistent with the baseline RD results. Third, I run the RD analysis on a subgroup of individuals who continually have insurance at least 5 years before retirement and continue to be insured after retirement to ascertain that retirement impacts on health are not driven by receiving health insurance coverage or retirees' access to health insurance coverage. The RD results from this subgroup are also consistent with the RD results from the low-income subgroup and the baseline RD results. However, the negative impacts of retirement on health outcomes are slightly lower in magnitude, especially for self-reported health, for both the low-income subgroup and the continually insured subgroup.

During the analysis of the effect of retirement on health in this paper, I analyzed the robustness of the fuzzy RD estimates by using a variety of distinct procedures. I found that the main RD findings are robust to different model specifications, parametric and nonparametric

regression strategies, i.e., the RD results do not rely on the chosen strategy. I also found that the RD estimates of the effect of retirement do not change across the three different weighting schemes: Triangular, Uniform, and Epanechnikov Kernel function, which give high weight to observations in the neighborhood of the cut-off but the low weight to observations farther from the cut-off. I checked the robustness of the fuzzy RD results under different polynomial procedures, linear and quadratic, and there was no significant change in the fuzzy RD results for parametric and non-parametric procedures. The robustness of the regression is also verified by estimating the effect of retirement at different bandwidths around the cut-off point that shows increments by 5 percentage points, 50% to 200% percent of the CV-optimal bandwidth, which lends substantial validity to the results. Also, the estimations showed that adding covariates to the main model in the non-parametric RD setting did not change the results dramatically but improved the efficiency of the estimated results.

The findings of this study have important policy implications. Most importantly, since the study suggests that retirement has a negative impact on health outcomes and improvement in health-related behavior, reducing drinking participation and frequency, any policy that prolongs the working period, such as increases in the retirement age, will be accompanied by social benefits which may be greater than the estimated savings in the government's pension burden. Increasing the retirement age could delay the decline in cognitive function and preserve the impairment due to aging and could also improve mental health through mitigating factors that increase depression. Moreover, since late retirement is health-preserving, as measured by self-reported health status and the other two health outcomes, labor force participation of older people can reduce health care utilization and expenditure, in addition to relieving the

government's financial pension burden. In contrast, delaying retirement age could increase drinking participation and frequency among working individuals.

The negative impact of retirement on health outcomes and the improvement in health-related behavior related to drinking participation and intensity in this study may be also related to the nature of the sample used, which dropped individuals who retired due to disability or who retired involuntarily. When individuals decide to retire voluntarily, they may desire to improve their health and enjoy their life by increasing participation in healthy behaviors, which may result in less smoking and drinking participation and intensity, among those who continue to be smokers or drinkers. Future research includes extending this analysis to investigate the impact of voluntary and involuntary retirement on health and health behaviors, sine involuntary retirement may be accompanied by more stress.

1.11. Appendices

APPENDIX A:

The observed outcome can be written as follows

$$Y_i = Y_i(0) + D_i.[Y_i(1) - Y(0)]$$

$$\mathbb{E}[Y|x] = \mathbb{E}[Y(0)|x] + \mathbb{E}\big[D\big(Y(1) - Y(0)\big)\big|x\big]$$

$$= \mathbb{E}[Y(0)|x + \mathbb{E}[D|x] \cdot \mathbb{E}[Y(1) - Y(0)|x]$$

$$= \mathbb{E}[Y(0)|x] + \mathbb{E}[D|x] \cdot ATE(x)$$

$$\lim_{x\downarrow\bar{x}} \mathbb{E}[Y|X = x] = \lim_{x\downarrow\bar{x}} \mathbb{E}[Y(0)|X = x] + \lim_{x\downarrow\bar{x}} \mathbb{E}[D|X = x] \cdot \lim_{x\downarrow\bar{x}} ATE(x)$$

$$\lim_{x\uparrow\bar{x}} \mathbb{E}[Y|X = x] = \lim_{x\uparrow\bar{x}} \mathbb{E}[Y(0)|X = x] + \lim_{x\uparrow\bar{x}} \mathbb{E}[D|X = x] \cdot \lim_{x\uparrow\bar{x}} ATE(x)$$

$$\lim_{x\downarrow\bar{x}} \mathbb{E}[Y|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[Y|X = x] =$$

$$\underbrace{\lim_{x\downarrow\bar{x}} \mathbb{E}[Y(0)|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[Y(0)|X = x]}_{which\ equal\ ZERO\ under\ continuity\ assumption} +$$

$$\lim_{x\downarrow\bar{x}} \mathbb{E}[D|X = x] \cdot \lim_{x\downarrow\bar{x}} ATE(x) - \lim_{x\uparrow\bar{x}} \mathbb{E}[D|X = x] \cdot \lim_{x\uparrow\bar{x}} ATE(x)$$

$$\lim_{x\downarrow\bar{x}} \mathbb{E}[Y|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[Y|X = x] =$$

$$\lim_{x\downarrow\bar{x}} \mathbb{E}[D|X = x] \cdot \lim_{x\downarrow\bar{x}} ATE(x) - \lim_{x\uparrow\bar{x}} \mathbb{E}[D|X = x] \cdot \lim_{x\uparrow\bar{x}} ATE(x)$$

$$\lim_{x\downarrow\bar{x}} \mathbb{E}[Y|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[Y|X = x] =$$

$$ATE\ [\lim_{x\downarrow\bar{x}} \mathbb{E}[D|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[D|X = x]$$

Therefore,

$$LATE = \frac{\lim_{x\downarrow\bar{x}} \mathbb{E}[Y|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[Y|X = x]}{[\lim_{x\downarrow\bar{x}} \mathbb{E}[D|X = x] - \lim_{x\uparrow\bar{x}} \mathbb{E}[D|X = x]}$$

APPENDIX B:

Table B1

Fuzzy RD estimates - cognitive functioning at different bandwidth - All sample at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.8545	0.1399	0.0000	-1.1286	-0.5804
*Fuzzy RD*_50	-0.8842	0.1739	0.0000	-1.2250	-0.5434
*Fuzzy RD*_55	-0.9281	0.1999	0.0000	-1.3199	-0.5363
*Fuzzy RD*_60	-0.8187	0.1760	0.0000	-1.1637	-0.4736
*Fuzzy RD*_65	-0.9049	0.1799	0.0000	-1.2575	-0.5523
*Fuzzy RD*_70	-0.8918	0.1394	0.0000	-1.1650	-0.6185
*Fuzzy RD*_75	-0.8916	0.1663	0.0000	-1.2176	-0.5657
*Fuzzy RD*_80	-0.8786	0.1588	0.0000	-1.1898	-0.5675
*Fuzzy RD*_85	-0.8665	0.1531	0.0000	-1.1667	-0.5664
*Fuzzy RD*_90	-0.8700	0.1495	0.0000	-1.1630	-0.5770
*Fuzzy RD*_95	-0.8444	0.1431	0.0000	-1.1248	-0.5639
*Fuzzy RD*_105	-0.8101	0.1292	0.0000	-1.0634	-0.5568
*Fuzzy RD*_110	-0.8205	0.1309	0.0000	-1.0770	-0.5641
*Fuzzy RD*_115	-0.8335	0.1005	0.0000	-1.0304	-0.6365
*Fuzzy RD*_120	-0.7997	0.1243	0.0000	-1.0432	-0.5561
*Fuzzy RD*_125	-0.7900	0.1213	0.0000	-1.0277	-0.5523
*Fuzzy RD*_130	-0.8116	0.0928	0.0000	-0.9935	-0.6297
*Fuzzy RD*_135	-0.8104	0.1166	0.0000	-1.0390	-0.5818
*Fuzzy RD*_140	-0.8122	0.0887	0.0000	-0.9862	-0.6383
*Fuzzy RD*_145	-0.7851	0.1079	0.0000	-0.9965	-0.5736
*Fuzzy RD*_150	-0.7720	0.1091	0.0000	-0.9858	-0.5582
*Fuzzy RD*_155	-0.8113	0.0831	0.0000	-0.9742	-0.6483
*Fuzzy RD*_160	-0.7633	0.1050	0.0000	-0.9691	-0.5574
*Fuzzy RD*_165	-0.7769	0.1003	0.0000	-0.9735	-0.5802
*Fuzzy RD*_170	-0.7980	0.1018	0.0000	-0.9976	-0.5985
*Fuzzy RD*_175	-0.7925	0.1001	0.0000	-0.9886	-0.5963
*Fuzzy RD*_180	-0.7854	0.0984	0.0000	-0.9783	-0.5925
*Fuzzy RD*_185	-0.7785	0.0968	0.0000	-0.9682	-0.5887
*Fuzzy RD*_190	-0.7487	0.0926	0.0000	-0.9303	-0.5672
*Fuzzy RD*_195	-0.7422	0.0913	0.0000	-0.9211	-0.5634
*Fuzzy RD*_200	-0.7575	0.0711	0.0000	-0.8967	-0.6182

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table B2:

Fuzzy RD estimates - cognitive functioning at different bandwidth - females only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.7776	0.1768	0.0000	-1.1240	-0.4311
*Fuzzy RD*_50	-0.6966	0.2536	0.0060	-1.1937	-0.1995
*Fuzzy RD*_55	-0.7331	0.2398	0.0020	-1.2031	-0.2631
*Fuzzy RD*_60	-0.8304	0.2455	0.0010	-1.3115	-0.3493
*Fuzzy RD*_65	-0.8399	0.2340	0.0000	-1.2984	-0.3813
*Fuzzy RD*_70	-0.7971	0.2105	0.0000	-1.2097	-0.3844
*Fuzzy RD*_75	-0.8420	0.2152	0.0000	-1.2638	-0.4202
*Fuzzy RD*_80	-0.8152	0.2027	0.0000	-1.2126	-0.4178
*Fuzzy RD*_85	-0.8184	0.1582	0.0000	-1.1285	-0.5083
*Fuzzy RD*_90	-0.7881	0.1883	0.0000	-1.1573	-0.4190
*Fuzzy RD*_95	-0.7659	0.1771	0.0000	-1.1131	-0.4187
*Fuzzy RD*_105	-0.7605	0.1670	0.0000	-1.0879	-0.4332
*Fuzzy RD*_110	-0.7893	0.1338	0.0000	-1.0516	-0.5270
*Fuzzy RD*_115	-0.7472	0.1629	0.0000	-1.0666	-0.4279
*Fuzzy RD*_120	-0.7581	0.1548	0.0000	-1.0616	-0.4547
*Fuzzy RD*_125	-0.7390	0.1551	0.0000	-1.0431	-0.4350
*Fuzzy RD*_130	-0.7945	0.1537	0.0000	-1.0957	-0.4933
*Fuzzy RD*_135	-0.7416	0.1484	0.0000	-1.0325	-0.4508
*Fuzzy RD*_140	-0.8008	0.1145	0.0000	-1.0253	-0.5763
*Fuzzy RD*_145	-0.7971	0.1441	0.0000	-1.0795	-0.5147
*Fuzzy RD*_150	-0.7641	0.1367	0.0000	-1.0320	-0.4961
*Fuzzy RD*_155	-0.7212	0.1370	0.0000	-0.9896	-0.4528
*Fuzzy RD*_160	-0.7151	0.1345	0.0000	-0.9787	-0.4516
*Fuzzy RD*_165	-0.7465	0.1293	0.0000	-1.0000	-0.4931
*Fuzzy RD*_170	-0.6992	0.1298	0.0000	-0.9536	-0.4448
*Fuzzy RD*_175	-0.7320	0.1250	0.0000	-0.9770	-0.4871
*Fuzzy RD*_180	-0.6813	0.1255	0.0000	-0.9273	-0.4353
*Fuzzy RD*_185	-0.7369	0.0956	0.0000	-0.9244	-0.5495
*Fuzzy RD*_190	-0.7122	0.1191	0.0000	-0.9457	-0.4787
*Fuzzy RD*_195	-0.7276	0.0927	0.0000	-0.9093	-0.5458
*Fuzzy RD*_200	-0.7245	0.0907	0.0000	-0.9024	-0.5467

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table B3:

Fuzzy RD estimates - cognitive functioning at different bandwidth - males only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-1.1586	0.2276	0.0000	-1.6047	-0.7126
*Fuzzy RD*_50	-1.2035	0.3472	0.0010	-1.8839	-0.5231
*Fuzzy RD*_55	-1.1104	0.2976	0.0000	-1.6938	-0.5271
*Fuzzy RD*_60	-1.2517	0.3136	0.0000	-1.8663	-0.6372
*Fuzzy RD*_65	-1.2462	0.2974	0.0000	-1.8290	-0.6633
*Fuzzy RD*_70	-1.2198	0.2245	0.0000	-1.6599	-0.7797
*Fuzzy RD*_75	-1.1059	0.2505	0.0000	-1.5970	-0.6149
*Fuzzy RD*_80	-1.1894	0.2608	0.0000	-1.7006	-0.6783
*Fuzzy RD*_85	-1.1717	0.2504	0.0000	-1.6625	-0.6808
*Fuzzy RD*_90	-1.1603	0.2422	0.0000	-1.6349	-0.6857
*Fuzzy RD*_95	-1.1548	0.2346	0.0000	-1.6146	-0.6951
*Fuzzy RD*_105	-1.1467	0.2213	0.0000	-1.5805	-0.7129
*Fuzzy RD*_110	-1.1159	0.2149	0.0000	-1.5371	-0.6947
*Fuzzy RD*_115	-1.1189	0.2093	0.0000	-1.5291	-0.7086
*Fuzzy RD*_120	-1.1096	0.1583	0.0000	-1.4198	-0.7994
*Fuzzy RD*_125	-1.0947	0.1984	0.0000	-1.4835	-0.7060
*Fuzzy RD*_130	-1.0642	0.1935	0.0000	-1.4435	-0.6850
*Fuzzy RD*_135	-1.0538	0.1891	0.0000	-1.4243	-0.6832
*Fuzzy RD*_140	-1.0436	0.1849	0.0000	-1.4059	-0.6812
*Fuzzy RD*_145	-1.0290	0.1810	0.0000	-1.3837	-0.6743
*Fuzzy RD*_150	-1.0269	0.1774	0.0000	-1.3745	-0.6792
*Fuzzy RD*_155	-1.0105	0.1739	0.0000	-1.3514	-0.6696
*Fuzzy RD*_160	-0.9840	0.1706	0.0000	-1.3184	-0.6496
*Fuzzy RD*_165	-0.9920	0.1677	0.0000	-1.3207	-0.6632
*Fuzzy RD*_170	-0.9675	0.1647	0.0000	-1.2904	-0.6447
*Fuzzy RD*_175	-0.9594	0.1619	0.0000	-1.2768	-0.6420
*Fuzzy RD*_180	-0.9755	0.1594	0.0000	-1.2879	-0.6630
*Fuzzy RD*_185	-0.9689	0.1568	0.0000	-1.2762	-0.6616
*Fuzzy RD*_190	-0.9308	0.1490	0.0000	-1.2228	-0.6388
*Fuzzy RD*_195	-0.9311	0.1518	0.0000	-1.2287	-0.6336
*Fuzzy RD*_200	-0.9239	0.1447	0.0000	-1.2074	-0.6403

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table B4:

Fuzzy RD estimates - cognitive functioning at different bandwidth - low-educated only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-1.5510	0.4336	0.0000	-2.4008	-0.7012
*Fuzzy RD*_50	-1.7171	0.7268	0.0180	-3.1416	-0.2927
*Fuzzy RD*_55	-1.7774	0.8452	0.0350	-3.4340	-0.1208
*Fuzzy RD*_60	-1.7567	0.7922	0.0270	-3.3095	-0.2040
*Fuzzy RD*_65	-1.7197	0.6168	0.0050	-2.9286	-0.5107
*Fuzzy RD*_70	-1.6908	0.7023	0.0160	-3.0673	-0.3142
*Fuzzy RD*_75	-1.6646	0.5478	0.0020	-2.7384	-0.5909
*Fuzzy RD*_80	-1.6748	0.6377	0.0090	-2.9247	-0.4249
*Fuzzy RD*_85	-1.6406	0.5007	0.0010	-2.6220	-0.6591
*Fuzzy RD*_90	-1.6356	0.5876	0.0050	-2.7872	-0.4839
*Fuzzy RD*_95	-1.6065	0.4513	0.0000	-2.4909	-0.7220
*Fuzzy RD*_105	-1.4887	0.4202	0.0000	-2.3124	-0.6651
*Fuzzy RD*_110	-1.4155	0.5351	0.0080	-2.4642	-0.3667
*Fuzzy RD*_115	-1.3389	0.3925	0.0010	-2.1081	-0.5697
*Fuzzy RD*_120	-1.2379	0.4967	0.0130	-2.2114	-0.2644
*Fuzzy RD*_125	-1.1960	0.3764	0.0010	-1.9338	-0.4582
*Fuzzy RD*_130	-1.0804	0.4675	0.0210	-1.9966	-0.1641
*Fuzzy RD*_135	-1.0681	0.3555	0.0030	-1.7649	-0.3714
*Fuzzy RD*_140	-0.9886	0.4542	0.0300	-1.8788	-0.0983
*Fuzzy RD*_145	-0.9250	0.4404	0.0360	-1.7881	-0.0619
*Fuzzy RD*_150	-0.9250	0.3354	0.0060	-1.5823	-0.2677
*Fuzzy RD*_155	-0.8871	0.4387	0.0430	-1.7470	-0.0273
*Fuzzy RD*_160	-0.8185	0.4156	0.0490	-1.6331	-0.0040
*Fuzzy RD*_165	-0.7874	0.4080	0.0540	-1.5870	0.0122
*Fuzzy RD*_170	-0.7838	0.3063	0.0100	-1.3840	-0.1835
*Fuzzy RD*_175	-0.7492	0.4047	0.0640	-1.5424	0.0439
*Fuzzy RD*_180	-0.6857	0.3847	0.0750	-1.4398	0.0684
*Fuzzy RD*_185	-0.6278	0.3778	0.0970	-1.3683	0.1126
*Fuzzy RD*_190	-0.6169	0.2791	0.0270	-1.1639	-0.0699
*Fuzzy RD*_195	-0.5785	0.3744	0.1220	-1.3122	0.1553
*Fuzzy RD*_200	-0.5489	0.2675	0.0400	-1.0731	-0.0247

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table B5:

Fuzzy RD estimates - cognitive functioning at different bandwidth - high-educated only (cutoff age 65)

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.7543	0.1998	0.0000	-1.1459	-0.3628
*Fuzzy RD*_50	-0.9188	0.3074	0.0030	-1.5213	-0.3162
*Fuzzy RD*_55	-0.8937	0.2831	0.0020	-1.4487	-0.3388
*Fuzzy RD*_60	-0.8974	0.2183	0.0000	-1.3252	-0.4696
*Fuzzy RD*_65	-0.8910	0.2560	0.0010	-1.3927	-0.3893
*Fuzzy RD*_70	-0.8752	0.2454	0.0000	-1.3562	-0.3943
*Fuzzy RD*_75	-0.8515	0.2355	0.0000	-1.3130	-0.3900
*Fuzzy RD*_80	-0.8582	0.2289	0.0000	-1.3069	-0.4095
*Fuzzy RD*_85	-0.8193	0.1755	0.0000	-1.1632	-0.4754
*Fuzzy RD*_90	-0.7944	0.2120	0.0000	-1.2098	-0.3789
*Fuzzy RD*_95	-0.7713	0.2051	0.0000	-1.1733	-0.3693
*Fuzzy RD*_105	-0.7302	0.1927	0.0000	-1.1080	-0.3525
*Fuzzy RD*_110	-0.7192	0.1462	0.0000	-1.0058	-0.4326
*Fuzzy RD*_115	-0.7061	0.1834	0.0000	-1.0657	-0.3466
*Fuzzy RD*_120	-0.6938	0.1789	0.0000	-1.0444	-0.3432
*Fuzzy RD*_125	-0.6850	0.1737	0.0000	-1.0254	-0.3445
*Fuzzy RD*_130	-0.6725	0.1698	0.0000	-1.0053	-0.3396
*Fuzzy RD*_135	-0.6607	0.1663	0.0000	-0.9866	-0.3348
*Fuzzy RD*_140	-0.6268	0.1562	0.0000	-0.9330	-0.3206
*Fuzzy RD*_145	-0.6370	0.1605	0.0000	-0.9516	-0.3223
*Fuzzy RD*_150	-0.6324	0.1575	0.0000	-0.9411	-0.3238
*Fuzzy RD*_155	-0.6445	0.1539	0.0000	-0.9461	-0.3430
*Fuzzy RD*_160	-0.6228	0.1516	0.0000	-0.9199	-0.3256
*Fuzzy RD*_165	-0.6359	0.1483	0.0000	-0.9266	-0.3452
*Fuzzy RD*_170	-0.6109	0.1463	0.0000	-0.8976	-0.3242
*Fuzzy RD*_175	-0.6071	0.1383	0.0000	-0.8782	-0.3359
*Fuzzy RD*_180	-0.6255	0.1080	0.0000	-0.8373	-0.4138
*Fuzzy RD*_185	-0.6011	0.1392	0.0000	-0.8739	-0.3283
*Fuzzy RD*_190	-0.5959	0.1322	0.0000	-0.8550	-0.3368
*Fuzzy RD*_195	-0.5895	0.1303	0.0000	-0.8449	-0.3340
*Fuzzy RD*_200	-0.5811	0.1330	0.0000	-0.8418	-0.3204

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table B6:

Fuzzy RD estimates - cognitive functioning at different bandwidth - low income only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.6735	0.1501	0.0000	-0.9677	-0.3793
*Fuzzy RD*_50	-0.7666	0.2278	0.0010	-1.2131	-0.3200
*Fuzzy RD*_55	-0.6831	0.2254	0.0020	-1.1249	-0.2414
*Fuzzy RD*_60	-0.6886	0.2128	0.0010	-1.1057	-0.2715
*Fuzzy RD*_65	-0.7016	0.1639	0.0000	-1.0228	-0.3803
*Fuzzy RD*_70	-0.6716	0.1825	0.0000	-1.0292	-0.3140
*Fuzzy RD*_75	-0.7154	0.1777	0.0000	-1.0637	-0.3672
*Fuzzy RD*_80	-0.6861	0.1698	0.0000	-1.0190	-0.3533
*Fuzzy RD*_85	-0.7232	0.1382	0.0000	-0.9941	-0.4523
*Fuzzy RD*_90	-0.7143	0.1672	0.0000	-1.0419	-0.3866
*Fuzzy RD*_95	-0.6736	0.1548	0.0000	-0.9769	-0.3702
*Fuzzy RD*_105	-0.6671	0.1458	0.0000	-0.9530	-0.3813
*Fuzzy RD*_110	-0.6598	0.1420	0.0000	-0.9381	-0.3815
*Fuzzy RD*_115	-0.6892	0.1124	0.0000	-0.9095	-0.4688
*Fuzzy RD*_120	-0.6844	0.1093	0.0000	-0.8987	-0.4701
*Fuzzy RD*_125	-0.6318	0.1319	0.0000	-0.8904	-0.3732
*Fuzzy RD*_130	-0.6222	0.1290	0.0000	-0.8749	-0.3694
*Fuzzy RD*_135	-0.6699	0.1305	0.0000	-0.9257	-0.4140
*Fuzzy RD*_140	-0.6699	0.1278	0.0000	-0.9204	-0.4193
*Fuzzy RD*_145	-0.6046	0.1212	0.0000	-0.8421	-0.3671
*Fuzzy RD*_150	-0.6737	0.1228	0.0000	-0.9144	-0.4330
*Fuzzy RD*_155	-0.6715	0.0936	0.0000	-0.8550	-0.4880
*Fuzzy RD*_160	-0.6681	0.1182	0.0000	-0.8998	-0.4364
*Fuzzy RD*_165	-0.6632	0.1161	0.0000	-0.8907	-0.4357
*Fuzzy RD*_170	-0.6572	0.0882	0.0000	-0.8301	-0.4843
*Fuzzy RD*_175	-0.6515	0.0867	0.0000	-0.8214	-0.4815
*Fuzzy RD*_180	-0.6277	0.1074	0.0000	-0.8382	-0.4171
*Fuzzy RD*_185	-0.6191	0.1058	0.0000	-0.8264	-0.4118
*Fuzzy RD*_190	-0.6264	0.0824	0.0000	-0.7879	-0.4650
*Fuzzy RD*_195	-0.6176	0.0810	0.0000	-0.7764	-0.4588
*Fuzzy RD*_200	-0.5959	0.1012	0.0000	-0.7941	-0.3976

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

APPENDIX C:

Table C1

Fuzzy RD estimates -Mental health at different bandwidth - All sample at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	0.4091	0.0831	0.0000	0.2463	0.5719
*Fuzzy RD*_50	0.4771	0.1207	0.0000	0.2406	0.7137
*Fuzzy RD*_55	0.4779	0.1149	0.0000	0.2528	0.7031
*Fuzzy RD*_60	0.4631	0.1092	0.0000	0.2490	0.6772
*Fuzzy RD*_65	0.4312	0.1082	0.0000	0.2191	0.6432
*Fuzzy RD*_70	0.4375	0.0999	0.0000	0.2418	0.6333
*Fuzzy RD*_75	0.4252	0.0958	0.0000	0.2375	0.6129
*Fuzzy RD*_80	0.4256	0.0777	0.0000	0.2733	0.5779
*Fuzzy RD*_85	0.3978	0.0830	0.0000	0.2350	0.5605
*Fuzzy RD*_90	0.4269	0.0862	0.0000	0.2580	0.5959
*Fuzzy RD*_95	0.4250	0.0836	0.0000	0.2611	0.5889
*Fuzzy RD*_105	0.4217	0.0789	0.0000	0.2670	0.5764
*Fuzzy RD*_110	0.4186	0.0641	0.0000	0.2929	0.5442
*Fuzzy RD*_115	0.4129	0.0763	0.0000	0.2634	0.5625
*Fuzzy RD*_120	0.3925	0.0691	0.0000	0.2571	0.5279
*Fuzzy RD*_125	0.3941	0.0675	0.0000	0.2619	0.5264
*Fuzzy RD*_130	0.4166	0.0707	0.0000	0.2781	0.5552
*Fuzzy RD*_135	0.4193	0.0678	0.0000	0.2864	0.5522
*Fuzzy RD*_140	0.4019	0.0634	0.0000	0.2777	0.5260
*Fuzzy RD*_145	0.4217	0.0662	0.0000	0.2920	0.5515
*Fuzzy RD*_150	0.4216	0.0649	0.0000	0.2944	0.5488
*Fuzzy RD*_155	0.4217	0.0636	0.0000	0.2970	0.5464
*Fuzzy RD*_160	0.4225	0.0614	0.0000	0.3022	0.5429
*Fuzzy RD*_165	0.4067	0.0580	0.0000	0.2930	0.5203
*Fuzzy RD*_170	0.4253	0.0472	0.0000	0.3327	0.5179
*Fuzzy RD*_175	0.4279	0.0584	0.0000	0.3135	0.5423
*Fuzzy RD*_180	0.4303	0.0574	0.0000	0.3178	0.5429
*Fuzzy RD*_185	0.4164	0.0545	0.0000	0.3096	0.5233
*Fuzzy RD*_190	0.4183	0.0537	0.0000	0.3130	0.5236
*Fuzzy RD*_195	0.4197	0.0530	0.0000	0.3159	0.5235
*Fuzzy RD*_200	0.4338	0.0547	0.0000	0.3265	0.5410

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table C2:

Fuzzy RD estimates -Mental health at different bandwidth - females only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	0.3979	0.1132	0.0000	0.1761	0.6197
*Fuzzy RD*_50	0.3593	0.1721	0.0370	0.0219	0.6967
*Fuzzy RD*_55	0.4031	0.1445	0.0050	0.1198	0.6864
*Fuzzy RD*_60	0.4247	0.1514	0.0050	0.1280	0.7214
*Fuzzy RD*_65	0.4081	0.1266	0.0010	0.1600	0.6562
*Fuzzy RD*_70	0.3854	0.1194	0.0010	0.1513	0.6194
*Fuzzy RD*_75	0.3356	0.1342	0.0120	0.0725	0.5987
*Fuzzy RD*_80	0.3823	0.1281	0.0030	0.1312	0.6335
*Fuzzy RD*_85	0.3855	0.1240	0.0020	0.1425	0.6285
*Fuzzy RD*_90	0.3492	0.1209	0.0040	0.1122	0.5862
*Fuzzy RD*_95	0.3957	0.1165	0.0010	0.1674	0.6241
*Fuzzy RD*_105	0.3759	0.1042	0.0000	0.1717	0.5801
*Fuzzy RD*_110	0.3758	0.1014	0.0000	0.1770	0.5746
*Fuzzy RD*_115	0.3988	0.0853	0.0000	0.2316	0.5659
*Fuzzy RD*_120	0.3903	0.1018	0.0000	0.1908	0.5898
*Fuzzy RD*_125	0.3944	0.0995	0.0000	0.1995	0.5894
*Fuzzy RD*_130	0.4128	0.0787	0.0000	0.2586	0.5670
*Fuzzy RD*_135	0.4166	0.0764	0.0000	0.2668	0.5663
*Fuzzy RD*_140	0.4104	0.0931	0.0000	0.2278	0.5930
*Fuzzy RD*_145	0.4210	0.0724	0.0000	0.2791	0.5628
*Fuzzy RD*_150	0.4166	0.0895	0.0000	0.2412	0.5919
*Fuzzy RD*_155	0.4105	0.0841	0.0000	0.2458	0.5753
*Fuzzy RD*_160	0.4266	0.0862	0.0000	0.2576	0.5955
*Fuzzy RD*_165	0.4356	0.0844	0.0000	0.2703	0.6009
*Fuzzy RD*_170	0.4386	0.0829	0.0000	0.2761	0.6011
*Fuzzy RD*_175	0.4372	0.0818	0.0000	0.2769	0.5976
*Fuzzy RD*_180	0.4390	0.0805	0.0000	0.2813	0.5968
*Fuzzy RD*_185	0.4444	0.0789	0.0000	0.2897	0.5991
*Fuzzy RD*_190	0.4440	0.0777	0.0000	0.2917	0.5962
*Fuzzy RD*_195	0.4437	0.0765	0.0000	0.2938	0.5936
*Fuzzy RD*_200	0.4289	0.0730	0.0000	0.2858	0.5720

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table C3:

Fuzzy RD estimates -Mental health at different bandwidth - males only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	0.3773	0.0992	0.0000	0.1828	0.5718
*Fuzzy RD*_50	0.4789	0.1356	0.0000	0.2131	0.7447
*Fuzzy RD*_55	0.4806	0.1621	0.0030	0.1630	0.7982
*Fuzzy RD*_60	0.4033	0.1318	0.0020	0.1450	0.6615
*Fuzzy RD*_65	0.4597	0.1448	0.0010	0.1759	0.7435
*Fuzzy RD*_70	0.3910	0.1209	0.0010	0.1540	0.6281
*Fuzzy RD*_75	0.4306	0.1257	0.0010	0.1843	0.6770
*Fuzzy RD*_80	0.4030	0.1124	0.0000	0.1827	0.6233
*Fuzzy RD*_85	0.4290	0.1163	0.0000	0.2009	0.6570
*Fuzzy RD*_90	0.3909	0.1052	0.0000	0.1846	0.5971
*Fuzzy RD*_95	0.4197	0.1131	0.0000	0.1980	0.6414
*Fuzzy RD*_105	0.3956	0.0826	0.0000	0.2338	0.5574
*Fuzzy RD*_110	0.3924	0.0993	0.0000	0.1979	0.5869
*Fuzzy RD*_115	0.3903	0.0766	0.0000	0.2402	0.5404
*Fuzzy RD*_120	0.3920	0.0938	0.0000	0.2081	0.5760
*Fuzzy RD*_125	0.3919	0.0728	0.0000	0.2493	0.5345
*Fuzzy RD*_130	0.3899	0.0705	0.0000	0.2516	0.5281
*Fuzzy RD*_135	0.3874	0.0873	0.0000	0.2163	0.5584
*Fuzzy RD*_140	0.3850	0.0854	0.0000	0.2177	0.5523
*Fuzzy RD*_145	0.3890	0.0859	0.0000	0.2207	0.5573
*Fuzzy RD*_150	0.3819	0.0819	0.0000	0.2213	0.5424
*Fuzzy RD*_155	0.3803	0.0804	0.0000	0.2229	0.5378
*Fuzzy RD*_160	0.3657	0.0758	0.0000	0.2172	0.5142
*Fuzzy RD*_165	0.3851	0.0793	0.0000	0.2297	0.5404
*Fuzzy RD*_170	0.3698	0.0733	0.0000	0.2261	0.5134
*Fuzzy RD*_175	0.3714	0.0721	0.0000	0.2301	0.5127
*Fuzzy RD*_180	0.3867	0.0575	0.0000	0.2740	0.4993
*Fuzzy RD*_185	0.3887	0.0564	0.0000	0.2783	0.4992
*Fuzzy RD*_190	0.3914	0.0712	0.0000	0.2517	0.5310
*Fuzzy RD*_195	0.3932	0.0701	0.0000	0.2557	0.5307
*Fuzzy RD*_200	0.3934	0.0691	0.0000	0.2579	0.5289

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table C4:

Fuzzy RD estimates -Mental health at different bandwidth - low-educated only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	0.5182	0.0925	0.0000	0.3368	0.6995
*Fuzzy RD*_50	0.5811	0.1645	0.0000	0.2586	0.9036
*Fuzzy RD*_55	0.5618	0.1559	0.0000	0.2563	0.8674
*Fuzzy RD*_60	0.5463	0.1481	0.0000	0.2561	0.8365
*Fuzzy RD*_65	0.5256	0.1220	0.0000	0.2865	0.7646
*Fuzzy RD*_70	0.4897	0.1259	0.0000	0.2429	0.7365
*Fuzzy RD*_75	0.5301	0.1294	0.0000	0.2764	0.7838
*Fuzzy RD*_80	0.5314	0.1247	0.0000	0.2870	0.7758
*Fuzzy RD*_85	0.5075	0.1221	0.0000	0.2681	0.7468
*Fuzzy RD*_90	0.5269	0.1165	0.0000	0.2987	0.7551
*Fuzzy RD*_95	0.5021	0.1143	0.0000	0.2782	0.7261
*Fuzzy RD*_105	0.5140	0.1062	0.0000	0.3058	0.7222
*Fuzzy RD*_110	0.4882	0.0983	0.0000	0.2956	0.6807
*Fuzzy RD*_115	0.5161	0.0845	0.0000	0.3505	0.6817
*Fuzzy RD*_120	0.4958	0.0935	0.0000	0.3126	0.6790
*Fuzzy RD*_125	0.4984	0.0914	0.0000	0.3193	0.6775
*Fuzzy RD*_130	0.5204	0.0932	0.0000	0.3378	0.7030
*Fuzzy RD*_135	0.5065	0.0920	0.0000	0.3261	0.6869
*Fuzzy RD*_140	0.5200	0.0891	0.0000	0.3453	0.6947
*Fuzzy RD*_145	0.5084	0.0881	0.0000	0.3357	0.6811
*Fuzzy RD*_150	0.5244	0.0856	0.0000	0.3567	0.6921
*Fuzzy RD*_155	0.5275	0.0687	0.0000	0.3928	0.6622
*Fuzzy RD*_160	0.5305	0.0824	0.0000	0.3689	0.6920
*Fuzzy RD*_165	0.5323	0.0810	0.0000	0.3736	0.6910
*Fuzzy RD*_170	0.5341	0.0647	0.0000	0.4074	0.6609
*Fuzzy RD*_175	0.5205	0.0789	0.0000	0.3659	0.6750
*Fuzzy RD*_180	0.5206	0.0776	0.0000	0.3685	0.6726
*Fuzzy RD*_185	0.5208	0.0764	0.0000	0.3711	0.6704
*Fuzzy RD*_190	0.5359	0.0747	0.0000	0.3895	0.6822
*Fuzzy RD*_195	0.5190	0.0715	0.0000	0.3789	0.6592
*Fuzzy RD*_200	0.5152	0.0729	0.0000	0.3722	0.6581

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table C5:

Fuzzy RD estimates -Mental health at different bandwidth - high-educated only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	0.1757	0.0963	0.0680	-0.0130	0.3644
*Fuzzy RD*_50	0.2002	0.1588	0.2070	-0.1110	0.5115
*Fuzzy RD*_55	0.2112	0.1447	0.1440	-0.0724	0.4947
*Fuzzy RD*_60	0.1837	0.1251	0.1420	-0.0614	0.4289
*Fuzzy RD*_65	0.1943	0.1312	0.1390	-0.0629	0.4515
*Fuzzy RD*_70	0.1791	0.1157	0.1210	-0.0476	0.4058
*Fuzzy RD*_75	0.1942	0.1209	0.1080	-0.0428	0.4312
*Fuzzy RD*_80	0.1917	0.0941	0.0420	0.0072	0.3762
*Fuzzy RD*_85	0.1888	0.1125	0.0930	-0.0318	0.4093
*Fuzzy RD*_90	0.1881	0.1090	0.0840	-0.0255	0.4017
*Fuzzy RD*_95	0.1917	0.1072	0.0740	-0.0185	0.4019
*Fuzzy RD*_105	0.1749	0.0936	0.0620	-0.0086	0.3584
*Fuzzy RD*_110	0.1981	0.0979	0.0430	0.0062	0.3899
*Fuzzy RD*_115	0.1918	0.0940	0.0410	0.0075	0.3760
*Fuzzy RD*_120	0.1929	0.0917	0.0350	0.0131	0.3726
*Fuzzy RD*_125	0.1855	0.0851	0.0290	0.0186	0.3523
*Fuzzy RD*_130	0.2036	0.0886	0.0220	0.0299	0.3773
*Fuzzy RD*_135	0.2040	0.0867	0.0190	0.0341	0.3739
*Fuzzy RD*_140	0.2034	0.0849	0.0170	0.0371	0.3697
*Fuzzy RD*_145	0.1879	0.0787	0.0170	0.0336	0.3422
*Fuzzy RD*_150	0.2022	0.0816	0.0130	0.0423	0.3620
*Fuzzy RD*_155	0.1898	0.0760	0.0130	0.0408	0.3389
*Fuzzy RD*_160	0.1998	0.0582	0.0010	0.0858	0.3138
*Fuzzy RD*_165	0.2017	0.0765	0.0080	0.0516	0.3517
*Fuzzy RD*_170	0.2032	0.0752	0.0070	0.0558	0.3505
*Fuzzy RD*_175	0.2044	0.0551	0.0000	0.0964	0.3124
*Fuzzy RD*_180	0.2054	0.0726	0.0050	0.0630	0.3477
*Fuzzy RD*_185	0.1986	0.0689	0.0040	0.0636	0.3336
*Fuzzy RD*_190	0.2065	0.0520	0.0000	0.1046	0.3084
*Fuzzy RD*_195	0.2104	0.0696	0.0030	0.0740	0.3468
*Fuzzy RD*_200	0.2118	0.0685	0.0020	0.0775	0.3461

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table C6:

Fuzzy RD estimates -Mental health at different bandwidth - low income only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf. Interval]	
Fuzzy RD	0.2580	0.1012	0.0110	0.0597	0.4564
*Fuzzy RD*_50	0.2633	0.1502	0.0800	-0.0310	0.5576
*Fuzzy RD*_55	0.2867	0.1341	0.0330	0.0238	0.5496
*Fuzzy RD*_60	0.2831	0.1413	0.0450	0.0062	0.5600
*Fuzzy RD*_65	0.2706	0.1359	0.0460	0.0042	0.5370
*Fuzzy RD*_70	0.2568	0.1132	0.0230	0.0350	0.4786
*Fuzzy RD*_75	0.2477	0.1065	0.0200	0.0390	0.4564
*Fuzzy RD*_80	0.2396	0.1209	0.0480	0.0026	0.4767
*Fuzzy RD*_85	0.2574	0.1110	0.0200	0.0399	0.4748
*Fuzzy RD*_90	0.2442	0.1084	0.0240	0.0317	0.4567
*Fuzzy RD*_95	0.2603	0.0921	0.0050	0.0798	0.4409
*Fuzzy RD*_105	0.2663	0.1046	0.0110	0.0613	0.4714
*Fuzzy RD*_110	0.2683	0.1021	0.0090	0.0682	0.4683
*Fuzzy RD*_115	0.2671	0.0995	0.0070	0.0721	0.4621
*Fuzzy RD*_120	0.2558	0.0936	0.0060	0.0723	0.4393
*Fuzzy RD*_125	0.2547	0.0914	0.0050	0.0756	0.4338
*Fuzzy RD*_130	0.2649	0.0923	0.0040	0.0840	0.4457
*Fuzzy RD*_135	0.2600	0.0872	0.0030	0.0890	0.4309
*Fuzzy RD*_140	0.2791	0.0834	0.0010	0.1155	0.4426
*Fuzzy RD*_145	0.2651	0.0838	0.0020	0.1009	0.4292
*Fuzzy RD*_150	0.2764	0.0683	0.0000	0.1425	0.4103
*Fuzzy RD*_155	0.2777	0.0831	0.0010	0.1149	0.4406
*Fuzzy RD*_160	0.2707	0.0793	0.0010	0.1153	0.4262
*Fuzzy RD*_165	0.2797	0.0801	0.0000	0.1226	0.4367
*Fuzzy RD*_170	0.2727	0.0767	0.0000	0.1224	0.4230
*Fuzzy RD*_175	0.2844	0.0735	0.0000	0.1404	0.4285
*Fuzzy RD*_180	0.2862	0.0723	0.0000	0.1445	0.4280
*Fuzzy RD*_185	0.2774	0.0733	0.0000	0.1338	0.4210
*Fuzzy RD*_190	0.2802	0.0722	0.0000	0.1387	0.4216
*Fuzzy RD*_195	0.2907	0.0690	0.0000	0.1554	0.4260
*Fuzzy RD*_200	0.2909	0.0680	0.0000	0.1576	0.4242

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

APPENDIX D:

Table D1

Fuzzy RD estimates -Self-reported health at different bandwidth - All sample at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.1331	0.0248	0.0000	-0.1818	-0.0845
*Fuzzy RD*_50	-0.1346	0.0397	0.0010	-0.2123	-0.0568
*Fuzzy RD*_55	-0.1450	0.0340	0.0000	-0.2116	-0.0784
*Fuzzy RD*_60	-0.1251	0.0283	0.0000	-0.1806	-0.0697
*Fuzzy RD*_65	-0.1257	0.0271	0.0000	-0.1788	-0.0725
*Fuzzy RD*_70	-0.1419	0.0293	0.0000	-0.1993	-0.0845
*Fuzzy RD*_75	-0.1281	0.0252	0.0000	-0.1775	-0.0787
*Fuzzy RD*_80	-0.1443	0.0271	0.0000	-0.1975	-0.0911
*Fuzzy RD*_85	-0.1292	0.0237	0.0000	-0.1756	-0.0828
*Fuzzy RD*_90	-0.1401	0.0253	0.0000	-0.1897	-0.0904
*Fuzzy RD*_95	-0.1244	0.0223	0.0000	-0.1682	-0.0807
*Fuzzy RD*_105	-0.1206	0.0211	0.0000	-0.1620	-0.0791
*Fuzzy RD*_110	-0.1193	0.0206	0.0000	-0.1597	-0.0789
*Fuzzy RD*_115	-0.1276	0.0227	0.0000	-0.1720	-0.0832
*Fuzzy RD*_120	-0.1275	0.0213	0.0000	-0.1691	-0.0858
*Fuzzy RD*_125	-0.1181	0.0192	0.0000	-0.1558	-0.0803
*Fuzzy RD*_130	-0.1263	0.0210	0.0000	-0.1674	-0.0851
*Fuzzy RD*_135	-0.1197	0.0185	0.0000	-0.1560	-0.0835
*Fuzzy RD*_140	-0.1292	0.0164	0.0000	-0.1614	-0.0970
*Fuzzy RD*_145	-0.1299	0.0190	0.0000	-0.1672	-0.0926
*Fuzzy RD*_150	-0.1307	0.0158	0.0000	-0.1616	-0.0997
*Fuzzy RD*_155	-0.1297	0.0189	0.0000	-0.1667	-0.0927
*Fuzzy RD*_160	-0.1318	0.0152	0.0000	-0.1615	-0.1021
*Fuzzy RD*_165	-0.1322	0.0149	0.0000	-0.1613	-0.1031
*Fuzzy RD*_170	-0.1248	0.0164	0.0000	-0.1569	-0.0927
*Fuzzy RD*_175	-0.1328	0.0171	0.0000	-0.1663	-0.0994
*Fuzzy RD*_180	-0.1335	0.0140	0.0000	-0.1610	-0.1060
*Fuzzy RD*_185	-0.1340	0.0165	0.0000	-0.1663	-0.1017
*Fuzzy RD*_190	-0.1344	0.0135	0.0000	-0.1609	-0.1080
*Fuzzy RD*_195	-0.1345	0.0164	0.0000	-0.1667	-0.1024
*Fuzzy RD*_200	-0.1288	0.0149	0.0000	-0.1581	-0.0996

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table D2

Fuzzy RD estimates -Self-reported health at different bandwidth – females only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.0876	0.0311	0.0050	-0.1485	-0.0267
*Fuzzy RD*_50	-0.0912	0.0556	0.1010	-0.2002	0.0179
*Fuzzy RD*_55	-0.0884	0.0475	0.0630	-0.1816	0.0048
*Fuzzy RD*_60	-0.0843	0.0493	0.0870	-0.1809	0.0123
*Fuzzy RD*_65	-0.0810	0.0415	0.0510	-0.1623	0.0003
*Fuzzy RD*_70	-0.0808	0.0391	0.0390	-0.1574	-0.0042
*Fuzzy RD*_75	-0.0818	0.0426	0.0540	-0.1653	0.0016
*Fuzzy RD*_80	-0.0870	0.0423	0.0400	-0.1699	-0.0040
*Fuzzy RD*_85	-0.0879	0.0340	0.0100	-0.1545	-0.0212
*Fuzzy RD*_90	-0.0808	0.0343	0.0190	-0.1481	-0.0136
*Fuzzy RD*_95	-0.0898	0.0323	0.0050	-0.1531	-0.0266
*Fuzzy RD*_105	-0.0792	0.0315	0.0120	-0.1409	-0.0175
*Fuzzy RD*_110	-0.0905	0.0342	0.0080	-0.1574	-0.0235
*Fuzzy RD*_115	-0.0799	0.0298	0.0070	-0.1384	-0.0214
*Fuzzy RD*_120	-0.0877	0.0316	0.0050	-0.1496	-0.0258
*Fuzzy RD*_125	-0.0876	0.0307	0.0040	-0.1479	-0.0273
*Fuzzy RD*_130	-0.0879	0.0300	0.0030	-0.1466	-0.0291
*Fuzzy RD*_135	-0.0891	0.0293	0.0020	-0.1465	-0.0317
*Fuzzy RD*_140	-0.0907	0.0241	0.0000	-0.1380	-0.0435
*Fuzzy RD*_145	-0.0861	0.0261	0.0010	-0.1373	-0.0348
*Fuzzy RD*_150	-0.0925	0.0279	0.0010	-0.1471	-0.0379
*Fuzzy RD*_155	-0.0940	0.0273	0.0010	-0.1475	-0.0405
*Fuzzy RD*_160	-0.0956	0.0268	0.0000	-0.1481	-0.0431
*Fuzzy RD*_165	-0.1001	0.0259	0.0000	-0.1509	-0.0492
*Fuzzy RD*_170	-0.1021	0.0255	0.0000	-0.1520	-0.0521
*Fuzzy RD*_175	-0.1037	0.0251	0.0000	-0.1528	-0.0546
*Fuzzy RD*_180	-0.1056	0.0246	0.0000	-0.1538	-0.0573
*Fuzzy RD*_185	-0.1071	0.0242	0.0000	-0.1546	-0.0597
*Fuzzy RD*_190	-0.1084	0.0199	0.0000	-0.1474	-0.0693
*Fuzzy RD*_195	-0.1099	0.0234	0.0000	-0.1558	-0.0640
*Fuzzy RD*_200	-0.1053	0.0218	0.0000	-0.1481	-0.0625

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table D3:

Fuzzy RD estimates -Self-reported health at different bandwidth – males only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
ghealth	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.1738	0.0334	0.0000	-0.2392	-0.1084
*Fuzzy RD*_50	-0.1931	0.0517	0.0000	-0.2944	-0.0917
*Fuzzy RD*_55	-0.2097	0.0419	0.0000	-0.2919	-0.1275
*Fuzzy RD*_60	-0.1906	0.0457	0.0000	-0.2802	-0.1010
*Fuzzy RD*_65	-0.2066	0.0399	0.0000	-0.2847	-0.1284
*Fuzzy RD*_70	-0.1922	0.0415	0.0000	-0.2736	-0.1108
*Fuzzy RD*_75	-0.2032	0.0370	0.0000	-0.2756	-0.1308
*Fuzzy RD*_80	-0.1797	0.0323	0.0000	-0.2429	-0.1164
*Fuzzy RD*_85	-0.1920	0.0345	0.0000	-0.2596	-0.1244
*Fuzzy RD*_90	-0.1694	0.0304	0.0000	-0.2290	-0.1098
*Fuzzy RD*_95	-0.1803	0.0323	0.0000	-0.2436	-0.1171
*Fuzzy RD*_105	-0.1739	0.0271	0.0000	-0.2271	-0.1208
*Fuzzy RD*_110	-0.1717	0.0298	0.0000	-0.2301	-0.1133
*Fuzzy RD*_115	-0.1586	0.0270	0.0000	-0.2114	-0.1057
*Fuzzy RD*_120	-0.1699	0.0284	0.0000	-0.2255	-0.1143
*Fuzzy RD*_125	-0.1687	0.0277	0.0000	-0.2230	-0.1143
*Fuzzy RD*_130	-0.1676	0.0235	0.0000	-0.2136	-0.1215
*Fuzzy RD*_135	-0.1667	0.0279	0.0000	-0.2214	-0.1120
*Fuzzy RD*_140	-0.1654	0.0225	0.0000	-0.2095	-0.1212
*Fuzzy RD*_145	-0.1636	0.0256	0.0000	-0.2138	-0.1134
*Fuzzy RD*_150	-0.1619	0.0251	0.0000	-0.2112	-0.1126
*Fuzzy RD*_155	-0.1616	0.0258	0.0000	-0.2121	-0.1111
*Fuzzy RD*_160	-0.1607	0.0205	0.0000	-0.2009	-0.1204
*Fuzzy RD*_165	-0.1606	0.0201	0.0000	-0.2001	-0.1211
*Fuzzy RD*_170	-0.1525	0.0222	0.0000	-0.1959	-0.1090
*Fuzzy RD*_175	-0.1531	0.0218	0.0000	-0.1958	-0.1103
*Fuzzy RD*_180	-0.1538	0.0215	0.0000	-0.1960	-0.1117
*Fuzzy RD*_185	-0.1626	0.0223	0.0000	-0.2063	-0.1189
*Fuzzy RD*_190	-0.1634	0.0182	0.0000	-0.1990	-0.1278
*Fuzzy RD*_195	-0.1558	0.0206	0.0000	-0.1962	-0.1154
*Fuzzy RD*_200	-0.1560	0.0204	0.0000	-0.1959	-0.1162

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table D4:

Fuzzy RD estimates -Self-reported health at different bandwidth – low-educated only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.1507	0.0294	0.0000	-0.2084	-0.0930
*Fuzzy RD*_50	-0.1638	0.0532	0.0020	-0.2680	-0.0595
*Fuzzy RD*_55	-0.1414	0.0443	0.0010	-0.2282	-0.0545
*Fuzzy RD*_60	-0.1580	0.0471	0.0010	-0.2503	-0.0657
*Fuzzy RD*_65	-0.1597	0.0395	0.0000	-0.2370	-0.0824
*Fuzzy RD*_70	-0.1475	0.0388	0.0000	-0.2234	-0.0715
*Fuzzy RD*_75	-0.1633	0.0412	0.0000	-0.2440	-0.0826
*Fuzzy RD*_80	-0.1467	0.0362	0.0000	-0.2176	-0.0758
*Fuzzy RD*_85	-0.1574	0.0382	0.0000	-0.2322	-0.0825
*Fuzzy RD*_90	-0.1420	0.0339	0.0000	-0.2085	-0.0756
*Fuzzy RD*_95	-0.1525	0.0356	0.0000	-0.2223	-0.0827
*Fuzzy RD*_105	-0.1494	0.0286	0.0000	-0.2054	-0.0935
*Fuzzy RD*_110	-0.1496	0.0326	0.0000	-0.2136	-0.0856
*Fuzzy RD*_115	-0.1485	0.0316	0.0000	-0.2105	-0.0866
*Fuzzy RD*_120	-0.1480	0.0308	0.0000	-0.2083	-0.0877
*Fuzzy RD*_125	-0.1494	0.0301	0.0000	-0.2084	-0.0903
*Fuzzy RD*_130	-0.1391	0.0276	0.0000	-0.1931	-0.0850
*Fuzzy RD*_135	-0.1396	0.0270	0.0000	-0.1925	-0.0866
*Fuzzy RD*_140	-0.1396	0.0265	0.0000	-0.1915	-0.0877
*Fuzzy RD*_145	-0.1477	0.0232	0.0000	-0.1932	-0.1023
*Fuzzy RD*_150	-0.1391	0.0255	0.0000	-0.1891	-0.0891
*Fuzzy RD*_155	-0.1482	0.0265	0.0000	-0.2001	-0.0963
*Fuzzy RD*_160	-0.1467	0.0259	0.0000	-0.1975	-0.0959
*Fuzzy RD*_165	-0.1396	0.0242	0.0000	-0.1870	-0.0922
*Fuzzy RD*_170	-0.1490	0.0250	0.0000	-0.1979	-0.1000
*Fuzzy RD*_175	-0.1490	0.0245	0.0000	-0.1971	-0.1010
*Fuzzy RD*_180	-0.1470	0.0240	0.0000	-0.1940	-0.0999
*Fuzzy RD*_185	-0.1496	0.0237	0.0000	-0.1961	-0.1031
*Fuzzy RD*_190	-0.1473	0.0232	0.0000	-0.1928	-0.1018
*Fuzzy RD*_195	-0.1413	0.0219	0.0000	-0.1842	-0.0985
*Fuzzy RD*_200	-0.1493	0.0226	0.0000	-0.1936	-0.1050

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table D5:

Fuzzy RD estimates -Self-reported health at different bandwidth – high-educated only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.0669	0.0275	0.0150	-0.1209	-0.0130
*Fuzzy RD*_50	-0.0578	0.0450	0.1990	-0.1460	0.0304
*Fuzzy RD*_55	-0.0791	0.0354	0.0250	-0.1484	-0.0097
*Fuzzy RD*_60	-0.0708	0.0326	0.0300	-0.1346	-0.0069
*Fuzzy RD*_65	-0.0713	0.0313	0.0230	-0.1327	-0.0100
*Fuzzy RD*_70	-0.0721	0.0302	0.0170	-0.1312	-0.0129
*Fuzzy RD*_75	-0.0761	0.0344	0.0270	-0.1436	-0.0086
*Fuzzy RD*_80	-0.0785	0.0274	0.0040	-0.1322	-0.0248
*Fuzzy RD*_85	-0.0746	0.0304	0.0140	-0.1341	-0.0151
*Fuzzy RD*_90	-0.0646	0.0265	0.0150	-0.1167	-0.0126
*Fuzzy RD*_95	-0.0686	0.0297	0.0210	-0.1269	-0.0104
*Fuzzy RD*_105	-0.0643	0.0279	0.0210	-0.1190	-0.0095
*Fuzzy RD*_110	-0.0592	0.0239	0.0130	-0.1061	-0.0123
*Fuzzy RD*_115	-0.0651	0.0218	0.0030	-0.1079	-0.0223
*Fuzzy RD*_120	-0.0611	0.0230	0.0080	-0.1061	-0.0161
*Fuzzy RD*_125	-0.0629	0.0225	0.0050	-0.1071	-0.0188
*Fuzzy RD*_130	-0.0696	0.0238	0.0030	-0.1162	-0.0231
*Fuzzy RD*_135	-0.0711	0.0194	0.0000	-0.1092	-0.0330
*Fuzzy RD*_140	-0.0724	0.0228	0.0010	-0.1171	-0.0278
*Fuzzy RD*_145	-0.0739	0.0187	0.0000	-0.1106	-0.0372
*Fuzzy RD*_150	-0.0702	0.0205	0.0010	-0.1105	-0.0299
*Fuzzy RD*_155	-0.0764	0.0215	0.0000	-0.1185	-0.0342
*Fuzzy RD*_160	-0.0753	0.0218	0.0010	-0.1180	-0.0327
*Fuzzy RD*_165	-0.0745	0.0195	0.0000	-0.1128	-0.0362
*Fuzzy RD*_170	-0.0804	0.0204	0.0000	-0.1203	-0.0405
*Fuzzy RD*_175	-0.0769	0.0189	0.0000	-0.1139	-0.0399
*Fuzzy RD*_180	-0.0782	0.0186	0.0000	-0.1147	-0.0418
*Fuzzy RD*_185	-0.0844	0.0158	0.0000	-0.1152	-0.0535
*Fuzzy RD*_190	-0.0858	0.0155	0.0000	-0.1161	-0.0554
*Fuzzy RD*_195	-0.0826	0.0178	0.0000	-0.1175	-0.0477
*Fuzzy RD*_200	-0.0841	0.0176	0.0000	-0.1185	-0.0496

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table D6:

Fuzzy RD estimates -Self-reported health at different bandwidth – low income only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.1114	0.0281	0.0000	-0.1665	-0.0564
*Fuzzy RD*_50	-0.1157	0.0409	0.0050	-0.1959	-0.0354
*Fuzzy RD*_55	-0.1361	0.0392	0.0010	-0.2129	-0.0592
*Fuzzy RD*_60	-0.1317	0.0347	0.0000	-0.1996	-0.0637
*Fuzzy RD*_65	-0.1277	0.0356	0.0000	-0.1974	-0.0580
*Fuzzy RD*_70	-0.1270	0.0342	0.0000	-0.1939	-0.0601
*Fuzzy RD*_75	-0.1137	0.0320	0.0000	-0.1765	-0.0509
*Fuzzy RD*_80	-0.1165	0.0298	0.0000	-0.1749	-0.0581
*Fuzzy RD*_85	-0.1210	0.0308	0.0000	-0.1814	-0.0606
*Fuzzy RD*_90	-0.1177	0.0273	0.0000	-0.1713	-0.0641
*Fuzzy RD*_95	-0.1141	0.0258	0.0000	-0.1646	-0.0636
*Fuzzy RD*_105	-0.1095	0.0273	0.0000	-0.1629	-0.0561
*Fuzzy RD*_110	-0.1021	0.0252	0.0000	-0.1515	-0.0527
*Fuzzy RD*_115	-0.1067	0.0259	0.0000	-0.1574	-0.0560
*Fuzzy RD*_120	-0.1024	0.0244	0.0000	-0.1501	-0.0546
*Fuzzy RD*_125	-0.1022	0.0236	0.0000	-0.1485	-0.0560
*Fuzzy RD*_130	-0.1028	0.0233	0.0000	-0.1484	-0.0572
*Fuzzy RD*_135	-0.1035	0.0228	0.0000	-0.1482	-0.0589
*Fuzzy RD*_140	-0.1049	0.0223	0.0000	-0.1486	-0.0612
*Fuzzy RD*_145	-0.1102	0.0198	0.0000	-0.1490	-0.0714
*Fuzzy RD*_150	-0.1105	0.0224	0.0000	-0.1544	-0.0665
*Fuzzy RD*_155	-0.1067	0.0212	0.0000	-0.1483	-0.0652
*Fuzzy RD*_160	-0.1069	0.0208	0.0000	-0.1477	-0.0661
*Fuzzy RD*_165	-0.1072	0.0205	0.0000	-0.1473	-0.0670
*Fuzzy RD*_170	-0.1076	0.0201	0.0000	-0.1471	-0.0682
*Fuzzy RD*_175	-0.1080	0.0198	0.0000	-0.1468	-0.0693
*Fuzzy RD*_180	-0.1096	0.0193	0.0000	-0.1475	-0.0718
*Fuzzy RD*_185	-0.1105	0.0190	0.0000	-0.1478	-0.0732
*Fuzzy RD*_190	-0.1115	0.0187	0.0000	-0.1483	-0.0748
*Fuzzy RD*_195	-0.1146	0.0163	0.0000	-0.1466	-0.0826
*Fuzzy RD*_200	-0.1154	0.0189	0.0000	-0.1524	-0.0784

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Appendix E:

Table E1

Fuzzy RD estimates -Drinking participation at different bandwidth - All sample at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.0914	0.0274	0.0010	-0.1450	-0.0378
*Fuzzy RD*_50	-0.1174	0.0480	0.0140	-0.2115	-0.0233
*Fuzzy RD*_55	-0.1323	0.0439	0.0030	-0.2184	-0.0462
*Fuzzy RD*_60	-0.1247	0.0414	0.0030	-0.2058	-0.0436
*Fuzzy RD*_65	-0.1060	0.0348	0.0020	-0.1743	-0.0377
*Fuzzy RD*_70	-0.1049	0.0377	0.0050	-0.1788	-0.0309
*Fuzzy RD*_75	-0.1152	0.0359	0.0010	-0.1856	-0.0448
*Fuzzy RD*_80	-0.1004	0.0311	0.0010	-0.1613	-0.0395
*Fuzzy RD*_85	-0.1075	0.0266	0.0000	-0.1597	-0.0553
*Fuzzy RD*_90	-0.1007	0.0319	0.0020	-0.1633	-0.0381
*Fuzzy RD*_95	-0.1001	0.0308	0.0010	-0.1606	-0.0397
*Fuzzy RD*_105	-0.0903	0.0266	0.0010	-0.1424	-0.0382
*Fuzzy RD*_110	-0.0967	0.0281	0.0010	-0.1518	-0.0417
*Fuzzy RD*_115	-0.0954	0.0219	0.0000	-0.1382	-0.0526
*Fuzzy RD*_120	-0.0943	0.0213	0.0000	-0.1361	-0.0526
*Fuzzy RD*_125	-0.0931	0.0207	0.0000	-0.1338	-0.0525
*Fuzzy RD*_130	-0.0919	0.0202	0.0000	-0.1316	-0.0523
*Fuzzy RD*_135	-0.0908	0.0247	0.0000	-0.1393	-0.0424
*Fuzzy RD*_140	-0.0899	0.0242	0.0000	-0.1373	-0.0425
*Fuzzy RD*_145	-0.0882	0.0235	0.0000	-0.1342	-0.0422
*Fuzzy RD*_150	-0.0823	0.0218	0.0000	-0.1250	-0.0396
*Fuzzy RD*_155	-0.0866	0.0225	0.0000	-0.1308	-0.0424
*Fuzzy RD*_160	-0.0816	0.0210	0.0000	-0.1228	-0.0404
*Fuzzy RD*_165	-0.0868	0.0172	0.0000	-0.1206	-0.0531
*Fuzzy RD*_170	-0.0867	0.0215	0.0000	-0.1288	-0.0445
*Fuzzy RD*_175	-0.0818	0.0200	0.0000	-0.1209	-0.0427
*Fuzzy RD*_180	-0.0819	0.0196	0.0000	-0.1203	-0.0434
*Fuzzy RD*_185	-0.0866	0.0204	0.0000	-0.1265	-0.0467
*Fuzzy RD*_190	-0.0869	0.0157	0.0000	-0.1176	-0.0562
*Fuzzy RD*_195	-0.0871	0.0154	0.0000	-0.1173	-0.0569
*Fuzzy RD*_200	-0.0869	0.0194	0.0000	-0.1248	-0.0490

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table E2:

Fuzzy RD estimates -Drinking participation at different bandwidth - females only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.0381	0.0406	0.3480	-0.1178	0.0415
*Fuzzy RD*_50	-0.1236	0.0650	0.0570	-0.2511	0.0038
*Fuzzy RD*_55	-0.0954	0.0532	0.0730	-0.1997	0.0090
*Fuzzy RD*_60	-0.0672	0.0567	0.2360	-0.1784	0.0439
*Fuzzy RD*_65	-0.0739	0.0478	0.1220	-0.1677	0.0198
*Fuzzy RD*_70	-0.0551	0.0509	0.2790	-0.1549	0.0447
*Fuzzy RD*_75	-0.0672	0.0490	0.1700	-0.1632	0.0289
*Fuzzy RD*_80	-0.0588	0.0470	0.2120	-0.1509	0.0334
*Fuzzy RD*_85	-0.0466	0.0410	0.2560	-0.1270	0.0339
*Fuzzy RD*_90	-0.0421	0.0397	0.2890	-0.1199	0.0357
*Fuzzy RD*_95	-0.0371	0.0414	0.3700	-0.1182	0.0440
*Fuzzy RD*_105	-0.0326	0.0362	0.3680	-0.1035	0.0384
*Fuzzy RD*_110	-0.0357	0.0375	0.3400	-0.1091	0.0377
*Fuzzy RD*_115	-0.0328	0.0370	0.3750	-0.1053	0.0397
*Fuzzy RD*_120	-0.0316	0.0360	0.3800	-0.1022	0.0390
*Fuzzy RD*_125	-0.0303	0.0351	0.3880	-0.0991	0.0385
*Fuzzy RD*_130	-0.0302	0.0343	0.3790	-0.0973	0.0370
*Fuzzy RD*_135	-0.0309	0.0335	0.3560	-0.0965	0.0347
*Fuzzy RD*_140	-0.0319	0.0259	0.2180	-0.0827	0.0188
*Fuzzy RD*_145	-0.0334	0.0320	0.2980	-0.0962	0.0295
*Fuzzy RD*_150	-0.0390	0.0307	0.2050	-0.0992	0.0213
*Fuzzy RD*_155	-0.0410	0.0301	0.1730	-0.1001	0.0180
*Fuzzy RD*_160	-0.0398	0.0239	0.0950	-0.0866	0.0070
*Fuzzy RD*_165	-0.0417	0.0296	0.1590	-0.0997	0.0163
*Fuzzy RD*_170	-0.0468	0.0284	0.0990	-0.1025	0.0088
*Fuzzy RD*_175	-0.0449	0.0285	0.1150	-0.1007	0.0109
*Fuzzy RD*_180	-0.0438	0.0265	0.0980	-0.0958	0.0082
*Fuzzy RD*_185	-0.0475	0.0274	0.0840	-0.1012	0.0063
*Fuzzy RD*_190	-0.0486	0.0269	0.0720	-0.1014	0.0043
*Fuzzy RD*_195	-0.0495	0.0265	0.0620	-0.1014	0.0025
*Fuzzy RD*_200	-0.0536	0.0257	0.0370	-0.1039	-0.0033

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table E3:

Fuzzy RD estimates -Drinking participation at different bandwidth - males only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.1490	0.0397	0.0000	-0.2268	-0.0711
*Fuzzy RD*_50	-0.1398	0.0603	0.0210	-0.2581	-0.0215
*Fuzzy RD*_55	-0.1408	0.0568	0.0130	-0.2520	-0.0295
*Fuzzy RD*_60	-0.1295	0.0480	0.0070	-0.2235	-0.0355
*Fuzzy RD*_65	-0.1459	0.0525	0.0050	-0.2487	-0.0430
*Fuzzy RD*_70	-0.1509	0.0491	0.0020	-0.2471	-0.0546
*Fuzzy RD*_75	-0.1377	0.0425	0.0010	-0.2210	-0.0545
*Fuzzy RD*_80	-0.1539	0.0460	0.0010	-0.2440	-0.0639
*Fuzzy RD*_85	-0.1524	0.0355	0.0000	-0.2219	-0.0828
*Fuzzy RD*_90	-0.1522	0.0336	0.0000	-0.2180	-0.0864
*Fuzzy RD*_95	-0.1505	0.0326	0.0000	-0.2144	-0.0866
*Fuzzy RD*_105	-0.1468	0.0385	0.0000	-0.2222	-0.0713
*Fuzzy RD*_110	-0.1443	0.0374	0.0000	-0.2175	-0.0710
*Fuzzy RD*_115	-0.1403	0.0363	0.0000	-0.2115	-0.0692
*Fuzzy RD*_120	-0.1407	0.0351	0.0000	-0.2094	-0.0719
*Fuzzy RD*_125	-0.1368	0.0342	0.0000	-0.2039	-0.0697
*Fuzzy RD*_130	-0.1325	0.0335	0.0000	-0.1980	-0.0669
*Fuzzy RD*_135	-0.1283	0.0260	0.0000	-0.1792	-0.0774
*Fuzzy RD*_140	-0.1213	0.0321	0.0000	-0.1843	-0.0583
*Fuzzy RD*_145	-0.1191	0.0315	0.0000	-0.1808	-0.0574
*Fuzzy RD*_150	-0.1136	0.0289	0.0000	-0.1702	-0.0569
*Fuzzy RD*_155	-0.1166	0.0302	0.0000	-0.1757	-0.0574
*Fuzzy RD*_160	-0.1159	0.0296	0.0000	-0.1739	-0.0579
*Fuzzy RD*_165	-0.1110	0.0273	0.0000	-0.1645	-0.0575
*Fuzzy RD*_170	-0.1163	0.0224	0.0000	-0.1602	-0.0724
*Fuzzy RD*_175	-0.1148	0.0278	0.0000	-0.1693	-0.0603
*Fuzzy RD*_180	-0.1075	0.0260	0.0000	-0.1584	-0.0566
*Fuzzy RD*_185	-0.1061	0.0256	0.0000	-0.1562	-0.0560
*Fuzzy RD*_190	-0.1045	0.0252	0.0000	-0.1538	-0.0551
*Fuzzy RD*_195	-0.1084	0.0262	0.0000	-0.1597	-0.0571
*Fuzzy RD*_200	-0.1059	0.0199	0.0000	-0.1449	-0.0669

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table E4:

Fuzzy RD estimates -Drinking intensity at different bandwidth -All sample at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.0479	0.0191	0.0120	-0.0854	-0.0105
*Fuzzy RD*_50	-0.0679	0.0311	0.0290	-0.1289	-0.0069
*Fuzzy RD*_55	-0.0566	0.0322	0.0790	-0.1197	0.0066
*Fuzzy RD*_60	-0.0643	0.0346	0.0630	-0.1321	0.0034
*Fuzzy RD*_65	-0.0641	0.0327	0.0500	-0.1282	-0.0001
*Fuzzy RD*_70	-0.0618	0.0241	0.0100	-0.1091	-0.0145
*Fuzzy RD*_75	-0.0502	0.0290	0.0840	-0.1071	0.0067
*Fuzzy RD*_80	-0.0487	0.0256	0.0580	-0.0989	0.0016
*Fuzzy RD*_85	-0.0451	0.0267	0.0910	-0.0974	0.0072
*Fuzzy RD*_90	-0.0485	0.0264	0.0660	-0.1003	0.0033
*Fuzzy RD*_95	-0.0437	0.0248	0.0780	-0.0923	0.0049
*Fuzzy RD*_105	-0.0478	0.0238	0.0440	-0.0945	-0.0012
*Fuzzy RD*_110	-0.0482	0.0230	0.0370	-0.0933	-0.0030
*Fuzzy RD*_115	-0.0492	0.0224	0.0280	-0.0930	-0.0054
*Fuzzy RD*_120	-0.0444	0.0211	0.0360	-0.0859	-0.0030
*Fuzzy RD*_125	-0.0466	0.0196	0.0170	-0.0851	-0.0082
*Fuzzy RD*_130	-0.0502	0.0157	0.0010	-0.0809	-0.0195
*Fuzzy RD*_135	-0.0427	0.0196	0.0290	-0.0810	-0.0043
*Fuzzy RD*_140	-0.0485	0.0149	0.0010	-0.0778	-0.0192
*Fuzzy RD*_145	-0.0478	0.0146	0.0010	-0.0765	-0.0192
*Fuzzy RD*_150	-0.0441	0.0176	0.0120	-0.0787	-0.0096
*Fuzzy RD*_155	-0.0465	0.0184	0.0120	-0.0826	-0.0104
*Fuzzy RD*_160	-0.0405	0.0176	0.0210	-0.0750	-0.0060
*Fuzzy RD*_165	-0.0406	0.0172	0.0190	-0.0744	-0.0067
*Fuzzy RD*_170	-0.0456	0.0133	0.0010	-0.0716	-0.0196
*Fuzzy RD*_175	-0.0449	0.0171	0.0080	-0.0784	-0.0115
*Fuzzy RD*_180	-0.0440	0.0167	0.0090	-0.0768	-0.0112
*Fuzzy RD*_185	-0.0388	0.0160	0.0160	-0.0702	-0.0073
*Fuzzy RD*_190	-0.0403	0.0153	0.0080	-0.0703	-0.0103
*Fuzzy RD*_195	-0.0419	0.0121	0.0010	-0.0656	-0.0183
*Fuzzy RD*_200	-0.0378	0.0153	0.0130	-0.0677	-0.0078

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table E5:

Fuzzy RD estimates -Drinking intensity at different bandwidth -females only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf.	Interval]
Fuzzy RD	-0.0142	0.0272	0.6010	-0.0676	0.0392
*Fuzzy RD*_50	-0.0519	0.0418	0.2140	-0.1338	0.0300
*Fuzzy RD*_55	-0.0483	0.0445	0.2780	-0.1355	0.0389
*Fuzzy RD*_60	-0.0406	0.0318	0.2020	-0.1029	0.0217
*Fuzzy RD*_65	-0.0362	0.0299	0.2260	-0.0949	0.0224
*Fuzzy RD*_70	-0.0266	0.0351	0.4470	-0.0954	0.0421
*Fuzzy RD*_75	-0.0232	0.0334	0.4880	-0.0887	0.0423
*Fuzzy RD*_80	-0.0208	0.0261	0.4250	-0.0719	0.0303
*Fuzzy RD*_85	-0.0159	0.0290	0.5840	-0.0726	0.0409
*Fuzzy RD*_90	-0.0156	0.0234	0.5050	-0.0615	0.0303
*Fuzzy RD*_95	-0.0147	0.0282	0.6030	-0.0700	0.0406
*Fuzzy RD*_105	-0.0147	0.0210	0.4850	-0.0558	0.0265
*Fuzzy RD*_110	-0.0151	0.0265	0.5680	-0.0670	0.0368
*Fuzzy RD*_115	-0.0144	0.0237	0.5430	-0.0609	0.0320
*Fuzzy RD*_120	-0.0147	0.0231	0.5230	-0.0599	0.0305
*Fuzzy RD*_125	-0.0131	0.0234	0.5740	-0.0590	0.0327
*Fuzzy RD*_130	-0.0154	0.0236	0.5140	-0.0617	0.0309
*Fuzzy RD*_135	-0.0148	0.0230	0.5210	-0.0599	0.0304
*Fuzzy RD*_140	-0.0146	0.0174	0.4010	-0.0487	0.0195
*Fuzzy RD*_145	-0.0145	0.0220	0.5120	-0.0576	0.0287
*Fuzzy RD*_150	-0.0123	0.0209	0.5560	-0.0531	0.0286
*Fuzzy RD*_155	-0.0149	0.0211	0.4800	-0.0564	0.0265
*Fuzzy RD*_160	-0.0134	0.0200	0.5050	-0.0526	0.0259
*Fuzzy RD*_165	-0.0144	0.0197	0.4650	-0.0529	0.0241
*Fuzzy RD*_170	-0.0152	0.0188	0.4190	-0.0521	0.0217
*Fuzzy RD*_175	-0.0160	0.0149	0.2830	-0.0451	0.0132
*Fuzzy RD*_180	-0.0147	0.0182	0.4190	-0.0503	0.0210
*Fuzzy RD*_185	-0.0150	0.0183	0.4130	-0.0508	0.0208
*Fuzzy RD*_190	-0.0146	0.0176	0.4050	-0.0491	0.0198
*Fuzzy RD*_195	-0.0154	0.0177	0.3830	-0.0501	0.0192
*Fuzzy RD*_200	-0.0155	0.0136	0.2550	-0.0422	0.0112

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

Table E6:

Fuzzy RD estimates -Drinking intensity at different bandwidth - males only at cutoff age 65

Regression discontinuity robustness of bandwidth choice					
	RD estimates	Bootstrap Std. Err.	P-value	[95% Conf. Interval]	
Fuzzy RD	-0.0665	0.0373	0.0750	-0.1396	0.0066
*Fuzzy RD*_50	-0.0631	0.0467	0.1770	-0.1546	0.0285
*Fuzzy RD*_55	-0.0598	0.0545	0.2720	-0.1665	0.0469
*Fuzzy RD*_60	-0.0672	0.0455	0.1400	-0.1564	0.0220
*Fuzzy RD*_65	-0.0757	0.0379	0.0460	-0.1500	-0.0015
*Fuzzy RD*_70	-0.0708	0.0466	0.1290	-0.1621	0.0206
*Fuzzy RD*_75	-0.0665	0.0351	0.0580	-0.1353	0.0024
*Fuzzy RD*_80	-0.0653	0.0329	0.0470	-0.1298	-0.0008
*Fuzzy RD*_85	-0.0646	0.0414	0.1190	-0.1457	0.0165
*Fuzzy RD*_90	-0.0586	0.0395	0.1380	-0.1360	0.0188
*Fuzzy RD*_95	-0.0655	0.0298	0.0280	-0.1238	-0.0072
*Fuzzy RD*_105	-0.0601	0.0358	0.0930	-0.1302	0.0100
*Fuzzy RD*_110	-0.0689	0.0351	0.0500	-0.1376	-0.0001
*Fuzzy RD*_115	-0.0642	0.0316	0.0420	-0.1262	-0.0023
*Fuzzy RD*_120	-0.0633	0.0308	0.0400	-0.1237	-0.0029
*Fuzzy RD*_125	-0.0572	0.0319	0.0730	-0.1196	0.0053
*Fuzzy RD*_130	-0.0560	0.0311	0.0720	-0.1170	0.0049
*Fuzzy RD*_135	-0.0595	0.0288	0.0390	-0.1160	-0.0031
*Fuzzy RD*_140	-0.0622	0.0301	0.0390	-0.1212	-0.0033
*Fuzzy RD*_145	-0.0615	0.0225	0.0060	-0.1057	-0.0173
*Fuzzy RD*_150	-0.0524	0.0285	0.0660	-0.1082	0.0034
*Fuzzy RD*_155	-0.0515	0.0279	0.0650	-0.1062	0.0032
*Fuzzy RD*_160	-0.0506	0.0273	0.0640	-0.1042	0.0030
*Fuzzy RD*_165	-0.0536	0.0256	0.0370	-0.1038	-0.0034
*Fuzzy RD*_170	-0.0525	0.0251	0.0370	-0.1018	-0.0032
*Fuzzy RD*_175	-0.0543	0.0198	0.0060	-0.0931	-0.0155
*Fuzzy RD*_180	-0.0470	0.0254	0.0640	-0.0967	0.0027
*Fuzzy RD*_185	-0.0458	0.0249	0.0660	-0.0947	0.0030
*Fuzzy RD*_190	-0.0501	0.0186	0.0070	-0.0866	-0.0136
*Fuzzy RD*_195	-0.0437	0.0241	0.0700	-0.0909	0.0036
*Fuzzy RD*_200	-0.0427	0.0237	0.0720	-0.0893	0.0038

Notes: (i) Each row is a separate non-parametric regression discontinuity point estimate. (ii) the first row reports the Fuzzy RD estimate at the CV-optimal bandwidth (iii) Fuzzy non-parametric RD treatment effect is estimated for 50 to 200 percent of the CV-optimal bandwidth (half and twice the optimal bandwidth) with a 5 percentage points incremental. (iv) Bootstrapped standard error are based on 120 simulations to construct the 95% CI.

1.12. References

1. Adam, S., Bonsang, E., Germain, S., and Perelman, S. (2007). Retirement and Cognitive Reserve: A Stochastic Frontier Approach Applied to Survey Data. CREPP working papers 2007/04. http://www.crepp.ulg.ac.be/papers/crepp-wp200704.pdf
2. Anderson, N. D., and Craik, F.I.M. (2000). Memory in the Aging Brain, In The Oxford Handbook of Memory, ed. E. Tulving and F. I. M. Craik, 411–425. Oxford: Oxford University Press.
3. Angrist, J. D., & Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. Quarterly Journal Of Economics, 114(2), 533-575.
4. Atchley, Robert C, and Judith L Robinson. (1982). Attitudes toward retirement and distance from the event." Research on Aging 4 (3): 299–313.
5. Ayyagari, P., and Sindelar, J.L. (2010). The Impact of Job Stress on Smoking and Quitting: Evidence from the HRS. The BE Journal of Economic Analysis & Policy, 10(1).
6. Banks, J., and Mazzonna, F. (2012). 'The effect of education on old age cognitive abilities: evidence from a regression discontinuity design. The Economic Journal. 122 (560), 418–448.
7. Barefoot, J. C., & Schroll, M. (1996). Symptoms of depression, acute myocardial infarction, and total mortality in a community sample. Circulation, 93(11), 1976-1980.
8. Bartalotti, O. and Brummet, Q. (2017). Regression discontinuity designs with clustered data. In Cattaneo, M. D. and Escanciano, J. C., editors, Regression Discontinuity Designs: Theory and Applications (Advances in Econometrics, volume 38). Emerald Group Publishing, forthcoming.

9. Behncke, S. (2009). How Does Retirement Affect Health? IZA Discussion Paper series, IZA DP, Institute for the Study of Labor, No. 4253. http://ftp.iza.org/dp4253.pdf

10. Behnke, S. (2012). "Does Retirement Trigger Ill Health?" Health Economics, 21(3), 282-300.

11. Belgrave LL, Haug MR, Gomez-Bellenge FX. (1987). Gender and race differences in effects of health and pension on retirement before 65. Comprehensive Gerontology [B]. 1(3):109-17.

12. Benitez-Silva, H., Buchinsky, M., Chan, H.M., Cheidvasser, S., and Rust, J. P. (2004). How Large is the Bias is Self-Reported Disability?. *Journal of Applied Econometrics*. 19, 649-670.

13. Benjamin, D., S. Brown, and J. Shapiro (2006). 'Cognitive Ability and Anomalous Preferences,' mimeo, Harvard University.

14. Black, N., Johnston, D.W., Suziedelyte, G. (2017). Justification bias in self-reported disability: New evidence from panel data. *Journal of Health Economics*, 54, 124-134.

15. Black, S. E. (1999). Do Better Schools Matter? Parental Valuation of Elementary Education. Quarterly Journal of Economics, 114(2), 577-599.

16. Bloemen, H., & Zweerink, J. (2013). The Causal Effect of Retirement on Mortality: Evidence from Targeted Incentives to Retire Early. IZA Discussion paper series, IZA DP NO. 7570

17. Bond, J. (1991). Self-Reported vs. Objective Measures of Health in Retirement Models. *Journal of Human Resources.* 26, 106-38.

18. Bonsang, E., Adam, S., & Perelman, S. (2012). Does Retirement Affect Cognitive Functioning?. Journal of Health Economics, 31(3), 490-501.

19. Börsch-Supan, A., and M. Schuth (2013). 'Early Retirement, Mental Health and Social Networks,' In Wise D.A. (ed.), Discoveries in the Economics of Aging, University of Chicago Press.

20. Bossé, Raymond, Carolyn M Aldwin, Michael R Levenson, and David J Ekerdt. (1987). Mental health differences among retirees and workers: findings from the Normative Aging Study. Psychology and Aging 2 (4): 383.

21. Bound, J. and Waidmann, T. (2007). 'Estimating the Health Effects of Retirement. University of Michigan Retirement Research Center working paper, WP 2007-168.

22. Bound, J., Schoenbaum, M., Stinebrickner, T. R., & Waidmann, T. (1998). The Dynamic Effects of Health on the Labor Force Transitions of Older Workers.

23. Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2016). On the effect of bias estimation on coverage accuracy in nonparametric inference. arXiv:1508.02973.

24. Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2018a). Coverage Error Optimal Confidence Intervals for Regression Discontinuity Designs. working paper, University of Michigan.

25. Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust data-driven inference in the regression-discontinuity design. The Stata Journal. 14(4), 909–946.

26. Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica, 82(6):2295–2326.

27. Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017a). Regression Discontinuity Designs Using Covariates. working paper, University of Michigan.

28. Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2017b). rdrobust: Software for Regression Discontinuity Designs. Stata Journal,17, 372-404.

29. Calonico, S., Cattaneo, M.D & Titiunik, R. (2014c). Notes and Comments Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. Econometrica, 82(6), 2295–2326

30. Calonico, S., Cattaneo, M.D., & Titiunik, R. (2015) Optimal Data-Driven Regression Discontinuity Plots. Journal of the American Statistical Association. 110(512), 1753-1769

31. Calonico, S., Cattaneo, M.D., Farrell, M.H. (2018). Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs.

32. Calvo, Esteban & Haverstick, Kelly & Sass, Steven, 2007. "What Makes Retirees Happier: A Gradual or 'Cold Turkey' Retirement?," MPRA Paper 5607, University Library of Munich, Germany.

33. Calvo, Esteban, Natalia Sarkisian, and Christopher Tamborini. (2013). "Causal Effects of Retirement Timing on Subjective Physical and Emotional Health." Journal of Gerontology Series B: Psychological Sciences and Social Sciences 68(1):73-84.

34. Card, D., and Pei, Z., Lee, D.S., and Weber, A. (2041). Local Polynomial Order in Regression Discontinuity Designs. University of California, Berkeley, department of economics, Working Paper. http://davidcard.berkeley.edu/papers/local-polynomial-2014.pdf

35. Card, D., C. Dobkin and N. Maestas, 2008, "The impact of nearly universal coverage on health care utilization: evidence from Medicare", American Economic Review, 98(5), 2242-2258.

36. Card, D., C. Dobkin and N. Maestas, 2009, "Does Medicare save lives?", Quarterly Journal of Economics, 124(2): 597-636.

37. Card, D., Dobkin, C., and Maestas, N. (2014). the impact of nearly universal insurance coverage on health care utilization and health: evidence from medicare. NBER Working Paper Series 10365. March 2004.

38. Cattaneo, M. D., and J. C. Escanciano, eds. (2017). Advances in Econometrics: Vol. 38 Regression Discontinuity Designs: Theory and Applications. Bingley, UK: Emerald.

39. Cattaneo, M. D., Frandsen, B., and Titiunik, R. (2015). Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. Journal of Causal Inference. 3, 1-24.

40. Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2018a). A Practical Introduction to Regression Discontinuity Designs: Part II, In preparation for Cambridge Elements: Quantitative and Computational Methods for Social Science, Cambridge University Press.

41. Cattaneo, M. D., Idrobo, N., and Titiunik, R. (2018b). rddensity: Manipulation Testing based on Density Discontinuity. Stata Journal, forthcoming.

42. Cattaneo, M. D., Jansson, M., and Ma, X. (2017a). Simple Local Polynomial Density Estimators. working paper, University of Michigan.

43. Cattaneo, M. D., R. Titiunik, and G. Vazquez-Bare. 2016. Inference in regression discontinuity designs under local randomization. Stata Journal 16: 331–367.

44. Cattaneo, M.D., and Vazquez-Bare, G. (2016). The Choice of Neighborhood in Regression Discontinuity Designs. Observational Studies, 2, 134-146

45. Cattaneo, M.D., Idrobo, N., and Titiunik, R. (2018). A Practical Introduction to Regression Discontinuity Designs: Volume I. Monograph prepared for Cambridge Elements: Quantitative and Computational Methods for Social Science. Cambridge University Press. Cambridge Elements: Quantitative and Computational Methods for Social Science.

Cambridge University Press. http://www.cambridge.org/us/academic/elements/quantitative-and-computational-methods-social-science

46. Cattaneo, M.D., Jansson, M., and Ma, X. (2018). Manipulation testing based on density discontinuity. The Stata Journal. 18(1), 234-261.
47. Cerulli, G. (2015). Econometric Evaluation of Socio-Economic Programs, Theory and Applications, Advanced Studies in Theoretical and Applied Econometrics 49. Springer-Verlag Berlin Heidelberg. Retrieved from http://www.springer.com/us/book/9783662464045.
48. Charles, K. K. (2004). 'Is retirement depressing? Labor force inactivity and psychological well-being in later life,' Research in Labor Economics, 23: 269–299.
49. Chay, K. Y., & Greenstone, M. (2005). Does Air Quality Matter? Evidence from the Housing Market. Journal Of Political Economy, 113(2), 376-424.
50. Chay, K. Y., McEwan, P. J., & Urquiola, M. (2005). The Central Role of Noise in Evaluating Interventions that Use Test Scores to Rank Schools. American Economic Review, 95(4), 1237-1258.
51. Cheng, M.-Y., Fan, J., and Marron, J.S. (1997). On Automatic Boundary Corrections. The Annals of Statistics, 25 (4), 1691–1708.
52. Christelis, D., Georgarakos, D., Sanz-de-Galdeano, A. (2014).The Impact of Health Insurance on Stockholding: A Regression Discontinuity Approach
53. Coe, N. B., & Zamarro, G. (2011). Retirement Effects on Health in Europe. Journal Of Health Economics, 30(1), 77-86.

54. Coe, N., & Lindeboom, M. (2008). Does Retirement Kill You. Evidence from Early Retirement Windows. IZA Discussion Paper 3817. DOI:10.2139/ssrn.1295315

55. Coe, N., H. von Gaudecker, M. Lindeboom, and J. Maurer (2012). The effect of retirement on cognitive functioning. Health Economics 21, 913-927

56. Coe, Norma B, and Gema Zamarro. 2011. Retirement Effects on Health in Europe. Journal of Health Economics 30 (1): 77–86.

57. Coe, Norma B, Hans-Martin von Gaudecker, Maarten Lindeboom, and Jürgen Maurer. (2012). The effect of retirement on cognitive functioning. Health Economics. 21 (8): 913–927.

58. Cousins, R.C., & Tang, N. (2004). Working time and work and family conflict in the Netherlands, Sweden, and the U.K. Work, Employment & Society, 18(3), 531-549.

59. Currie, J., and Madrian, B.C. (1999). Health, Health Insurance and the Labour Market. In Ashenfelter, O, and D. Card, D. (eds.), Handbook of Labour Economics v. 3c, New York: Elsevier, 3309-3416.

60. Dave, D., I. Rashad and J. Spasojevic (2008). The Effects of Retirement on Physical and Mental Health Outcomes. Southern Economic Journal, 75, 497-523.

61. DiNardo, J., & Lee, D. S. (2004). Economic Impacts of New Unionization on Private Sector Employers: 1984-2001. Quarterly Journal of Economics, 119(4), 1383-1441.

62. Disney, R., Emmerson, C., & Wakefield, M. (2006). Ill Health and Retirement in Britain: A Panel Data-Based Analysis. Journal of Health Economics, 25(4), 621-649.

63. Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2007). Are Risk Aversion and Impatience Related to Cognitive Ability? IZA Discussion Paper No. 2735.

64. Eibich, P. (2014). Understanding the effect of retirement on health using Regression Discontinuity Design. Retrieved from Available at SSRN: http://ssrn.com/abstract=2463240 or http://dx.doi.org/10.2139/ssrn.2463240

65. Ekerdt, D. J., R. Bosse, and J. S. Locastro (1983). Claims that retirement improves health. Journal of Gerontology, 38, 231–236.

66. Ettner SL, Frank RG, Kessler RC. (1997). The Impact of Psychiatric Disorders on Labor Market Outcomes. Industrial and Labor Relations Review. 51(1):64-81.

67. Evans D.A., L.A. Beckett, M.S. Albert, et al. (1993). 'Level of education and change in cognitive function in a community population of older persons,' Ann Epidemiol 3, 71-7.

68. Evenson, Kelly R., Wayne D. Rosamond, Jianwen Cai, Ana V. Diez-Roux, and Frederick L. Brancati. (2002). "Influence of Retirement on Leisure-Time Physical Activity. The Atherosclerosis Risk in Communities Study." American Journal of Epidemiology 155(8): 692-699.

69. Farmer, M M., & Ferraro K.F. (1997). Distress and Perceived Health: Mechanisms of Health Decline. Journal of Health and Social Behavior, 38(3), 298-311.

70. Fitzpatrick, M., and Moore, T., (2018). The mortality effects of retirement: Evidence from Social Security eligibility at age 62. Journal of Public Economics. 157, 121- 137.

71. Flicker, L. (2009). 'Life Style Interventions to Reduce the Risk of Dementia. Maturitas 63, 319-322.

72. Fratiglioni, L., S. Paillard-Borg, B. Winblad (2004). An Active and Socially Integrated Lifestyle in late Life might Protect against Dementia. The Lancet, Neurology 3, 343-353.

73. Frederick, Shane. 2005. Cognitive Reflection and Decision Making. Journal of Economic Perspectives, 19(4): 25-42.

74. Gelman, A., and Imbens, G. (2014). Why High-Order Polynomials Should Not Be Used In Regression Discontinuity Designs. NBER Working Paper Series, Working Paper 14726, August 2014. http://www.nber.org/papers/w20405.pdf

75. George, L.K., Maddox, G.L. (1977). Subjective adaptation to loss of the work role: a longitudinal study. Journal of Gerontology. 32, 456–462.

76. Godard, M. (2013). Gaining weight through retirement? Results from the SHARE survey. University of Paris Dauphine Mimeo, (November), 1–20. Retrieved from http://www.iza.org/conference_files/older_workers_2013/godard_m8733.pdf

77. Goldman, D., Lakdawalla, D. and Zheng, Y., 2008. Retirement and weight, mimeo, available online via http://www.aeaweb.org/assa/2009/retrieve.php?pdfid=219, last accessed February 20, 2015.

78. Gorry, Aspen, Devon Gorry, and Sita Slavov. 2016. Does retirement improve health and life satisfaction?" NBER Working Paper.

79. Grâce, Gaston-René de, Purushottam Joshi, René Pelletier, and Claudine Beaupré. 1994. Consequences psychologiques de la retraite en fonction du sexe et du niveau occupationnel anterieur." Canadian Journal on Aging/La Revue canadienne du vieillissement. 13 (02): 149–168.

80. Grady, C.L., Horwitz, B., Pietrini, P., Mentis, M.J., Ungerleiter, L., Rapoport, S.I., & Haxby, J. (1996). The effect of task difficulty on cerebral blood flow during perceptual matching of faces. Human Brain Mapping, 4, 227–239.

81. Grasby, P.M., Frith, C.D., Friston, K.J., Simpson, J.F.P.C., Frack-owiak, R.S.J., & Dolan, R.J. (1994). A graded task approach to functional mapping of areas implicated in auditory-verbal memory. Brain, 117, 1271–1282.

82. Grip, De. A., Dupuy, A., Jolles, J., and Boxtel, M.V. (2015). Retirement and cognitive development in the Netherlands: Are the retired really inactive? Economics and Human Biology. 19, 157- 169.

83. Grossman M, Kaestner R. (1997). Effects of education on health. The Social Benefits of Education. Behrman JR, Stacey N (eds). University of Michigan Press: Ann Arbor. 69-123.

84. Gur, R.C., Gur, R.E., Skolnick, B.E., Resnick, S.M., Silver, F.L., Chawluk, J.M.L., Obrist, W.D., & Reivich, M. (1988). Effects of task difficulty on regional cerebral blood flow: Relationships with anxiety and performance. Psychophysiology, 25, 392–399.

85. Hachinski, V. (1999). Post-stroke depression, not to be underestimated. The Lancet, May 22, 353(9166), 1728.

86. Hahn, J., P. Todd, and W. Van der Klaauw (2001). Identification and estimation of treatment effects with a regression-discontinuity design. Econometrica, 69 (1), 201-209.

87. Hahn, J., Todd, P., Van Der Klaauw, W., 1999. Evaluating the effect of an anti-discrimination law using a regression-discontinuity design. NBER Working Paper No. 7131.

88. Health and Retirement Study Staff. (2017). Sample Sizes and Response Rates. Survey Research Center, Institute for Social Research, University of Michigan. Report 9042 https://hrs.isr.umich.edu/sites/default/files/biblio/ResponseRates_2017.pdf

89. Henkens, K., Van Solinge, H., & Gallo, W. T. (2008). Effects of retirement voluntariness on changes in smoking, drinking and physical activity among Dutch older workers. European Journal of Public Health, 18(6), 644–649. DOI:10.1093/eurpub/ckn095

90. Hertzog C., Kramer, A.F., Wilson, R.B., and Lindenberger, U. (2008). Enrichment effects on adult cognitive developments. Can the functional capacity of older adults be preserved and enhanced? Psychological Science in the Public Interest. 9, 1–65.

91. Herzog, A.R., and Wallace, R.B. (1997). Measures of cognitive functioning in the AHEAD study, Journals of Gerontology Series B, 52, 37–48.

92. HHS, (2018). Annual Update of the HHS Poverty Guidelines. Health and Human Services Department, Federal Register/Vol. 83, No. 12/Thursday, January 18, 2018/Notices, document number:2018-00814,2642-2644.https://www.gpo.gov/fdsys/pkg/FR-2018-01-18/pdf/2018-00814.pdf

93. Hultsch, D.F., Hertzog, C., Small, B.J., Dixon, R.A. 1999. Psychology and Aging, 14(2), 245-263

94. Hurd, Michael and Kathleen McGarry (1993). "The Relationship Between Job Characteristics and Retirement," NBER Working Paper No. 4558.

95. Imbens, G. W., & Lemieux, T. (2008). Regression Discontinuity Designs: A Guide to Practice. Journal Of Econometrics, 142(2), 615-635.

96. Imbens, G., and Kalyanaraman, K. (2009). Optimal Bandwidth Choice for The Regression Discontinuity Estimator. NBER Working Paper Series, Working Paper 14726. http://www.nber.org/papers/w14726.pdf

97. Insler, M., 2014. The Health Consequences of Retirement, Journal of Human Resources, 49(1), 195-233.

98. Jacob, R., Zhu, P., Somers, M., & Bloom, H. (2012). A Practical Guide to Regression Discontinuity, (July). Retrieved from http://www.mdrc.org/publications/644/full.pdf

99. Jenkins, K. R., Ofstedal, M. B., Weir, D., Wallace, R. B., Crimmins, E. M., Faul, J. D., … Arbor, A. (2008). HRS / AHEAD Documentation Report Report prepared by, (February). Journal, E. (2001). Economic Journal. The New York Times, 9(4), 2002–2003.

100. Johnston, David W., and Wang-Sheng Lee. 2009. Retiring to the Good Life? The Short Term Effects of Retirement on Health." Economics Letters.

101. Jones, A. M., Rice, N., & Roberts, J. (2010). Sick of Work or Too Sick to Work? Evidence on Self-Reported Health Shocks and Early Retirement from the BHPS. Economic Modelling, 27(4), 866-880.

102. Kajitani, S., Sakata, K., & Mckenzie, C. R. (2013). Does the Definition of Retirement Matter in Estimating the Effects of Retirement on Cognitive Functioning ?, (December), 1–6.

103. Kämpfen, F., and Maurer, J. (2016). Time to burn (calories)? The impact of retirement on physical activity among mature Americans. Journal of Health Economics. 45: 91-102.

104. Karoly, Lynn A. and Jeannette A. Rogowski (1994). the Effect of Access to Post-Retirement Health Insurance on the Decision to Retire Early." Industrial and Labor Relations Review, 48(1), 103-123.

105. Kunst, A. E., Bos, V., Lahelma, E., Bartley, M., Lissau, I., Regidor, E., Mielck, A., Cardano, M., Dalstra, J. A., Geurts, J. M., Helmert, U., Lennartsson, C., Ramm, J., Spadea, T., Stronegger, W. J., Mackenbach, J.P. (2005). Trends in socioeconomic inequalities in self-assessed health in 10 European countries. International Journal of Epidemiology, 34, 295-305.

106. Latif, E. (2012). The impact of retirement on health in Canada. Canadian Public Policy, 38(1), 15–29. DOI:10.3138/cpp.38.1.15

107. Latif, E. (2013). The impact of retirement on mental health in Canada. Journal of Mental Health Policy and Economics, 16, 35–46.

108. Lechner, M., & Vazquez-Alvarez, R. (2011). The Effect of Disability on Labour Market Outcomes in Germany. Applied Economics, 43(4-6), 389-412.

109. Lee, D. S. (2008). Randomized Experiments from Non-random Selection in U.S. House Elections. Journal Of Econometrics, 142(2), 675-697.

110. Lee, D. S., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. Journal Of Economic Literature, 48(2), 281-355.

111. Mackenbach, J.P., Simon, J.G., Looman, C. WN., and Joung, I. MA. (2002). Self-assessed health and mortality: could psychosocial factors explain the association?. International Journal of Epidemiology, 31 (6), 1162-1168.

112. Madrian, Brigitte C. (1994). Employment-Based Health Insurance and Job Mobility: Is There Evidence of Job Lock?. Quarterly Journal of Economics. 109(1), 27-54.

113. Mannheim, B., Baruch, Y., & Tal, J. (1997). Alternative models for antecedents and outcomes of work centrality and job satisfaction of high-tech personnel. Human Relations, 50(12), 1537-1562.

114. Maurer, J. (2010). Height, Education and Cognitive Function at Older Ages: International Evidence from Latin America and the Caribbean. Economics and Human Biology, 8: 168-176.

115. Mazzonna, F., and Peracchi, F. (2012). Ageing, cognitive abilities and retirement. European Economic Review, 56, 691-710.

116. Mazzonna, F., and Peracchi, F. (2012). Ageing, Cognitive Abilities and Retirement. European Economic Review. 56(4), 691-710.

117. McFadden, D. (2008), 'Human Capital Accumulation and Depreciation,' Applied Economic Perspectives and Policy, vol. 30(3), 379-385.

118. McGarry K. (2004). Health and retirement: do changes in health affect retirement expectations? Journal of Human Resources. 39(3):624-48.

119. Midanik, L. T., Soghikian, K., Ransom, L. J., & Tekawa, I. S. (1995). The effect of retirement on mental health and health behaviors: the Kaiser Permanente Retirement Study. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 50 B(1), 59–S61.

120. Motegi, Hiroyuki, Yoshinori Nishimura, and Kazuyuki Terada. 2016. Does Retirement Change Lifestyle Habits?" The Japanese Economic Review 67 (2): 169–191.

121. Müller, Tobias, and Mujaheed Shaikh. 2017. Your Retirement and My Health Behavior: Evidence on Retirement Externalities from a fuzzy regression discontinuity design. Journal of Health Economics.

122. Neuman K. (2008). 'Quit your job and get healthier? The effect of retirement on health,' Journal of Labor Research, 29: 177–201.

123. Neuman, K. (2008). Quit your job and get healthier? The effect of retirement on health. Journal of Labor Research 29, 177-201.

124. Neuman, K., & Davis, J. R. (2015). Preliminary, Please to Not Cite Miracle Drug Or Daily Vitamin? The Health Effects of Retirement over Time Prepared for the Labor and Employment Relations/ASSA 2015 Meeting, (715).

125. Nolan-Hoeksema S, Larson J, and Grayson C. (1999). Explaining the gender difference in depressive symptoms. Journal of Personality and Social Psychology. 77(5):1061-1072.

126. Ofstedal, M.B., Fisher, G.G., Herzog, A.R. (2005). Documentation of cognitive functioning measures in the health and retirement study. The HRS Health Working Group, Survey Research Center, University of Michigan, HRS Documentation Report DR-006-March, 2005.

127. Otsu, T., Xu, Ke-Li., and Matsushita, Y. (2015). Empirical likelihood for regression discontinuity design. Journal of Econometrics, 186, 94-112.

128. Physical Activity Guidelines Advisory Committee. (2008). Physical Activity Guidelines Advisory Committee Report. U.S. Department of Health and Human Services, Washington, DC

129. Porter, J. (2003). Estimation in the regression discontinuity model. Dep. Econ. Univ. Wis. Mimeo.

130. Portnoi, VA. (1983). Postretirement depression: myth or reality. Comprehensive Therapy. 9 (7): 31–37.

131. Postlethwaite, Bennett Eugene. "Fluid ability, crystallized ability, and performance across multiple domains: a meta-analysis." Ph.D. (Doctor of Philosophy) thesis, University of Iowa, 2011. http://ir.uiowa.edu/etd/1255.

132. Prull, M. W., J. D. E. Gabrieli, and S. A. Bunge (2000). 'Age-related Changes in Memory: A Cognitive Neuroscience Perspective,' In Handbook of aging and cognition II, ed. Fregus I. M. Craik and Timothy A. Salthouse, 91–153. Mahwah, NJ: Lawrence Erlbaum Associates.

133. Queiroz, B. L., & Tom, S. E. (2005). How Does Voluntary and Involuntary Retirement Affect Psychological Well-Being and Health Status Later in Life ?, 1–11.

134. Riphahn, R. T. (1999). Income and Employment Effects of Health Shocks: A Test Case for the German Welfare State. Journal Of Population Economics, 12(3), 363-389.

135. Roberts, B.A., Fuhrer, R., Marmot, M., Richard, M. 2011. Does retirement influence cognitive performance? The Whitehall II Study. Journal of Epidemiology and Community Health. 65(11), 958-963.

136. Rohwedder, S., and Willis, R.J. 2010. Mental Retirement. Journal of Economic Perspectives, 24(1), 119–138.

137. Rohwedder, Susann, and Robert J Willis. 2010. Mental Retirement. Journal of Economic Perspectives 24 (1): 119–38.

138. Rubin, D. B. (1977). Assignment to treatment group on the basis of a covariate. Journal of Educational Statistics, 2, 1–26.

139. Rubin, D., (1974). Estimating causal effects of treatments in randomized and non-randomized studies. Journal of Educational Psychology, 66, 688–701.

140. Salthouse T. (1991). Theoretical perspectives on cognitive aging, Hillsdale, NJ: L. Erlbaum Associates.

141. Salthouse, T. A. (2009). 'When does age-related cognitive decline begin?' Neurobiology of Aging, 30:507–514.

142. Salthouse, T.A., (2006). Mental Exercise and Mental Health Aging. Perspectives on Psychological Science. 1(10), 66-87

143. Scarmeas, N., and Y. Stern .(2003). 'Cognitive Reserve and Lifestyle,' Journal of Clinical and Experimental Neuropsychology 25, 625-633.

144. Sekhon, J.S., Titiunik, R. (2017). On Interpreting the Regression Discontinuity Design as a Local Experiment. In M.D. Cattaneo & J.C. Escanciano (Eds.), Regression Discontinuity Designs: Theory and Applications (pp. 1-28). United Kingdom: Emerald Publishing.

145. Sharabi, M., Harpaz, I., 2011. Gender and the relative centrality of major life domains: changes over the course of time. Community, Work & Family 14 (1), 57–62.

146. Stancanelli, E. (2012). Spouses' Retirement and Hours Outcomes: Evidence from Twofold Regression Discontinuity with Differences-in-Differences. IZA Discussion Paper Series, IZA DP No.6791.

147. Stancanelli, E., & Soest, A. (2012). Retirement and Home Production: A Regression Discontinuity approach, American Economic Review, Papers and Proceedings.

148. Stancanelli, E., & Soest, A. (2014). Joint Leisure Before and After Retirement: A Double Regression Discontinuity Approach, IZA Discussion Paper No 6698.

149. Steffick, D.E. (2000). Documentation of Affective Functioning Measures in Health and Retirement Survey. The HRS Health Working Group, Survey Research Center, University of Michigan, HRS Documentation Report DR-005.

150. Stern, Y. (2002). 'What is Cognitive Reserve? Theory and Research Application of the Reserve Concept,' Journal of the International Neuropsychological Society, 8: 448–460.

151. Stern, Y. (2003). 'The Concept of Cognitive Reserve: A Catalyst for Research,' Journal of Clinical and Experimental Neuropsychology, Vol. 25, No. 5, pp. 589–593

152. Stock, J., & Yogo, M. (2002). Testing for weak instruments in linear IV regression. NBER Technical Working Paper 284.

153. Stone, C. J. (1982). Optimal global rates of convergence for nonparametric regression. Annals of Statistics, 10, 1040–1053.

154. Sturm, R. (2002). The effects of obesity, smoking, and drinking on medical problems and costs. Health Affairs, 21, 245{253.

155. Thistlethwaite, D. L., and Campbell, D.T. (1960). Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment, Journal of Educational Psychology, 51, 309–317.

156. Trochim, W. M. K. (1984). Research design for program evaluation. Beverly Hills, CA: Sage Publications.

157. Tumino, Alberto, et al. 2016. "Retirement and cognitive abilities." Working Paper.

158. Upadhayay, N., & Guragain, S. (2014). Comparison of Cognitive Functions Between Male and Female Medical Students: A Pilot Study. Journal of Clinical and Diagnostic Research : JCDR, 8(6), BC12–BC15.

159. Van Der Klaauw, W. (2002). Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach. International Economic Review, 43(4), 1249-1287.

160. Vitae, C., & Scott, A. (2013). Agency and incentives in General Practice., 1–37.

161. Vogel, T., P.-H. Brechat, P.-M. Leprêtre, G. Kaltenbach, M. Berthel, J. Lonsdorfer (2009). health benefits of physical activity in older patients: a review. International Journal of Clinical. Practices. 63 (2).

162. WHO (2002) Active Ageing: A Policy Framework. WHO/NMH/NPH/02.8, http://apps.who.int/iris/bitstream/10665/67215/1/WHO_NMH_NPH_02.8.pdf

163. Wing, C., and Bello-Gomez, R.A. (2018) . Regression Discontinuity and Beyond: Options for Studying External Validity in an Internally Valid Design. American Journal of Evaluation. 39(1), 91 – 108.

164. Yates, J.F. (1990). Judgment and decision making. Englewood Cliffs, NJ: Prentice-Hall.

165. Zaboli, R., Seyedin, S. H., & Malmoon, Z. (2014). Original Article Macroeconomic policies and increasing social-health inequality in Iran, 3(3), 129–134. DOI:10.15171/ijhpm.2014.70

166. Ziebarth, N.R. (2010). Measurement of Health, Health Inequality, and Reporting Heterogeneity. Social Science & Medicine, 71(1), 116-124.

167. Ziebarth, N.R. and Grabka, M. M., (2009). In Vino Pecunia? The Association Between Beverage Specific Drinking Behavior and Wages, Journal of Labor Research, 30(3), 219-244.

Chapter 2

Impact of the Timing of the Initial Exposure to Maternal Depression on Children's School Readiness

Abstract

This study investigates the impact of the timing of initial exposure to maternal depression on a comprehensive measure of children's school readiness that incorporates multidimensional developmental domains that underlie school class adaptation and later success. The Early Development Instrument (EDI) scores of 59,413 children are linked to their mothers and followed over time from five years before the child's birth to the child's 5th birthday. I found that exposure to maternal depression was associated with developmental vulnerability in emotional, physical, social, and cognitive domains. When controlling for health (measured by major Adjusted Diagnosis Groups (ADGs) and minor ADGs, and hospital admission frequency) of the child at birth, and through early childhood, and mother's health prior to pregnancy, however, the effects of exposure to maternal depression on children's abilities in the emotional, physical, and social domains were attenuated across the different exposure periods. That is, although maternal depression is a risk factor for children's school readiness, children's health and socioeconomic adversity remained an important factor for early child development. Exposure to depression during pregnancy has the strongest effect on developmental vulnerability, followed by the preschool period. Emotional maturity is the most sensitive domain across the different exposure periods. In contrast, cognitive and communication domains are the least sensitive to depression. Finally, there is gender and marital status heterogeneity in the effect of maternal depression on the emotional, physical, and social domains. These findings underscore the need for early detection of maternal depression, ideally by obstetricians during pregnancy, and in programs that focus on the mother and child together. Intervention programs should commence prior to the start of school to mitigate early developmental difficulties, which exacerbate if they are not addressed.

Disclosure

The author acknowledges the Manitoba Centre for Health Policy for use of data contained in the Manitoba Population Research Data Repository under project H2018:299/HS22021 (HIPC#2018/2019-16). The results and conclusions are those of the author and no official endorsement by the Manitoba Centre for Health Policy, Manitoba Health, or other data providers is intended or should be inferred. Data used in this study are from the Manitoba Population Research Data Repository housed at the Manitoba Centre for Health Policy, University of Manitoba, and were derived from data provided by Manitoba Health, the Department of Education and Training, the Department of Families, Healthy Child Manitoba, the Vital Statistics Agency and Statistics Canada. No endorsement by the University of Manitoba is intended nor should be inferred.

Chapter 2

Impact of the Timing of the Initial Exposure to Maternal Depression on Children's School Readiness

2.1. Introduction:

School readiness can be defined as a child's profile of competencies that includes cognitive, emotional, social, behavioral, and communication skills, as well as knowledge, at the time of school entry. These are important factors that influence a child's learning, adjustment and later success (Snow, 2006; Forget-Dubois et al., 2007). Theoretical models of human capital formation show that in addition to cognitive skills formation during childhood, a variety of non-cognitive skills, i.e., personality, social and emotional traits, and communication skills are also crucial factors in shaping children's potential social and economic success (Heckman & Krueger, 2003; Heckman et al., 2006; Cunha et al., 2010; Mukherjee, 2011). Furthermore, growing evidence from neuroscience recognizes that interaction between genes (nature) and early childhood experiences from the surrounding environment (nurture) - which includes children's connection to their parents, relatives, teachers, peers - shapes the architecture of the developing brain in a manner that serve as a foundation for potential social, emotional, and cognitive skills formation (NSCDC, 2012; Pieterse, 2015). Any disruption to this developmental process may impair a child's capacity to acquire and accumulate skills required to learn at school and to integrate into society.

Maternal depression has been recognized as one of the most important risk factors that disrupt a child's developmental processes, which span physiological, behavioral, emotional, social, communicational, and cognitive dimensions. These dimensions, in turn, are strong predictors of a child's later school performance (Duncan et al., 2007; Forget-Dubois et al., 2007;

Grimm et al., 2010; Romano et al., 2010). While a small, but growing, body of research has investigated the relationship between maternal depression and child development, little attention has been paid to the association between the timing of maternal depression exposure and child's skills formation. The timing of initial exposure to maternal depression may have a heterogeneous impact on children's developmental outcomes. For instance, maternal depression during the first year of life after birth may adversely affect secure infant attachment or the healthy emotional bond between children and their mothers. The lack of a secure mother-infant attachment may impair or delay the emotional and cognitive development in early childhood (NIHCM, 2010), which may lead to long-lasting human capital accumulation gaps and socioeconomic inequality. Moreover, many studies found that depressed women during the early postpartum period are less likely to initiate breastfeeding and may be at high risk of infant-breastfeeding problems, including reduced breastfeeding duration and a lower level of breastfeeding self-efficacy (Dennis & McQueen, 2009). According to the American Academy of Pediatrics (1997), consumption of human milk via breastfeeding during the first six months is associated with better health and neurological development.

A variety of conceptual and empirical research from a broad range of disciplines, including economics, neurobiology, neuroscience, genetics, developmental psychology, and sociology, has shown adverse effects of children's exposure to maternal depression on developmental competencies in middle childhood and adolescence. However, there is much less research on the association between the timing of maternal depression and developmental functioning during the early childhood period, a period that is characterized by rapid growth in a child's emotional, cognitive, social, and behavioral capacities. Consequently, children who lack a strong foundation in these skills and abilities at school entry may have lower academic

trajectories than those with a high level of school readiness. Therefore, if maternal depression negatively affects the formation of these skills during the optimal acquisition period, then children may lack the skills necessary for classroom adaptation and academic development (Brinkman et al., 2013; Quach and Barnett, 2015). However, few studies from this diverse literature have examined the outcomes of children on a comprehensive measure of school readiness that incorporates the multidimensional nature of human developmental domains that underlie academic success.

The various dimensions of depression exposure, chronic illness, and maltreatment are significantly associated with later academic achievements in school (Romano et al., 2015), so these factors may be also important risk factors that affect child development to start school. In fact, later academic achievements may be the long-lasting consequence of adverse events and developmental delays early in life. Some studies have shown that the period from conception to the child 5^{th} birthday is crucial in early child development because the interactions between environment and genetics shape the architecture of the brain, making it a highly sensitive period to stress (Shonkoff and Phillips 2000; Knudsen et al 2006; Panzer, 2008).

This study aims to fill the gaps in the literature by investigating the impact of the timing of initial exposure to maternal depression on the multidimensionality of child capabilities and skills formation at kindergarten age, which reflects the child's readiness to learn at school. I adopt the human capital accumulation model as a framework (Heckman and Krueger, 2003; Heckman et al., 2006; Heckman, 2007) to understand the association between maternal depression exposure and developmental vulnerabilities in early childhood, with emphasis on distinguishing the effects of the timing of first exposure to maternal depression on several

developmental indices at age 5 years, which include emotional, physical, social, cognitive, and communication abilities.

This paper estimates the causal effect of maternal depression on a wide range of child outcomes and makes four key contributions. First, I examine the impact of maternal depression exposure on a comprehensive measure of school readiness, the Early Developmental Instrument (EDI), which incorporates multidimensional developmental domains that are the foundation for academic adaptation and success. Second, I investigate the impact of the timing of children's initial exposure to maternal depression on children's school readiness as measured by the different domains in the EDI. There is a general consensus in the economics and health literature of the importance of experiences during the first five years of a child's life on the long-lasting development of human capabilities. There are even some studies that include the time before birth in risk factors that affect school readiness (Aktar et al., 2019; Deave et al., 2008; Plant et al., 2015; O'Connor et al., 2002). Therefore, it is important to understand when maternal depression is more likely to have the greatest impact on children's abilities to acquire and accumulate skills, which, in turn, affect readiness to learn at school. Third, although child health through early childhood has a very important impact on a child's academic performance later in life, it is less clear whether this impact is evident at kindergarten age. It is therefore important to assess the impact of child health on school readiness, which will adjust for any confounding in the effect of maternal depression exposure. The Johns Hopkins "Adjusted Clinical Group®" (ACG®) system uses thousands of International Classification of Disease (ICD) diagnosis codes to assign each diagnosis code to one of 32 "Aggregate Diagnosis Groups" (ADGs), resulting in 8 Major ADG's and 24 Minor ADG's for children. The diagnosis-to-ADG mapping system in the ACG® software may assign one single ICD diagnosis code to more than one ADG group, which

indicates more than one underlying morbidity type (John Hopkins ACG, 2014). Fourth, the mother's health prior to birth, including before pregnancy, can affect child health biologically (genetic endowment) or child development. Mother health profiles five years before birth are included as controls by using the Johns Hopkins ACG® system to assign each diagnosis code from hospital abstracts and physician visit claims to the major illness and minor illness group. Empirically, some studies found that mother pre-pregnancy obesity has an increased risk for negative emotionality and inattentiveness in early childhood (Robinson et al., 2013). There is also evidence that obesity and metabolic syndrome are related to the elevated risk of depression, which suggests that there is a potential link between maternal metabolic health and fetal brain development in utero[1] (Robinson et al., 2013; Koponen et al., 2008). Therefore, it is reasonable to postulate that children's early development could be affected by the mother's health before birth. To date, however, studies of maternal depression have not considered the consequences of maternal health pre-pregnancy on child developmental outcomes.

Specifically, the study addresses the following three research questions:

- *Does the timing of a child's initial exposure to maternal depression have an impact on the child's school readiness at kindergarten age?*

  That is, is there an association between the timing of children's initial exposure to maternal depression and their cognitive, physical health, emotional, social and communicational capabilities at kindergarten age, as a foundation for school class adaptation and success?
- *What is the period of maternal depression exposure that has the strongest effect on the child's school readiness?*

  That is, is there one period of maternal depression exposure that has stronger effects on children's development than others and hence should be a priority for policy interventions?
- *Is there socioeconomic or demographic heterogeneity in the impact of timing of maternal depression exposure on the child's school readiness?*

[1] Metabolic syndrome is a cluster of risk factors for type 2 diabetes and cardiovascular diseases.

That is, is the development trajectory a result of the dynamic interplay between nature (genes) and nurture (environment), which includes whether the mother is a teenager at birth, whether the family depends on income assistance (poverty), neighborhood-area socioeconomic status, education level of the mother, and having a single mother? For example, children's development outcomes of depressed mothers may be worse among children living in families that depend on income assistance or among children with less-educated mothers. Although exposure to early maternal depression may be a significant predictor of later developmental trajectories, response to early childhood adversity may vary across children.

I overcome limitations of previous research by using a population-based administrative database from the Manitoba Centre for Health Policy (MCHP) Repository, which allows me to follow depressed mothers and their children from before pregnancy through kindergarten. The database includes the Early Development Instrument (EDI), which is a population-based assessment of children's development in five basic domains: health, social competence, emotional, language and cognitive development, and communication skills and knowledge. Hence, I can investigate the impact of maternal depression on more than one dimension of children's development, rather than just one dimension as in previous studies. Another limitation in the literature that I address is the definition of maternal depression, which was measured as an indicator of depressive symptoms rather than an official diagnosis (Claessens et al., 2015). Rather, with the longitudinal population-based database at MCHP, I use the international classification of disease (ICD) coding system to identify maternal depression and the timing of depression exposure. An advantage of the Repository database is that I can link the child-mother

master file with different datasets that include rich information about child and mother health, social, and demographic characteristics.

Why is this study important? Previous research found that some indicators of the early development instrument (EDI) at age five are strong predictors of later school performance. These predictors (sub-domains of EDI) include math, reading, attention (Duncan et al., 2007; Grimm et al., 2010; Hooper et al., 2010); language and cognitive development domain (Forget–Dubois et al., 2007); social and emotional behaviours (Grimm et al., 2010; Pagani et al., 2010; Romano et al., 2010); and general knowledge (Grissmer et al., 2010). In addition, empirical models of human capital formation predict that an early gap in a child's skills formation at age five will exacerbate long-lasting socioeconomic status inequalities. Therefore, identifying the impact of maternal depression on different dimensions of child skills formation and capabilities at age five will help to pinpoint early life channels that impair a child's readiness to learn later at school and help reduce the pre-school human capital accumulation gap. This should be one of the top cross-sectoral priorities for policymakers. Moreover, prevention programs that target early adversity in children's development should give high priority to the timing of the intervention, because research on sensitive periods suggest that children who do not get the appropriate skills and nurturing at the right timing may experience difficulties in acquiring these skills in a later time (Cunha et al., 2005; Heckman et al., 2006; Cunha and Heckman 2007; Heckman 2007; Heckman et al., 2010; Heckman et al., 2013).

The remainder of this paper is organized as follows. Section 2 summarizes the related literature; section 3 describes the data and presents summary statistics for the sample; section 4 provides the empirical econometric specification and methodology; sections 5 and 6 show the

results of the estimation and heterogeneity of maternal depression on school readiness; section 7 discusses and concludes.

2.2. Literature Review

The new wave of literature in human capital development during the last decade reached general consensus on five principles of early childhood development: (i) child development is a foundation for community development and economic development (National Scientific Council on the Developing Child at Harvard University, 2007); (ii) early parents-child interactions and genetic endowment literally shape the architecture of the developing brain, which is composed of highly integrated sets of neural circuits that can wire fast based on continuous interaction between environment and genetics influences during early childhood, and hence human capital accumulation (Greenough, 1991; Greenough and Black, 1992; National Scientific Council on the Developing Child at Harvard University, 2010); (iii) the most sensitive period of child development spans from the time of conception through the fifth birthday (Houston, 2014; Georgiadis et al., 2016); iv) interventions in development that takes place during early childhood is likely to be more effective and less costly than targeting at a later age (Cicchetti et al., 2000; Field, 2002); and v) there is a negative association between toxic stress[2] in early childhood and both the nervous system and stress hormone systems that can damage developing brain architecture, which creates the foundation for later problems in learning, behavior, and both physical and mental health ((Diego et al., 2009; Mattes et al., 2009; Center on the Developing Child at Harvard University, 2010; National Scientific Council on the Developing Child at Harvard University, 2014; Shonkoff and Phillips 2000; Knudsen et al 2006).

[2] For more information about the toxic stress effect on child development, see the Center on the Development Child online guide - Harvard University. https://developingchild.harvard.edu/science/key-concepts/toxic-stress/

Maternal depression is a worrisome issue that around 10 to 20 percent of mothers experience, either during pregnancy or at some point during their lives, and these figures are worse for mothers with a history of previous depression episodes (NIHCM, 2010). Approximately 50% of women who received a referral for depression treatment accessed follow-up treatment (Godman & Tyer-Viola, 2010).

According to the Public Health Agency of Canada (2014), 15.5% of women are either diagnosed with depression or treated with anti-depressant drugs before they become pregnant, and 12.5 % of women reported some kind of stress symptoms during the 12 months before birth. Population-based studies from the Manitoba Center for Health Policy showed that the prevalence of children whose mothers had mood or anxiety disorder was stable across the study period at 20.2% in 2000/01 – 2001/02 and 21% in 2008/09- 2009/10. Moreover, the prevalence was 20.3% in the first time period and 21.7% in the last time period for mothers of children 0-5 years old (Brownell et al., 2012). These figures make maternal depression an important public health issue, which should be investigated and assessed for potential impact on children's development.

Children's developmental trajectories during their early life have been widely documented in psychology, neuroscience, medicine, and economics as an outcome of a series of dynamic interactions between genes (inherited at conception) and the environment with which children interact. For instance, the effect of genes on fetal development during pregnancy and birth outcome trajectories is contingent on the environment in the fetus and uterus. Similarly, the effect of fetal outcomes, such as growth rate, low birth weight, and preterm birth, on infants' development and health trajectories are contingent on the surrounding environment during pregnancy, and adulthood outcomes are contingent on the lifestyle and different environmental factors during childhood (Wadhwa, 2005).

A mother's behavior profile, such as depression, smoking, drugs, and alcohol consumption, can impact children's developmental trajectories at different points in time because the mother's profile of behaviors during pregnancy and her interactions with her children after birth has a significant impact on children's brain development. For instance, maternal depression, part of the mother's biochemical profile, may begin to affect children's brain development in the fetus. Lundy et al. (1999) and Field et al. (2004) noted that prenatally depressed mothers have a higher level of both cortisol (stress hormones) and norepinephrine (neurotransmitters), and a low level of dopamine during pregnancy (Field et al., 2004). The mother's elevated level of cortisol may cross the placenta and directly affect the fetus's neurotransmitters/neurohormone levels (Gitau et al., 1998; Glover et al., 1999), and the elevated levels of catecholamine (norepinephrine and epinephrine) may reduce the uterine blood flow to the fetus (Glover et al., 1999) and directly affect the neurobehavioral development. Hence, exposure to a biochemically dysregulated environment in utero affects early brain development.

Exposure to maternal depression during the prenatal period (time between conception and birth) was found to be associated with increased infant cortisol levels (Gutteling et al., 2005; Brennan et al., 2008), low birth weight (Field, 2011), preterm birth (Field et al., 2008), decreased breastfeeding initiation, disorganized sleep and less responsiveness to stimulation in the neonate (Field, 2011), and shorter gestational age (Field et al., 2008). Emory and Dieter (2006) found that fetuses of prenatally depressed mothers are more active during mid-gestation and experience a lower baseline heart rate and less total movement during late-term vibratory stimulation. To identify whether this excessive fetal activity may delay fetal growth, Diego et al (2009) used fetal weight and birth weight data to measure fetal growth rate and found that prenatally depressed mothers have elevated cortisol level and that their fetuses were smaller with lower

fetal growth rates and lower birth weight, which are risk factors for impaired cognitive and social development (Wadhwa, 2005). Romeo et al (2010) investigated the cognitive function of late preterm (LPT) babies during the infancy period and found that children born at 33-36 weeks experienced a delay in cognitive function relative to full-term infants at 12 and 18 months. Chyi et al (2008) compared school outcomes for infants born at 32-33 weeks (moderate preterm (MPT)), 34-36 weeks (late preterm (LPT)) and full-term (FT) infants using an early childhood longitudinal study-kindergarten cohort. The study found that LPT infants have lower reading scores than FT infants in kindergarten to grade one; MPT infants perform generally poorer than FT infants over all other periods.

Latendreise et al (2015) found that women whose depression lasted for a long period during pregnancy are more likely to experience an increase in admission of their infants to the neonatal intensive care unit (NICU). To investigate whether NICU-admission adversely affects a child's cognitive development, Baron et al (2011) assessed the cognitive outcomes of a cohort of complicated late preterm (CLPT) children at 35-36 weeks against full-term infants using the Differential Ability Scales-Second Edition (DAS-II), a well-standardized multi-subtest measure of General Conceptual Ability (GCA). The study found that children who got admitted to the NICU did poorly on non-verbal reasoning (e.g., picture similarities) and spatial (coping, pattern construction) cluster score, and poor performance on both GCA and verbal reasoning (verbal comprehension, naming vocabulary) than full-term children at preschool age. In another study, Barons et al (2009) found that NICU-admitted infants are more likely to have negative neuropsychological sequelae. Field et al (2004) assessed both maternal and neonatal biochemistry and found that mothers with depressive symptoms have a higher level of cortisol

and lower level of dopamine and their newborns mimic their hormone profile. In addition, they found that cortisol levels were associated with prematurity.

Some studies investigated the effect of prenatal depression on cognitive development as assessed by scores on the Bayley Mental Development Index (MDI), school grade, or by a language development measure (Talge, 2007). They found an association between antenatal maternal depression and child's temperament in early childhood and low school marks at age six (Niederhofer & Reiter, 2004), as well as low score on the MDI at 3 and 8 months of age (Huizink et al., 2003). Laplante et al (2004) investigated the impact of the prenatal level of stress on a child's intellectual and language development at age 2 years using the Ice storm in Quebec during 1998 as a natural disaster to identify the severity of women's stress. They found that the birth weight and age of the children account for 12% and 14.8% of the variance in the Bayley MDI and in the language development, respectively. In addition, prenatal maternal stress accounts for 12.1%, 11.4%, and 17.3% of the variance in Bayley MDI score, productive language, and receptive language abilities, respectively, after controlling for children's other health factors at birth.

In addition to effects on social and cognitive development outcomes, prenatal maternal depression can cause disability and health-related problems during the toddler age. For example, one study found that women who experienced antenatal stress at 18-week gestation are more likely to have a mixed-handedness child when they were 42-months of age (Glover et al., 2004). Raposa et al (2014) used longitudinal data to investigate the impact of depression during pregnancy on child health during their first five years. The study found that there is a direct association between prenatal depression and child health before the age of five. They also measured maternal depression at three time periods: 3-4 days, 6 months, and 5 years after birth

and found that maternal depression during these periods was associated with child's physical health at age five, which in turn, predicted increased health-related stress and poor social function at age 20.

The first year after birth is recognized as a very sensitive period of child development. From a developmental cognitive neuroscience perspective, emotion and cognition are dynamically linked and work together during the first year of the child's life to shape the base for subsequent development (Bell & Wolf, 2004). Therefore, how children relate to and interact with their mothers are very important for developing the child's emotional and cognitive skills. John Bowlby (1982) termed the affectional bonds that develop between children and their primary caregivers, usually their mothers, during their early time of life "attachment". Mother-infant attachment refers to the deep and lasting emotional bonds between the child and her/his mother. When children are frightened, fatigued, or sick, they seek their attachment figures (mothers) for help, protection, and soothing. When children find caregivers available and responsive to their needs in time of stress, their attachment system, which refers to psychological organization hypothesized to be within children, will activate feelings of security and safety toward their caregivers (Bretherton, 1985). In addition, accumulated feelings of security during attachment relationships are the main factor that regulates children's motivation to maintain proximity to caregivers and explore the surrounding environment (McCormick et al., 2016). Securely attached children trust that the primary caregiver is physically and psychologically available and responsive to their needs; their attachment systems will be deactivated and, thus, they can use their caregivers as a secure base from which to play and explore the world (Bretherton, 1985). In contrast, insecurely attached children do not trust in an appropriate physical and psychological availability of their caregivers in times of distress or threat. Hence, insecure children's

attachment systems are more activated, and their exploratory systems are more deactivated (McCormick et al., 2016). Maternal depression can affect a child's skills formation during the early years after birth through mother-infant attachment because mother-infant attachment facilitates the association between maternal depression and child's multidimensional skills formation and hence school readiness.

Various studies suggested that affective quality of mother-child attachment influence child's cognitive skills formation at early childhood where children depend on their mothers for social and intellectual stimulation and any long-lasting disturbance to the capacity of making affectional bonds, or repeated disturbance for existing bonds, are potential causes of psychiatric disturbance in childhood (Bowlby, 1970; Cowan, 1982). Relative to unsecured mother-child attachment, securely attached children have more willingness to approach and persist in tasks, accept their caregivers' assistance and support in solving problems, and they have social competence which enhances the greater flow of information between themselves and caregivers (Estrada, 1987). Bretherton (1985) suggests that securely attached children approach cognitive tasks in ways that enhance cognitive development and they show more curiosity, enthusiasm, persistence in their problem-solving styles than insecurely attached infants.

There is a relatively large empirical literature that suggests an association between mother-child attachment and cognitive skills formation, which includes intelligence, memory, and reasoning (Williams et al., 1988; Spieker et al., 2003). McCormick et al (2016) examined the association between mother-child attachment styles and cognitive development, as measured by math and reading skills, at age 54 months to fifth grade. They found that insecurely attached infants experience a low average level of reading and math skills. They also found that children's task engagement in the classroom partially mediates the association between insecure attachment

and reading and math skills in middle childhood. West et al (2013) investigated the mechanisms that mediate the association between mother-infant attachment and children's cognitive performance, including both academic performance and IQ. They found that securely attached children at ages 24 or 36 months show better cognitive performance and this association is significantly mediated by maternal teaching quality and academic support to children, children's social relationships with peers, support at school, and children's regulatory characteristics. In contrast, they also found that both insecure-ambivalent attached and disorganized attached children at 36 months had lower grades and IQ scores in middle childhood.

Murray (1992) used a sample of 113 mother-infant pairs and screened mothers for maternal depression after childbirth. The sample was followed up to 18 months when their infants were assessed on a measure of cognitive, language, social and behavior development. The study found that infants of postnatally depressed mothers were insecurely attached to their mothers, perform worse on object concept tasks, had behavioral problems, but there is no impact on language and general health development.

In addition to the psychological mechanism, maternal depression can impact children's skills formation through children's brain physiological mechanisms. Some studies in neuroscience have been working on brain biological vulnerabilities as a mechanism through which maternal depression can affect children's cognitive and language development. Using an electroencephalogram (EEG) as a measure of infant and toddler brain activity, they noted that positive emotion was accompanied by greater relative left frontal EEG electrical activity (Dawson et al., 2001; Field et al., 2004; Diego et al., 2006) and negative emotion (like sadness and stress) and the withdrawal was accompanied by greater relative right frontal EEG electrical activity (Jones et al., 2000). For instance, Davidson and Fox (1989) noted that 10-month-old

infants who cried during the period of mother separation showed higher relative right frontal EEG activation relative to infants who did not cry. Dawson et al (1997) supported this finding in a study of infants 13-15 months old and found that infants of depressed mothers exhibited lower levels of left frontal EEG electrical activity than infants of non-depressed mothers. In addition, Jones et al (2000) found that infants of depressed mothers showed more relative right frontal activity.

These findings show that infants of depressed mothers exhibit greater right frontal EEG asymmetry (greater right than left EEG power), which, in turn, increases the risk of withdrawal and negative affect (Sohr-Preston and Scaramella, 2006). Field and Diego (2008) noted that the infant's EEG profiles are stable from the neonatal stage to early infancy to preschool age and there is an association between the EEG profiles of depressed mothers and the EEG of their infants. These findings have also been interpreted as a biological marker for symptoms associated with depression in infants (Jones et al, 2001). Consequently, since withdrawal and negative affect disturb infants' capabilities and readiness to process and interact with stimuli from the surrounding environment, infants of depressed mothers are more likely to experience a delay in cognitive and language development (Dawson et al., 1992). Moreover, Field (1995) found that infants of depressed mothers have lower vagal tone and high right frontal EEG asymmetry, and both are associated with reduced self-regulation and emotional expression.

The literature on maternal depression reported three aspects of maternal depression that should be controlled to clearly identify whether there are sensitive periods during which maternal depression has the strongest effect on the child's capabilities at school age. These aspects include the timing, severity, and chronicity of depression. The timing of initial exposure to maternal depression is an important issue because there may be crucial periods during which maternal

depression significantly impacts a child's skills formation, which, in turn, impacts cognitive and non-cognitive development. For instance, Essex et al (2001) examined the impact of initial exposure to maternal depression on the child's internalizing and externalizing behaviors [3]in two time periods that include the first-year postpartum and the toddler years, age 2 to 4.5. Using a prospective community-based study, they found that children's exposure to maternal depression during the first year after birth was associated with high internalizing symptoms, and exposure to depression during the toddler period was associated with an increased risk of externalizing symptoms among girls. Similarly, Luoma et al (2001) used a sample of 349 mother-child pairs to examine the association between maternal depressive symptoms at a different point in time (prenatal, postnatal, and children age 8 to 9) and the level of children's psychosocial function and emotional and behavioral problems in school-age children. The study found that prenatal depression was a strong predictor of children's high externalizing problems, and postnatal depressive symptoms were a good predictor of children's low social competencies. In addition, they found that concurrent and prenatal depressive symptoms were associated with the lowest child outcomes.

Some studies found an association between the timing of maternal depression and emotional disorder. For instance, Naicker et al (2012) found that adolescents whose exposure to maternal depression was between the ages of 2-3 and 4-5 years are twice as likely to have emotional disorders than adolescents of non-depressed mothers. Moreover, they found that the sensitive period of initial exposure to maternal depression occurs between ages 2 and 5 and not

[3] Externalizing behaviors are easily observable by others because they are actions in the external world, such as acting out, antisocial behavior, hostility, aggression, disruption, and destruction of property. Internalizing behaviors are quiet and often invisible because they are internalized and are generally not disruptive to others, such as feeling unloved or unwanted, being withdrawn, anxiety, felling sad and lonely, somatization, and depression (Winters et al., 2008).

during the first year after birth as documented in previous studies. In contrast, Brennan et al (2000) examined the impact of timing of maternal depressive symptoms on a child's vocabulary test, as a measure of cognitive functioning, and behavior in a cohort of 4,953 Australian children. They found that the timing of maternal depression was not significantly related to the child's vocabulary score, but the chronicity and severity of maternal depression were related to more behavior problems and lower vocabulary scores. Although some studies investigated the impact of maternal depression at different points in time, there is inconsistency in their findings, which is partially due to investigators using different measures to assess the outcome of maternal depression, and partially because they used different periods of time for initial exposure to maternal depression.

In brief, the impact of maternal depression on child's development is well-documented in the literature but generally focuses on just one dimension or domain of the child's development. However, the evidence regarding the impact of maternal depression on different dimensions of child's skills, as measured, for example, by a comprehensive measure of child school readiness, are unclear and a number of questions regarding the association between maternal depression at different points in time and child's school readiness at kindergarten age remain unanswered. For example, an important question that remains to be answered is when maternal depression has the greatest impact on the child's profile of competencies (Timing of initial exposure). The current study addresses these gaps in the literature by using a population-based database that provides an excellent opportunity to follow up mothers and their children from the prenatal period to kindergarten age. This allows me to investigate the impact of timing of the child's initial exposure to maternal depression on different aspects of a child's skills and development. Specifically, the study examines which period of children's initial exposure to maternal

depression has the greatest impact on the profile of competencies that shape school readiness at kindergarten age. Four specific periods that are recognized in the neuroscience literature and have developmental characteristics that make them independent from each other were utilized. Additionally, the study investigates whether a high-risk family environment in which the child develops exacerbates the impact of the timing of initial exposure to maternal depression. That is, I investigate to what extent the impact of maternal depression timing differs among different socioeconomic and demographic groups.

2.3. Sample and Study Period:

2.3.1. Sample size and study period

This study uses six cohorts of children who are born in Manitoba, Canada, and who received the Early Development Instrument (EDI) assessment by their kindergarten teachers between February and March in all 37 public school divisions in Manitoba. Depending on the collection years (2005/06; 2006/07; 2008/09; 2010/11; 2012/13; 2014/2015), I linked children born between (2000 - 2009) to their biological mothers through a unique identifier, Personal Health Identification Number (PHIN), which gives a sample of approximately 70,683 mother-child pairs.

Mother-child pairs are followed over time from five years before the child's birth to the time at which the Early Development Instrument (EDI) assessment is completed for the child. Since the first cohort starts in 2000 and mother-child pairs will be followed starting five years prior to the child's birth, the study period will cover the time from 1995 – 2015. Children cohorts used in the study, based on the available EDI assessments in 2005/06; 2006/07; 2008/09; 2010/11; 2012/13; 2014/2015), can be identified as follows:

- The first cohort (2000-2005/2006): children born in 2000 and can be followed over time until their EDI assessment is completed in the school year 2005/06.
- The second cohort (2001-2006/2007): children born in 2001 and can be followed over time until their EDI assessment is completed in the school year 2006/07.
- Third cohort (2003-2008/09): children born in 2003 and can be followed over time until their EDI assessment is completed in the school year 2008/2009.
- Fourth cohort (2005- 2010/11): children born in 2005 and can be followed over time until their EDI assessment is completed in the school year 2010/2011.
- Fifth cohort (2007 - 2012/13: children born in 2007 and can be followed over time until their EDI assessment is completed in the school year 2012/2013.
- Sixth cohort (2009 - 2014/15): children born in 2009 and can be followed over time until their EDI assessment is completed in the school year 2014/2015.

In order to carry out the analysis in this study, I made the following exclusions: (i) children with invalid, duplicate, or missing PHIN, because this was required to link children to their mothers; (ii) children with invalid EDI; (iii) children without continuous health coverage from pregnancy to kindergarten age, since I needed to follow their health status. This includes children born out of province, not born in a hospital, or moved out of province before the EDI assessment date; (iv) children of multiple births (i.e., twins or triplets) were excluded from the analysis where they maybe not independent observations (Fransoo, 2007); (v) children linked to postal codes of a public trustee; and (vi) children's mothers who did not have a valid PHIN or complete health status profile. After these exclusions, the final working sample that I used in my analysis includes 59,413 children.

2.3.2. Datasets

This study uses data from the population Health Research Data Repository, which includes a collection of administrative databases housed at the Manitoba Centre for Health Policy (MCHP), University of Manitoba (Brownell, 2016). The repository data are population-based dataset which includes nearly all residents in Manitoba. This repository database is well suited for this study because it contains databases across multiple domains, including health, education, social services, and health services utilization, that can be linked for analysis using an encrypted Personal Health Identification Number (PHIN). Also, the administrative data on health services utilization contain detailed information on every health contact with health service providers over time which allows following up mothers and their children using the unique identifier number (PHIN). The datasets that I used in this study include:

(i) Hospital discharge abstracts. This data is a hospital form/computerized record filled out upon the patient's discharge from the hospital and includes detailed information about the hospitalized patients in Manitoba. The abstract contains information about patient sex, date of admission and discharge, length of stay, diagnosis codes, and type of service (inpatient, outpatient, or surgery).

(ii) Manitoba Health Insurance Registry (Population Health Registry). This is a population-based registry of all persons registered with Manitoba Health since 1970. The Registry allows for database linkages using scrambled PHINs. It also provides individual and family-level information such as birth date, sex, postal code, family size and mother's marital status.

(iii) Early Development Instrument (EDI). This represents a province-wide population-based assessment of children's early development at kindergarten age. For the current study, this database provides school readiness results in 5 domains: physical health and well-being, social competence, emotional maturity, language and cognitive development, and communication skills

and general knowledge for all Kindergarten children assessed in 2005/06; 2006/07; 2008/09; 2010/11; 2012/13; and 2014/2015 and collected by Healthy Child Manitoba Office (HCMO).

(iv) Drug Program Information Network (DPIN) Dataset. The DPIN database is maintained by Manitoba Health and contains transaction-based prescription drug claims from all pharmacies in Manitoba. The DPIN system contains information about the dispense date of medication, medication history, sex, and the unique patient identification number for all Manitoba residents. This information is linked to hospital discharge abstracts to identify mothers diagnosed with maternal depression and the time of the first diagnosis.

(v) Enrollment, Marks and Assessments Database. This dataset provides information on school enrollment, courses, marks, standards tests, graduation status, and educational assessments for all Manitoba students from Kindergarten to Grade 12. The data are maintained by Manitoba Education and is used to identify children's mothers who have a high school degree.

(vi) Baby First Screen and Families First Screen. This data is from a social survey that collects information from the newborn's family within a week of the newborn's discharge from the hospital and is maintained by the Healthy Child Manitoba Office (HCMO). The survey includes questions about parent's ethnic background, alcohol use, drug use, child abuse, and education. This dataset is used to identify the educational degree of children's mothers who could not be identified from the enrolment, marks, and assessments dataset, and to identify mothers who smoked during pregnancy. The family first screen replaced the baby first screen in 2003.

(vii) Employment/Income Assistant program (SAMIN). This dataset is maintained by the Department of Families and provides information on individuals and families who receive income assistance. This data set was used to identify families that receive income assistance

from the Employment and Income Assistance (EIA) Program, an individual-level measure of socio-economic status (SES).

2.3.3 Outcome (Dependent) Variable

Children's school readiness was assessed using the Early Development Instrument (EDI). The EDI is a population-based, community-level measure of children's development in five major domains that include: physical health and well-being, social competence, emotional maturity, cognitive and language development, and communication skills and general knowledge (Janus and Offord, 2007). In Manitoba, the EDI is a 103-item close-ended questionnaire collected province-wide every two years on behalf of the Health Child Manitoba Office (HCMO) by all kindergarten teachers (Santos et al., 2012). Consequently, teachers complete the assessment of a child's skills based on their knowledge of the children in their classrooms. Internal reliability, test-retest reliability, and external validity of EDI as a strong predictor of child's readiness to learn at school have been examined in several studies (Forget-Dubois et al., 2007; Janus and Offord, 2007; Janus and Duku, 2011; Guhn et al., 2011; Forer and Zumbo, 2011). Moreover, extant literature reports adequate evidence on the important impact of each domain in EDI on children's adjustment and success at school, as well as the predictive power of EDI for short- or long-term school performance (Brinkman et al. 2013; Guhn et al. 2016; Janus and Reid-Westoby, 2016; Davies et al. 2016;). The following are the five domains of school readiness to learn included in the Early Development Instrument:

1- physical health and well-being domain: this consists of 13 items that examine three sub-domains i) physical readiness for the school day, ii) physical independence, iii) gross and fine motor skills, and 4) energy level. Children who are "vulnerable" or "not ready" have average or

poor fine and gross motor skills, are sometimes tired or hungry, often have flagging energy levels and less than average overall physical development.

2-Social competence and knowledge domain. The social competence domain consists of 26 items that examining four sub-domains: i) Overall social competence and cooperation, ii) responsibility and respect, iii) approaches to learning, and iv) readiness to explore new things (Janus & Offord, 2007). Children considered "vulnerable" on this domain have challenges that include poor overall social skills, with regular problems in getting along with other children, problems following rules and class routines, showing respect for adults, children, and other's property. They can lack self-confidence, self-control, and may be unable to work independently.

3-Emotional maturity domain. The emotional maturity domain consists of 30 items that examine four sub-domains: i) pro-social and helping behavior, ii) anxious and fearful behavior, iii) aggressive behavior, and iv) hyperactivity and inattention (Janus & Offord, 2007). Children considered "vulnerable" or "not ready" on this domain regularly have problems managing aggressive behavior, are prone to disobedience, and/or easily distractible, inattentive, impulsive, usually unable to show helping behavior towards other children, and are sometimes upset when left by a caregiver.

4-The language and cognitive development domain. This consists of 26 items that examine four sub-domains: i) basic literacy, ii) interest in literacy/numeracy and uses memory, iii) advanced literacy, and iv) basic numeracy (Janus & Offord, 2007). Children considered "vulnerable" or "not ready" on this domain have challenges in reading/writing and numeracy; have difficulty remembering things; are unable to read and write simple words; are unable to attach sounds to letters, cannot count to 20, are unable to recognize or compare numbers, and are not interested in numbers.

5-Communication skills and general knowledge. Communication skills and general knowledge, unlike the other domains, does not consist of subdomains. This domain examines children's ability to communicate their needs and ideas effectively and the children's interest in the surrounding world (Janus & Offord, 2007). Children considered "vulnerable" or "not ready" on this domain have poor communication skills and articulation, a limited command of English/French, have difficulties in talking to others, understanding and being understood, and have poor knowledge understanding about their world.

The five domains of the EDI are evaluated separately to recognize the weak and strong dimensions of child development. Therefore, a child can show vulnerability in one or more domains and not in other ones. The core questions in each of the five domains of EDI are scored from "0" (lowest score or vulnerability) to "10" (highest score) and the domain total score is calculated as an average score of all questions in the sub-domains, scored out of "10" (Janus and Offord, 2007). Logically, since the score of each of the five domains is out of "10", the total score of EDI is out of "50".

Following Janus and Offord (2007), children at kindergarten age can be classified as "not ready", "ready", and "very ready", based on the score distribution inside each domain. That is, children who score in the bottom 10th percentile cut-off score are classified as 'developmentally vulnerable' or "not ready" for school, based on the skills in that domain. In contrast, children who score in the top 30th percentile cut-off score are classified as "very ready" for school. Children who score between the 11th and 69th percentile are classified as "ready" for school. For this study, these three categories were collapsed into two 'vulnerable/not ready' and 'ready' to capture established developmental vulnerability, based on documented or theoretical relevance to the EDI domains.

2.3.4 Covariate (Independent) Variables:

2.3.4.1 Maternal Depression

2.3.4.1.1 Algorithm of Identifying Depressed Mothers

Using administrative data collected over time represents a unique opportunity to recognize depressed mothers who experienced depression for the first time in each of the sensitive periods of early child development in this study. I used algorithms from previous MCHP studies to recognize depressed mothers based on official diagnoses (Martens et al., 2004; Fransoo et al., 2009; Martens et al., 2010; Chartier et al., 2012). These algorithms use the international classification of disease (ICD) coding system to identify depressed mothers as follows (i) at least one hospitalization with any of ICD-9-CM diagnosis codes 296.1-296.8, 300.0, 300.2-300.4, 300.7, 309, 311or ICD-10-CA codes F31, F32, F33, F34.1, F38.0, F38.1, F40, F41.0-F41.3, F41.8, F41.9, F42, F43.1, F43.2, F43.8, F45.2, F53.0, F93.0; OR (ii) at least one hospitalization (any diagnosis (dx) code) with ICD-9-CM code 300 or ICD-10-CA codes F32, F34.1, F40, F41, F42, F44, F45.0, F45.1, F48, F68.0, F99 AND one or more prescription for antidepressant or mood stabilizer. (iii) at least one physician visit (prefix=7) with ICD-9-CM codes 296, 311 (iv) at least one physician visit (prefix=7) with ICD-9-CM code 300 AND one or more prescription for antidepressant or mood stabilizer; OR (v) at least three physician visits (prefix=7) with ICD-9-CM codes 300 or 309 (must be 3 of same dx code). Drugs to treat mood and anxiety disorders include (i) Antidepressants, ATC code N06A; (ii) Benzodiazepine Derivatives Anxiolytics, ATC code N05BA and (iii) Lithium, ATC code N05AN01.

2.3.4.1.2 Algorithm of Identifying Sensitive Periods of Child's First Exposure to Maternal Depression

Initial exposure to maternal depression is defined as the child's age at which maternal depression occurs for the first time during the mother's pregnancy and the child's first five years of life (immediately after birth through kindergarten age). Following children's initial exposure to maternal depression implies children will be excluded from the remaining periods in the study. For instance, children who experience initial maternal depression during their first six months of postpartum and their mothers' depression continued to their second postpartum will be included in the first postpartum initial exposure to maternal depression group and excluded from the second postpartum initial exposure group.

This study investigates the impact of children's initial exposure to maternal depression during a different sensitive period of child development. By investigating the literature on neuroscience, psychology and early child development, I found two streams of literature. Both agree that the crucial development period of a child's early development spans the period from pregnancy through age 5. However, the two streams are different in the time covered in each stage in child development within the common period. For example, some literature in psychology that investigated the effect of maternal depression or maltreatment on early development used the postpartum period to cover the time from birth to 6 months as a distinguished developmental period, while other literature used the postpartum period to cover the time from birth to 12 months. The time periods that are most repeated in the first stream of literature can be broken down into seven sensitive periods in early child's life, which have some developmental characteristics that make them independent of other periods. Briefly, these sensitive periods cover the period from pregnancy onset to kindergarten age as follows: (i)

pregnancy period, which includes the nine months prior to the child's date of birth; (ii) postpartum period1, which includes the time from the child's date of birth to the last day of the sixth month of the child's age; (iii) postpartum period2, which includes the time from the first day of the seventh month of the child's age up to the day before the child's first birthday; (iv) Toddler1, which begins from the child's first birthday up to the last day of the eighteenth month of the child's age; (v) Toddler2, which includes the time from the first day of the nineteenth month of the child's age up to the day before the child's second birthday; (vi) Toddler3, which begins from the child's second birthday up to the child's third birthday; and (vii) Toddler4, which begins from the child's third birthday to kindergarten age. This classification is used in research that focuses on a specific time period or two to compare behavior or development in a specific development domain.

The second stream argues that child development in multiple domains is affected by brain development and early parents' interactions are significant input in human capital formation because it shapes the child's brain architecture. Therefore, the period of crucial development spans the period from the time of conception through the child's fifth birthday. Although brain development spans a few days after conception to early adulthood, the time between conception and a child's third birthday is the most critical time because this is when the foundational neural circuits are set up (Shonkoff and Phillips, 2000). According to this stream, the distinguished periods that include significant effects on brain development and thus child developmental trajectories can be classified as (i) the pregnancy period that starts at conception to birth; (ii) the postnatal period that spans the first year after childbirth; (iii) the toddler period that spans the period from a child's first birthday to the third birthday; and (iv) the preschool period that spans the period from a child's third birthday to the fifth birthday.

I did some analysis and I found the second approach more convenient in this study because there is no difference between the distinguished time periods when identified according to the first stream of literature, inside each broad time period in the second stream. For example, the second approach uses longer periods of time (12 month periods from 0 up to age 1) versus shorter periods of time in the first approach (6 month periods). The results, not included here, indicated that there is no difference between the two shorter periods. Moreover, the frequency of the depressed mother for each time period was small based on the first approach. Therefore, I decided to follow the second stream of literature in specifying the distinguished periods of child development.

2.3.4.2. Health Status at Birth:

The literature on early child health reports many variables that have a significant impact on early children's health and skills formation. This study adopts two credible categories of risk factors that were developed in previous MCHP studies by Brownell et al (2016), Santos et al (2012), and Fransoo et al (2008), based on variables available in the repository database. Child's health status at birth, which are available on the hospital discharge abstract, can be captured by six latent risk factors: (i) low birth weight, which is less than 2500 grams (Cohen and MacWilliam, 1994; Jefferis et al., 2002; Chen et al., 2014; Brownell 2016); (ii) preterm birth, which is birth before 37-week gestational age (Bhutta et al., 2002; De Jong et al., 2012; Chen et al., 2014); (iii) Small for gestational age (SGA), which is defined as at or below the 10^{th} percentile in birthweight from an infant population of the same sex and gestational age (Kramer et al., 2001; Jutte et al., 2010; Chen et al., 2014); (iv) Long hospital stay at birth, which exceeds more than 6 days; (v) baby admitted to the intensive care unit (ICU) for more than 3 days (Fransoo et al., 2008); (vi) Five minute Apgar score, which measures the physiological well

being of new babies at five minutes after birth: a score of zero, one, or two is given for each of five vital signs, which include the infant's heart rate, respiration, muscle tone, reflex, and color, and a total score less than 7 out of 10 reflects a problem (Oreopoulos et al., 2008; Jutte et al., 2010; Roos et al., 2011).

Low birthweight is represented by a binary variable that has value "1" if the child's birth weight was less than 2500 grams and '0' otherwise. Preterm birth is captured by a binary variable that has value "1" if the baby's gestational age is less than 37 weeks at birth and "0" otherwise. Apgar score at 5 minutes after the child's birth is captured by a binary variable "Apgar_5m" which is "1" if the APGAR score is greater than 8 and "0" otherwise. Delivery of child by emergency c-section is captured by a binary variable that has value "1" if the child is delivered by emergency c-section and "0" if naturally delivered or planned c-section. The total length of stay for birth hospitalization is a binary variable that has value "1" if the length of stay at the hospital is greater than 6 days (based on the 90^{th} percentile value).

2.3.4.3 Health Status during Childhood

To capture the child's health status through the early childhood period, after birth through the child's 5^{th} birthday, I used four variables. The main two variables in this study are derived using the John Hopkins ACG® system, which is a statistically valid, case-mix methodology that provides a number of markers derived from individual patient's diagnosis codes history over a one-year period. Based on thousands of International Classification of Disease (ICD) diagnosis codes, the ACG grouping system assigns each diagnosis code to one of 32 diagnosis groups called Aggregate Diagnosis Groups (ADGs). According to MCHP (2013), ICD-9 and ICD-10 codes are assigned to one of 32 different ADGs, of which 8 are considered Major for children, and 24 are considered Minor. Moreover, the diagnosis-to-ADG mapping system in the ACG

software may assign one single ICD diagnosis code to more than one ADG group, which indicates more than one underlying morbidity type (John Hopkins ACG, 2014). The ADGs are not categorized by organ system or disease, but there are five clinical criteria that guide the assignment of each diagnosis code into an ADG: Duration, Severity, Diagnostic certainty, type of etiology, and expected need for specialty care.

Each year, every child could be assigned any number of ADGs based on the number of diagnosis codes assigned to them from physician visits or hospital abstracts. These ADGs are classified into two main sub-groups: Major ADGs and Minor ADGs. Then the number of major ADGs or minor ADGs is counted for each child and summed over the entire year and repeated for the period from birth to the child's 5^{th} birthday. This process gives two variables for each child: the number of major ADGs-year and the number of minor ADGs-year. Children who have more than two major ADGs during childhood are classified as having Major ADG illness (using the 90^{th} percentile value of the major ADGs distribution). That is, the child's major illness is represented by a binary variable that has value "1" if the number of ADGs is greater than two major ADGs and "0" otherwise. Similarly, children who have more than the 90^{th} percentile value of the minor ADGs distribution during early childhood are classified as having a minor ADG. The other two variables that capture the child's health status during early childhood are the number of hospital admissions and the number of physician visits from birth to their 5^{th} birthday. Both are dichotomized based on the 95^{th} percentile value of the number of hospital admissions and 95^{th} percentile value of the number of physician visits from birth to their 5^{th} birthday.

2.3.4.4 Mother Health Status Prior Birth

Heckman and others proposed a model in which capabilities are formed in multistage throughout the child's life and this dynamic process begins with the health of the mother before the time of conception. Consequently, individuals are born with heterogeneous endowments of capabilities (e.g., cognitive capabilities, non-cognitive capabilities, and stock of mental and physical health at birth). The technology of skill formation at any time point in the child's life is given by parental capabilities (e.g., education, genes, income, IQ, psychological factors), parental and government investments in child skills, and the child accumulated capabilities from previous periods. Based on this dynamic model, the child's early development at any time point can be traced back, through recursive backward substitution in the skill formation technology, to a child's initial capabilities, which are affected by a mother's health before conception. The technology of capability production exhibits two important implications: (i) self-productivity, where skills attainment at one stage in a child's life raises skills attainment at later stages in the life cycle and (ii) dynamic complementarity, where skills produced at one stage increase productivity of investment in subsequent stages in a child's life cycle (early investment increases the productivity of later investment). In other words, skills beget skills and abilities beget abilities. Empirically, some studies found that mother pre-pregnancy obesity has an increased risk for negative emotionality and inattentiveness in early childhood (Robinson et al., 2013). There is also evidence that obesity and metabolic syndrome are related to an elevated risk of depression which postulates that there is a potential link between maternal metabolic and fetal brain development in utero[4] (Robinson et al., 2013; Koponen et al., 2008). Therefore, it is reasonable to postulate that the effects of maternal depression on children's early development could be confounded by the mother's health before pregnancy. To date, however, studies of

[4] Metabolic syndrome is a cluster of risk factors for type 2 diabetes and cardiovascular diseases.

maternal depression have not considered the consequences of maternal pre-pregnancy health on child developmental outcomes.

Mother health status is represented by three latent variables through the five years preceding the child's birth. Using the John Hopkins ACG® system, each year, mothers could be assigned any number of ADGs based on the number of diagnosis codes assigned to them from physician visits or hospital abstracts. Moreover, the ADGs are classified into two main sub-groups: Major ADGs and Minor ADGs. Then the number of major ADGs or minor ADGs are counted for each mother and then summed over the entire year and repeated for the five years preceding the child's birth. This process gives two variables for each mother: the number of major ADGs/year and the number of minor ADGs/year. The mother's major illness is dichotomized using the 90^{th} percentile value of the major ADGs distribution and minor illness is dichotomized using the 90^{th} percentile value of the minor ADGs distribution. The third variable is the number of hospital admissions prior to childbirth. It is dichotomized based on the 95^{th} percentile of the distribution of the number of hospital admissions.

2.3.4.5 Child's and Family's Demographic and Socioeconomic Characteristics

Child variables include (i) age as a continuous variable: age in months was calculated using birthdate from the population registry and date of assessment from the EDI database; and (ii) gender, captured by a binary variable which has value "1" if the child is male and "0" if the child is female.

Family risk factors include variables that demonstrate an association with the child's health and educational outcome. I used four latent variables to capture a high-risk environment in which a child develops, and this data is taken from different data sources in the repository. The latent variables include: (i) mother is a teenager at first birth. Some studies found that children of

teen mothers are more likely to have poor health, low educational attainment and social outcomes (Jutte et al., 2010). Teenage mothers are captured by a binary variable that has value "1" if the mother's age at child's birth is less than 19 years, and "0" otherwise. (ii) Family receives income assistance, which is a measure of low-income. Using the Social Assistance Management Information Network (SAMIN) data, receiving employment and income assistance (EIA) is defined as a binary measure. Mothers receiving at least two consecutive months of EIA are identified as receiving EIA (Brownell et al., 2016). (iii) Family with four or more children. Family size was first measured based on counting the number of children born to the same mother, identified by matching the mother's and baby's records using the Manitoba Health Registration Number. These values were dichotomized into two categories with value "1" if the number of children is more than four and "0" otherwise. (iv) Mother's education level. A binary variable of whether the mother has less than high school or graduated from high school at the time of the birth of her child. This information is obtained from Family First Screen and Baby First Screen; (v) Neighbourhood-level socioeconomic status. Socioeconomic status was estimated using the Socioeconomic Factor Index – Version 2 score (SEFI-2). This area-level measure uses the following variables from the Census: average household income (age 15+), percent of single-parent households, unemployment rate (age 15+), and high school education rate for a specified dissemination area (Metge et al., 2009). The continuous SEFI-2 scores are interpreted as follows: Scores less than zero indicate more favorable socioeconomic conditions, while scores greater than zero indicate lower socioeconomic status condition. The SEFI-2 is linked to a mother's postal code to assign a socioeconomic status to the child's family. (vi) Smoking During Pregnancy (categorical). This is captured by a binary variable that has value "1" if the mother reported that she smoked during pregnancy on the BF/FF Screening Form and "0"

otherwise. (vii) Mother's initiation of breastfeeding at birth. This is captured by a binary variable which has value "1" if the mother initiated breastfeeding or breastfeeding mixed with artificial feeding and "0" if the mother initiated only artificial feeding (formula). This data is obtained from the hospital abstract (viii) Mother marital status, captured by a binary variable with value "1" if the mother is married or has a partner and "0" otherwise. This data is obtained from the Manitoba Health Insurance Registry (for all variables, see appendix A "Variable Descriptions" for more details).

2.4. Econometric Methodology

Due to the categorized nature of the EDI domain as a measure of school readiness, linear model regression techniques are not appropriate. In a logistic regression setting, if Y is a binary outcome variable and X is a set of explanatory variables, the conditional mean of Y given X when a logistic distribution is used can be given by

$$\mathbb{E}(Y = 1|X) \equiv pr(Y = 1|X) \equiv \pi(x)$$

And the specific form of the logistic regression model that I used is

$$\pi(x) = \frac{e^{\alpha+\beta' X}}{1 + e^{\alpha+\beta' X}}$$

assuming that $g(x) = ln\left[\frac{\pi(x)}{1-\pi(x)}\right]$,

Then the logit, or log odds, of being not ready/vulnerable in EDI domains can be modeled by the following

$$g(\cdot) = \alpha_0 + \alpha_1\, timing + \alpha_2\, birth + \alpha_3\, childhood + \alpha_4\, mother + \alpha_5 X$$

where $g(\cdot)$ denotes a set of EDI outcomes for child i, i.e., a set of a logit of being vulnerable in EDI domains, including emotional, physical, social, cognitive, and communication outcomes domains. $Timing$ denotes the timing of the first exposure to maternal depression. Timing of maternal depression was entered into the model as a categorical predictor variable: children not exposed to maternal depression at any time point (reference group), children exposed to maternal depression during pregnancy, children exposed to maternal depression during the postnatal period (birth–12 months), children exposed to maternal depression during the toddler period (12 months - 36 months), and children exposed to maternal depression during the preschool period (36 months - 60 months). $birth$ denotes child's health at birth as indicated by low birth weight, preterm birth, 5-minute Apgar, length of stay in hospital after birth, and emergency c-section. $Childhood$ denotes a child's health through early childhood as indicated by a major illness, minor illness, and hospitalizations. $Mother$ denotes mother health status five years before the child's birth as indicated by the mother's major illness, mother's minor illness, and mother's hospitalization. X denotes a set of child and family variables, including child gender, child age, mother's marital status, teenage mother at birth, mother's education, residence, family size, receiving income assistance, and neighborhood-area.

Covariates were selected for inclusion in the models because of their documented or theoretical relevance to the EDI domains. Moreover, covariates were entered into the base multivariable models in four steps: (i) Baseline maternal depression periods (pregnancy, postnatal, toddler, and preschool), (ii) child health status through early childhood and child health status at birth, (iii) mother's health status five years before child's birth (mother's major illness, mother's minor illness, and hospitalization through the five years before child's birth), and (iv) child and family characteristics (child gender, child age, mother's marital status, teenage

mother at birth, mother's education, residence, family size, receiving income assistance, and a neighborhood-level variable - adjusting for socioeconomic status using the SEFI-2). I used the techniques described above for two reasons. First, to construct a series of multivariate analyses to test the association between maternal depression and EDI outcomes after adjusting for the potential confounders or mediators. Second, to test the hypothesis that controlling for child health through early childhood, child's health at birth, mother's health before birth, and social and demographic variables would decrease the measured association between maternal depression timing and EDI outcomes.

The fact that the model specification of maternal depression and child's development trajectories are taken at different points in time may include the possibility of reverse causality, i.e., a child's poor health may cause maternal depression. The model includes some factors that control for a child's health at birth and during childhood to allay the issue related to reverse causality.

2.5. Descriptive statistics:

Tables 2.1 through 2.4 provide descriptive statistics of the working sample. Table 2.1 shows the descriptive statistics of the main EDI domains, the outcome (independent) variable in the study, with frequencies presented separately for children exposed to maternal depression and non-exposed children. The share of children who identified as not ready/vulnerable on the emotional maturity domain is 12.4% (87.6% ready), on the physical health domain is 13.8% (86.3% ready), on social competence domain is 12.5% (87.5% ready), on the cognitive domain is 12.8% (87.2% ready), and on communication domain is 11.5% (88.5% ready). This indicates that most of the children in the sample are identified as 'ready' in one or more EDI domains. The

number of children who identified as not ready in at least one EDI domain is 17,466 (around 29.4%), while the number of children who are ready in at least one domain is 41,947.

The results indicate a significant difference in EDI vulnerability between the children exposed to maternal depression and those in the non-exposed group for all EDI domains. For example, among children exposed to maternal depression, 15.1% were vulnerable in the emotional domain compared to 11.2% in the non-exposed group. However, this does not mean that there is a causal relationship between the exposure to maternal depression and the developmental vulnerability in EDI domains.

Table 2.1: Descriptive Statistics of the Sample on the Main Early Development Instrument (EDI)

	Mean Not ready 'vulnerable'	St.Dev	Num. Ready	Num. not ready	Mean depression exposure	Mean non-depressed exposure	t-statistic
Emotional	0.124	0.329	51898	7327	0.151	0.112	13.29***
Physical	0.138	0.344	51096	8151	0.169	0.124	14.71***
Social	0.125	0.331	51826	7422	0.155	0.112	14.58***
Cognitive	0.128	0.224	51628	7598	0.148	0.120	9.76***
Communication	0.115	0.319	52458	6791	0.120	0.112	2.52**
Number of not ready/vulnerable in EDI domains							
0	0.703	0.457	41790	17623	0.658	0.723	-15.91***
1	0.130	0.337	51664	7749	0.142	0.125	5.65***
2	0.069	0.254	55302	4111	0.078	0.065	6.06***
3	0.042	0.200	56931	2482	0.055	0.036	10.52***
4	0.029	0.169	57665	1748	0.034	0.027	4.41***
5	0.023	0.150	58037	1376	0.028	0.021	5.17***
At least one domain	0.294	0.456	41947	17466	0.338	0.275	15.5***

Notes: Columns 1 and 2 give the overall mean and standard deviation of the EDI variable. Columns 3 and 4 provide the number of children who are ready and not ready on the EDI domains, respectively. Column 5 gives the frequency of the vulnerability for children exposed to maternal depression and the means for non-exposed children are given in column 6. Column 7 gives the t-statistic for the equality of means of both groups. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.2 displays the child's health, mother's health prior to birth, and sociodemographic characteristics of the cohort, with frequencies presented separately for children exposed to maternal depression and non-exposed children. Regarding child health through early childhood, 20% of children in the sample had more than two major ADGs from birth discharge to their 5th birthday, 8% had a number of minor illnesses greater than the 90th percentile of the number of minor ADGs in the population sample, 4% experienced hospital admission counts from birth to their 5th birthday above the 90th percentile, and 2% were admitted to the intensive care unit. For the child health at birth latent variables, approximately 4% of children are low birth (less than 2500 gm), 6% born less at than 37 weeks of gestation, 97% perform well on the five minutes Apgar score at birth, 4.4% stayed longer than six days in the hospital at birth, and 6% had birth complications. For mother's health prior to birth, approximately 19% of mothers had more than two major illnesses during the five years preceding childbirth, 12.8% had minor illnesses above the 90th percentile of the number of minor ADGs distribution, and 3.6% had hospital admission counts above the 90th percentile of the number of hospital admission in the population distribution. For demographic and socioeconomic variables in the sample, the sample was practically evenly divided between boys (50.9%) and girls (49.1%). The mean age of the children at the time of the EDI assessment was 68.75 months. Approximately 40% of mothers were married at birth, 15.5% had less than high school degree, 25.8% were teenagers at birth, and 16.4% smoked during pregnancy. On average, 18.8% of the children's families had four or more children at birth, 14.3% received income assistance for three consecutive months from birth to the child's 5th birthday, and 47% were living in a more favorable SES neighborhood area.

Table 2.2: Descriptive Statistics of the Study Explanatory Variables

	Mean	St.Dev	Exposed to depression	No-exposure to depression	t-statistic
Timing of first exposure to maternal depression					
Pregnancy (In uterus)	0.046	0.210	---	---	---
Postnatal (birth -12 months)	0.053	0.224	---	---	---
Toddler (12 months-36 months)	0.093	0.290	---	---	---
Preschool (36 months-60 months)	0.119	0.324	---	---	---
Child's health through childhood					
Major illness (2+ major ADGs)	0.207	0.405	0.2390	0.1935	12.57***
Minor illness (90th+ minor ADGs	0.087	0.282	0.1195	0.0730	18.48***
Child hospital admission (>95th)	0.036	0.186	0.0404	0.0337	4.05***
ICU admission during childhood	0.019	0.137	0.0218	0.0181	3.03***
Child's health at birth					
Low Birth Weight (< 2500g)	0.038	0.191	0.0432	0.0358	4.32***
Preterm (< 37 weeks)	0.061	0.061	0.0717	0.0570	6.85***
5-minutes Apgar (>=8)	0.966	0.181	0.9638	0.9669	-1.86*
5-minutes Apgar (1-10)	8.86	0.65	8.8476	8.8647	-2.93***
Length of hospital staying (>6days)	0.044	0.205	0.0528	0.0401	6.92***
Emergency c-section	0.031	0.173	0.0329	0.0302	1.74*
Birth complications	0.065	0.246	0.0703	0.0622	3.67***
Mother health 5 years before the birth					
Major illness (>2+ ADGs)	0.192	0.394	0.2820	0.1540	36.69***
Minor illness (>90th minor ADGs)	0.128	0.334	0.2294	0.0839	49.73***
Mother hospital admission (>95th)	0.036	0.186	0.0886	0.0463	20.09***
Child, Mother, family characteristics					
Child's age (in months)	68.75	3.56	68.812	68.729	2.62**
Child gender (male=1)	0.509	0.500	0.5069	0.5093	0.54
Breast feeding initiation	0.823	0.382	0.7815	0.8408	-17.37***
Mother married	0.400	0.490	0.3291	0.4309	-23.30***
Mother less than HS	0.155	0.362	0.1867	0.1416	13.94***
Teenage mother at birth	0.258	0.437	0.3210	0.2305	23.24***
Smoking during pregnancy	0.164	0.370	0.2289	0.1362	28.16***
Neighbourhood-area SES	0.470	0.499	0.4367	0.4843	-10.66***
Urban residence	0.505	0.500	0.6263	0.4531	39.19***
Family size (> 4)	0.188	0.350	0.2059	0.1804	7.310***
Family receive income assistance	0.143	0.350	0.2076	0.1153	29.67***

Notes: Columns 1 and 2 give the overall mean and standard deviation of the model variables. Column 3 and 4compares depressed mothers and their children with non-depressed mothers and their children in the sample under study. Column 5 gives the t-statistic for the equality of means of both groups. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Columns 3 and 4 in table 2.2 compare depressed mothers and their children with non-depressed mothers and their children in the sample under study. As is evident from the last column of the table, these groups are quite different. On average, depressed mothers are more likely to be less healthy prior to their pregnancy (more likely to have a major and minor illness), to be a single mom, less educated, and teenager at birth. Their children are more likely to be less healthy at birth and through childhood. Compared to children who were not exposed to maternal depression during any period, children exposed to depression were more likely to be less than normal weight at birth, to be born before reaching normal gestation age, to have lower Apgar scores, and to have more birth complications. Families of children in the depressed group are more likely to live in less favorable SES neighborhood areas and to receive income assistance. Finally, there were 2739 children (4.6% of the total sample) exposed to maternal depression during pregnancy (in the uterus), 3135 (5.28% of the total sample) children exposed to depression during the postnatal period, 5510 (9.27% of the total sample) children exposed to depression during the toddler period, and 7094 children (11.94% of the total sample) exposed to depression during the preschool period.

Table 2.3 presents the results of crosstabulations comparing all the EDI domains across family socioeconomic status and child gender. Children identified as ‘not ready’/’vulnerable’ on the domains of the EDI generally came from families with less favorable neighborhood socioeconomic status than children identified as ‘ready’. Results of Chi-square analyses showed that children identified as ‘ready’ and ‘not ready’ differed significantly by neighborhood-area socioeconomic status on each measure of the EDI domain. For example, among children who are classified as not ready/vulnerable in one or more of the EDI domains, 63.32% of them were in families in low SES neighborhoods compared to 36.68% in families in favorable SES

neighborhoods. The proportion of children whose families were living in less favorable SES neighborhoods among children who classified as not ready in the separate EDI domain were 61.65% in the emotional domain, 66.86% in the physical health domain, 65.60% in the social competence domain, 68.86% in the cognitive domain, 65.59% in the communication domain. The associated p-value of the Chi-Square statistic is less than 0.0001, which means that there is significant evidence of an association between child developmental vulnerability in EDI domains and family socioeconomic status. Children from all SES backgrounds can be vulnerable. However, children from low SES backgrounds may have a far greater risk of being vulnerable.

A greater proportion of boys than girls are identified as ‘not ready’ on all sub-domains of the EDI. Results of Chi-square statistics showed that children identified as ‘ready’ and ‘not ready’ differed significantly by gender on each sub-domain of the EDI. Among children who identified as not ready/vulnerable in one EDI domain or more, 11136 (63.76%) children were boys while 6330 (36.24%) were girls. The proportion of boys among children who classified as not ready/vulnerable in separate EDI domains were 73.84% in the emotional domain, 63.15% in the physical health domain, 69.95% in the social competence domain, 64.21% in the cognitive and language domain, and 64.78% in the communication and knowledge domain. The associated p-value is less than 0.0001, which means that there is significant evidence of an association between child developmental vulnerability in EDI domains and child gender. The other chi-square statistics (Likelihood Ratio Chi-Square Continuity Adj. Chi-Square Mantel-Haenszel Chi-Square) have similar values and are asymptotically equivalent.

Table 2.3: Descriptive Statistics of EDI by Neighborhood-area SES and Child Gender

	Socioeconomic status			Child Gender		
EDI domain	Less favorable	More favorable	Chi-Square	Boys	Girls	Chi-Square
One or more EDI domain						
Ready	20396 (48.70%)	21485 (51.30%)	1056.3652	19079 (45.48%)	22868 (54.52%)	1647.76
Not ready	11037 (63.32%)	6393 (36.68%)	p<0.0001	11136 (63.76%)	6330 (36.24%)	p<0.0001
Emotional Domain						
Ready	26801 (51.73%)	25011 (48.27%)	253.0481	24697 (47.59%)	27201 (52.4%)	1770.0100
Not ready	4507 (61.65%)	2804 (38.35%)	p<0.0001	5410 (73.84%)	1917 (26.16%)	p<0.0001
Physical Domain						
Ready	25884 (50.74%)	25130 (49.26%)	731.1753	24974 (48.88%)	26122 (51.12%)	572.6656
Not ready	5436 (66.86%)	2695 (33.14%)	p<0.0001	5147 (63.15%)	3004 (36.85%)	p<0.0001
Social domain						
Ready	26463 (51.15%)	25278 (48.85%)	543.6291	24927 (48.10%)	26899 (51.90%)	1240.925
Not ready	4858 (65.60%)	2547 (34.40%)	p<0.0001	5192 (69.95%)	2230 (30.05%)	p<0.0001
Cognitive Domain						
Ready	26084 (50.61%)	25454 (49.39%)	884.2983	25227 (48.86%)	26401 (51.14%)	624.5107
Not ready	5224 (68.86%)	2362 (31.14%)	p<0.0001	4879 (64.21%)	2719 (35.79%)	p<0.0001
Communication Domain						
Ready	26876 (51.32%)	25491 (48.68%)	490.4982	25722 (49.03%)	26736 (50.97%)	596.2706
Not ready	4447 (65.59%)	2333 (34.41%)	p<0.0001	4399 (64.78%)	2392 (35.22%)	p<0.0001

Notes: (i) The Pearson chi-square statistic is used to assess the association between the child's gender and the vulnerability in the EDI domains. (ii) The values printed under the cell count are the row percentage (iii) the null hypothesis is H_0: $no\ association\ between\ variables$. p-value < 0.05 means that there is significant evidence of an association between the two variables.

2.6. Regression Results

Table 2.4 presents the unadjusted and the fully adjusted ORs for being vulnerable/not ready on the emotional maturity domain, as an outcome of the timing of the first exposure to maternal depression. The unadjusted model (model 1) shows that maternal depression has a negative effect on a child's emotional maturity skills across all four time periods. Compared to children who were not exposed to maternal depression at any time, children exposed to maternal depression in the uterus (pregnancy) are 1.58 more likely to be not ready on the emotional maturity domain at kindergarten age. Children exposed to maternal depression in the postnatal period had significantly higher odds (34%) of developmental vulnerability on emotional maturity skills. Children exposed to maternal depression during the toddler period had significantly higher odds (29% increase) to be not ready in emotional maturity. Children exposed to maternal depression during the preschool period had significantly higher odds (48% increase) of development vulnerability on their emotional maturity skills.

After adjusting for child's health at birth, child's health through early childhood (model 2), and mother's health status during the five years before birth (model 3), the negative effect of maternal depression on child's emotional maturity skills remained statistically significant across the four time periods. However, the effect of maternal depression on the emotional domain was attenuated in all periods. The odds of being not ready on emotional maturity are 18-38% higher across the different time points of maternal depression compared with 29-58% in the unadjusted model.

Table 2.4: Unadjusted and Fully-adjusted Odds Ratio of being Developmentally Vulnerable in the Emotional Maturity Domain

	Model 1 Unadjusted OR	Model 2 Partially adj OR	Model 3 Partially adj OR	Model 4 Fully adj OR
Timing of first exposure to maternal depression				
Pregnancy (In uterus)	1.58***	1.49***	1.38***	1.32 [1.18-1.49] ***
Postnatal (birth -12months)	1.34***	1.28***	1.21***	1.13 [1.01-1.26] **
Toddler I (12months-36m)	1.29***	1.24***	1.18***	1.10 [1.01-1.20] **
Preschool (36months-60m)	1.48***	1.42***	1.35***	1.23 [1.14-1.32] ***
Child's health through childhood				
Major illness (2+ major ADGs)		1.34***	1.32***	1.17 [1.10-1.23] ***
Minor illness (90th+ minor ADGs		1.32***	1.27***	1.26 [1.18-1.34] ***
Hospital admission (>95th)		1.70***	1.68***	1.44 [1.28-1.61] ***
Child's health at birth				
Low Birth Weight (< 2500g)		1.19**	1.19**	1.33 [1.15-1.54] ***
Preterm (< 37 weeks)		1.02	0.99	0.93 [0.83-1.06]
5-minutes Apgar (>=8)		0.99	0.98	0.98 [0.86-1.12]
Length of hospitalization (>6)		1.19**	1.18**	1.12 [0.97-1.28]
Emergency c-section		1.10	1.10	1.20 [1.05-1.38] **
Mother health 5 years before the birth				
Major illness (>2+ ADGs)			1.15***	1.09 [1.03-1.15] ***
Minor illness (>90th minor ADGs)			1.17***	1.08 [1.01-1.15] **
Hospital admission (>95th)			1.19***	1.12 [1.01-1.24] **
Child, and family characteristics				
Child's age (in months)				0.97 [0.96--0.98] ***
Child gender (male=1)				3.14 [2.97-3.32] ***
Breastfeeding initiation				0.89 [0.84-0.95] ***
Mother married				0.68 [0.65-0.73] ***
Mother has less than HS				1.22 [1.13-1.32] ***
Teenage mother				1.21 [1.13-1.29] ***
Smoking during pregnancy				1.06 [0.99-1.14] *
Neighbourhood-area SES				0.84 [0.71-0.89] ***
Urban residence				1.06 [1.01-1.12] **
Family size (>4)				1.19 [1.12-1.27] ***
The family receive income assistance				1.29 [1.19-1.39] ***

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) OR is the odds ratios, CI is the confidence interval, (iv) vulnerability is the scoring in the bottom 10^{th} percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Column 4 shows the further adjustment for a child's characteristics (age, sex), family variables (mother's marital status, teenage mother, a mother with less than high school education, family size, family on income assistance), and the neighborhood-area socioeconomic status. Although there is a statistically significant association between maternal depression and child's emotional maturity across all maternal depression time points, the odds of vulnerability in emotional skills are moderated after adjusting for these variables. The odds of being not ready on emotional maturity skills are 10-32% greater across the different maternal depression time points. Compared to children who were not exposed to maternal depression during any period, the odds of developmental vulnerability on emotional skills are 32%, 13%, 10%, and 35% greater for children exposed to maternal depression during pregnancy, postnatal, toddler, and preschool time points, respectively. However, the influence of maternal depression appears to be stronger in the pregnancy period, followed by the preschool period than the other two time points. According to the fully-adjusted ORs (model 4), children exposed to maternal depression in the uterus could be as little as 1.32 times or as much as 1.49 times more likely to be not ready on emotional skills at kindergarten age compared to those who were not exposed to maternal depression at any time point. Similarly, children exposed to maternal depression during the preschool period could be as little as 1.14 times or as much as 1.32 times more likely to be vulnerable in emotional skills compared with those who were not exposed to maternal depression during any time point.

The results show that the child's health through early childhood has a statistically significant effect on the child's emotional domain. The odds of being developmentally vulnerable in emotional skills are 17-70% higher for children with health vulnerability during early childhood. According to fully-adjusted model, the odds of being vulnerable are 1.17 times

greater for children with a major illness (as indicated by 2+ major ADGs) than those without major illness in early childhood, 1.26 times greater for children with minor illness (as indicated by 95th percentile minor ADGs) than those without minor illness in early childhood, and 1.44 times greater for children who required hospitalization in early childhood compared with children with no hospitalization. Similarly, the odds of being vulnerable at kindergarten age in emotional skills are statistically significantly higher with a child's health status at birth. The odds of being vulnerable are 33% higher for low birth weight than normal-weight children, 20% higher for children delivered through emergency c-section than natural delivery.

The results also indicate that a child's mother's health status during the five years before the child's birth has a significant effect on a child's emotional skills at kindergarten age even after controlling for other covariates. The odds of being vulnerable on the emotional domain is 9% greater for children whose mother had the major illness before childbirth, 8% greater for children whose mother had minor illness prior child's birth, and 12% greater for children whose mother was hospitalized more frequently than the 95th percentile of Manitobans compared to children whose mother did not experience any of these health problems before their birth.

Among the child's individual-level factors, a child's sex has a strong effect on emotional maturity skills. The association appears to be stronger in boys than in girls. The odds of being vulnerable are 3.14 times greater for boys than girls and decrease with child age (OR= 0.97). The fully-adjusted ORs indicate that a child's family characteristics have a powerful influence on the child's performance on the child's emotional skills at kindergarten age. The odds of being vulnerable in the emotional domain are 22% greater for children whose mothers do not have high school than those whose mothers have a secondary degree or higher, 21% greater for children whose mothers were teenagers at their birth compared to those whose mothers were not

teenagers. In contrast, children whose mothers were married are less likely to be vulnerable in the emotional domain than those whose mothers are not married, and children whose mothers initiated breastfeeding at birth are less likely to be vulnerable in the emotional domain than those whose mothers did not initiate breastfeeding. The odds of being vulnerable on the emotion domain are 29% greater for children in families on income assistance than for those not on income assistance, and 19% greater for children having more than four siblings at birth compared with those having less than four siblings. The neighborhood-area socioeconomic status of a child's family was significantly associated with child emotional skills. Children whose family has favorable neighborhood-area socioeconomic status (as indicated by SEFI) have lower odds of being not ready in emotional skills.

Put together, these results suggest a significant and strong influence of the child's health and family environment on children's school readiness at kindergarten age. The reduction in the magnitude of the odds of being vulnerable in emotional skills after controlling for child health and family characteristics supports the hypothesis that these variables could have a potential confounding effect in the association between maternal depression and child's school readiness.

Table 2.5 shows the effect of maternal depression timing on children's physical health and well- being domain at kindergarten age. The adjusted ORs (model 1) indicate that exposure to maternal depression for the first time has a negative effect on the child's physical health and well-being domain across all time points. Moreover, the strongest effect of maternal depression exposure on children's physical health was found during pregnancy and the preschool period. However, maternal depression seemed to exert a greater effect on children's physical health domains during the preschool period than did the pregnancy period and other periods of maternal depression exposure.

Table 2.5: Unadjusted and Fully-adjusted odds Ratio of Being Developmentally Vulnerable in the Physical Health Domain.

	Model 1 unadjusted OR	Model 2 partially adj OR	Model 3 partially adj OR	Model 4 fully adj OR
Timing of first exposure to maternal depression				
Pregnancy (In uterus)	1.440***	1.362***	1.230***	1.157[1.035-1.294]**
Postnatal (birth -12months)	1.317***	1.27***	1.186***	1.074[0.965-1.195]
Toddler I (12months-36m)	1.349***	1.306***	1.23***	1.096[1.008-1.191]**
Preschool (36months-60m)	1.606***	1.559***	1.459***	1.228[1.141-1.322]***
Child's health through childhood				
Major illness (2+ major ADGs)		1.246***	1.222***	1.104[1.046-1.64]***
Minor illness (90th+ minor ADGs		1.144***	1.086**	1.141[1.069-1.219]***
Child hospital admission (>95th)		2.490***	2.44***	2.028[1.825-2.253]***
Child's health at birth				
Low Birth Weight (< 2500g)		1.244***	1.251***	1.305[1.137-1.497]***
Preterm (< 37 weeks)		1.203***	1.75***	1.08[0.096-1.212]
5-minutes Apgar (>=8)		0.847***	0.844***	0.768[0.678-0.870]***
Length of hospitalization(>6 days)		1.358***	1.351***	1.318[1.159-1.500]***
Emergency c-section		0.973***	0.977	1.1450.997-1.316]*
Mother health before birth				
Major illness (>2+ ADGs)			1.225***	1.089[1.033-1.147]***
Minor illness (>90th minor ADGs)			1.170***	1.022[0.961-1.088]
Mother hospital admission (>95th)			1.28***	1.109[1.009-1.219]**
Child, Mother, family char				
Child's age (in months)				0.962[0.955-0.968]***
Child gender (male=1)				1.844[1.753-1.940]***
Breastfeeding initiation				0.878[0.827-0.933]***
Mother married				0.615[0.579-0.652]***
Mother has less than HS				1.203[1.122-1.290]***
Teenage mother				1.398[1.318-1.483]***
Smoking during pregnancy				1.175[1.101-1.254]***
Neighbourhood-area SES				0.751[0.712-0.793]***
Urban residence				1.037[0.986-1.092]
Family size (>4)				1.687[1.590-1.790]***
The family receive income assistance				1.508[1.404-1.620]

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) OR is the odds ratios, CI is the confidence interval, (iv) vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

The odds of being vulnerable in physical health at kindergarten age were 44% and 60% greater in children exposed to maternal depression for the first time during pregnancy (in utero) and during preschool time point, respectively, compared to those who were not exposed to maternal depression during any time point. Similarly, children exposed to maternal depression during the postnatal or toddler periods had significantly higher odds of being developmentally vulnerable in physical health and the well- being domain by 31% and 34%, respectively, compared to those who were not exposed to maternal depression during any time point.

After controlling for child health through early childhood and child health at birth, the odds of being vulnerable in the physical health domain for children exposed to maternal depression for the first time remained statistically significant across all the maternal depression periods. Also, the strongest effect of maternal depression on the child's physical health domain remained through the pregnancy and preschool time points. After further controlling for child and family characteristics, the fully-adjusted odds of being not ready on physical health remained significant across the four periods and stronger for children exposed to maternal depression during pregnancy and preschool. For example, children exposed to maternal depression during pregnancy and preschool had odds of being developmentally vulnerable on physical health at kindergarten age 16% and 23% greater than children who were not exposed to maternal depression during any time.

The results of the fully-adjusted ORs indicate that children's vulnerability in the physical health domain can be traced back to biological vulnerability at birth and the child's mother's health before birth. For example, the odds of being vulnerable in physical health at kindergarten age are 1.10 times greater for children with major illness than children without major illness in early childhood, 1.14 times greater for the children with minor illness than those without minor

illness in early childhood, and 2.03 times greater for children who were frequently hospitalized in early childhood than those who were not. Similarly, the odds of being vulnerable in the physical health domain are 30% greater for low birth weight children than normal birth weight children, 31% greater for children who spent more than six days in hospital after birth compared with those who had a normal length of stay in hospital after birth, 15% greater for children who were delivered through emergency c-section than those born normally. In contrast, children who had a high 5-minute Apgar score were less likely to be not ready on physical health compared with those who had low 5-minute Apgar scores.

Among child individual-level variables, the odds of being vulnerable in physical health appears to be stronger in boys than girls. Boys are 1.84 times more likely to be vulnerable in the physical health domain than girls and decrease with child age (OR = 0.96). Among child family characteristics, children whose mothers did not have high school had significantly higher odds of being vulnerable in physical health by 20% than those whose mothers have a high school or higher degree. Children whose mothers were teenagers at their birth had odds of being vulnerable in physical health 40% higher than those whose mothers were not teenagers. In contrast, children whose mothers-initiated breastfeeding at their birth are less likely to be vulnerable in physical health (OR= 0.87) than those whose mothers did not initiate breastfeeding at birth. Children in two-parent families are less likely to be not ready on the physical domain (OR= 0.65) than those in single-mother families. the odds of being vulnerable on the physical health domain are 69% greater for children who had more than four siblings at birth than those who had less than four siblings, and 51% greater for children in families on income assistance than those not on income assistance.

Table 2.6: Unadjusted and Fully-adjusted Odds Ratio of Being Developmentally Vulnerable in the Social Competence Domain.

	Model 1 Unadjusted OR	Model 2 partially adj OR	Model 3 partially adj OR	Model 4 fully adj OR
Timing of first exposure to maternal depression				
Pregnancy (In uterus)	1.544***	1.464***	1.333***	1.240[1.108-1.389]***
Postnatal (birth -12m)	1.439***	1.385***	1.297***	1.185[1.064-1.320]***
Toddler I (12m-36m)	1.313***	1.269***	1.199***	1.072[0.983-1.169]
Preschool (36m-60m)	1.570***	1.519***	1.425***	1.224[1.135-1.321]***
Child's health through childhood				
Major illness (2+ major ADGs)		1.296***	1.272***	1.128[1.068-1.193]***
Minor illness (90th+ minor ADGs		1.203***	1.144***	1.158[1.083-1.238]**
Child hospital admission (>95th)		1.973***	1.936***	1.593[1.425-1.178]***
Child's health at birth				
Low Birth Weight (< 2500g)		1.145**	1.152**	1.240[1.072-1.434]***
Preterm (< 37 weeks)		1.067	1.044	0.970[0.859-1.096]
5-minutes Apgar (>=8)		0.918	0.916	0.877[0.769-1
Length of hospital stay(>6 days)		1.322***	1.316***	1.243[1.086-1.422]**
Emergency c-section		1.047	1.052	1.178[1.023-1.356]**
Mother health 5 years before the birth				
Major illness (>2+ ADGs)			1.216***	1.112[1.053-1.174]***
Minor illness (>90th minor ADGs)			1.179***	1.062[0.997-1.132]*
Mother hospital admission (>95th)			1.186***	1.069[0.968-1.180]
Child and family characteristics				
Child's age (in months)				0.962[0.968-0.969]***
Child gender (male=1)				2.590[2.453-2.735]***
Breastfeeding initiation				0.807[0.751-0.859]***
Mother married				0.649[0.611-0.690]***
Mother has less than HS				1.120[1.040-1.206]***
Teenage mother				1.363[1.281-1.450]***
Smoking during pregnancy				1.150[1.074-1.231]***
Neighbourhood-area SES				0.743[0.703-0.786]***
Urban residence				1.106[1.049-1.166]***
Family size (>4)				1.309[1.229-1.395]***
The family receive income assistance				1.388[1.288-1.497]***

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) OR is the odds ratios, CI is the confidence interval, (iv) vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.6 shows results of unadjusted ORs (model 1) indicate that the odds of being vulnerable/not ready on the social competence domain at kindergarten age are statistically significantly higher across all the time points of maternal depression.

Compared to children who were not exposed to maternal depression during any time point, the odds of being developmentally vulnerable in the social domain are 54% greater for children exposed to maternal depression during the pregnancy period, 44% greater for children exposed during the postnatal period, 31% greater for children exposed to maternal depression during the toddler period, and 57% greater for children exposed during the preschool period. Exposure to maternal depression during pregnancy and preschool periods seemed to exert a greater effect on physical health than exposure during postnatal and toddler periods.

After adjusting for child health through early childhood, child health at birth, and mother's health before birth (models 2 & 3), the odds of being vulnerable in the social competence domain remained statistically significant (19%-51% increase) across all four time periods Also, the strongest effect of maternal depression on the odds of being not ready on the social competence domain remained throughout the pregnancy period and the preschool period.

After further adjustment for child and family characteristics, the odds of being vulnerable in social competence at kindergarten age remained statistically significant across all maternal depression time points except during the toddler period. However, the odds of being vulnerable are attenuated after the adjustments. The odds ratios indicate that the odds of being vulnerable/at risk in social competence at kindergarten age are statistically significantly higher with child biological vulnerability across all models. The odds of being not ready on the social competence domain are 14-97% higher for children with vulnerable health through early childhood. For example, the fully-adjusted odds of being not ready on social competence are 1.13 times greater

for children with major illness than those without major illness, 1.16 times greater for children with a minor illness, and 1.59 times greater for children who required frequent hospitalization during early childhood. Similarly, the odds of being not ready on social competence are 1.24 times greater for children who had low birth weight than those who had normal birth weight, and for children who required an extended hospital stay at birth. Children whose mothers had health issues before the child's birth had higher odds of being not ready on the social competence domain at kindergarten age. The odds of being vulnerable in the social competence domain are 1.11 times greater for children whose mothers had a major illness than those whose mothers did not have a major illness, 1.07 times greater for children whose mothers had minor illness than those whose mothers did not have minor illness during the five years before birth.

Among child individual-level variables, the odds of being not ready on social competence at kindergarten age are 2.6 times greater for boys than for girls and decrease with child age (OR=0.96). Among family variables, the odds of being not ready on social competence are 1.12 times greater for children whose mothers did not finish high school compared to those whose mothers had a high school or higher degree, and 1.36 times greater for children whose mothers were teenagers at their birth than those their mothers were adults. In contrast, children in two-parent families are less likely to be vulnerable in social competence than children in single-mother families, and children whose mothers initiated breastfeeding at birth are less likely to be not ready in social competence. Similarly, the odds of being not ready on social competence are 1.39 times greater for children in families on income assistance than for those not on income assistance, and 1.31 times greater for children who had more than four siblings at birth compared to those with less than four siblings. These results confirm the significant effect of the childhood family environment on their school readiness.

Table 2.7: Unadjusted and Fully-adjusted Odds Ratio of Being Developmentally Vulnerable in the language and Cognitive Domain.

	Model 1 Unadjusted OR	Model 2 Partially adj OR	Model 3 Partially adj OR	Model 4 Partially adj OR
Timing of first exposure to maternal depression				
Pregnancy (In uterus)	1.279***	1.207***	1.089	1.032[0.916-1.161]
Postnatal (birth -12months)	1.334***	1.284***	1.197***	1.096[0.983-1.223]
Toddler I (12months-36m)	1.275***	1.231***	1.157***	1.043[0.957-1.137]
Preschool (36months-60m)	1.337***	1.240***	1.204***	1.006[0.930-1.088]
Child's health through childhood				
Major illness (2+ major ADGs)		1.310***	1.286***	1.162[1.099-1.228]***
Minor illness (90th+ minor ADGs)		1.147***	1.085***	1.155[1.080-1.237]***
Child hospital admission (>95th)		2.640***	2.589***	2.063[1.854-2.296]***
Child's health at birth				
Low Birth Weight (< 2500g)		1.264***	1.271***	1.368[1.106-1.577]***
Preterm (< 37 weeks)		1.182***	1.156**	1.047[0.929-1.181]
5-minutes Apgar (>=8)		0.916	0.913	0.850[0.745-0.970]**
Length of hospital stay(>6 days)		1.260***	1.253***	1.217[1.055-1.392]**
Emergency c-section		0.947	0.951	1.099[0.950-1.271]
Mother health 5 years before the birth				
Major illness (>2+ ADGs)			1.168***	1.025[0.971-1.083]
Minor illness (>90th minor ADGs)			1.214***	1.074[1.008-1.145]**
Mother hospital admission (>95th)			1.320***	1.125[1.020-1.240]**
Child and family characteristics				
Child's age (in months)				0.925[0.918-0.931]***
Child gender (male=1)				1.956[1.856-2.062]***
Breastfeeding initiation				0.751[0.707-0.798]***
Mother married				0.701[0.660-0.745]***
Mother has less than HS				1.278[1.190-1.373]***
Teenage mother				1.549[1.458-1.646]***
Smoking during pregnancy				1.006[0.939-1.077]
Neighborhood-area SES				0.690[0.652-0.730]***
Urban residence				0.994[0.943-1.048]
Family size (>4)				1.709[1.060-1.815]***
The family receive income assistance				1.502[1.395-1.618]***

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) OR is the odds ratios, CI is the confidence interval, (iv) vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.7 shows the unadjusted and fully-adjusted ORs of being vulnerable in the language and cognitive domain at kindergarten age as an outcome variable of the timing of maternal depression. The unadjusted odds show that children exposed to maternal depression had significantly higher odds of being vulnerable in language and cognitive domain across all the periods. For example, the odds of being not ready on language and cognitive domain are 28% greater for children exposed to maternal depression during pregnancy compared to children who were not exposed to maternal depression during any time point, 33% greater for children exposed to maternal depression during the postnatal period, 27% greater for children exposed to maternal depression during the toddler period, and 34% greater for children exposed to maternal depression during the preschool time period.

After adjusting for child health through early childhood and mother's health before birth, the odds of being vulnerable in language and cognitive domain remained significant across all initial exposure periods except during pregnancy. However, once the model is adjusted for child and family characteristics, the odds of being not ready were statistically insignificant across all the periods of initial exposure to maternal depression.

Child health through early childhood and at birth has a significant effect on the language and cognitive domain even after adjusting for the other covariates in the model. The odds of being vulnerable in the physical health domain are 16% greater for children without major illness, 15% greater for children with minor illness than those without minor illness, and 106% greater for children with a hospital stay during early childhood than those without hospitalization. Similarly, the odds of being not ready on the language and cognitive domain are 37% higher for low birth weight children than normal birth weight at birth, 22% greater for children who required longer hospital stay at birth than those with normal stay.

Among child individual-level variables, the odds of being vulnerable in the language and cognitive domain at kindergarten age are 1.96 times greater for boys than girls and decrease with child age (OR= 0.92). Among child family characteristics, the odds of being vulnerable in the language and cognitive domain are 28% greater for children whose mother with less than high school education, and 54% greater for children whose mothers were teenagers at birth than those whose mothers were not teenagers. Children whose mothers initiated breastfeeding at birth are less likely to be vulnerable in the language and cognitive domain. Children whose mother is married were less likely to be not ready on the cognitive domain than those whose mothers were single. Similarly, the odds of being vulnerable in the language and cognitive domain are 71% greater for children in families with more than four children than those in families with less than four children at birth, 50% greater for children in families on income assistance than those in families not on income assistance. Regarding family socioeconomic status (as indicated by SEFI2), children in families with favorable socioeconomic status are less likely to be vulnerable in the cognitive domain (OR= 0.69).

Table 2.8 indicates the unadjusted and fully-adjusted ORs of being vulnerable in communication and knowledge skills across the different time points of maternal depression. The odds of being not ready in the communication domain are not statistically significant across periods except for the preschool period. Children exposed to maternal depression during the preschool period for the first time are 1.19 times more likely to be vulnerable in the communication domain. After adjusting for child health and mother health before birth, the odds of being not ready in the communication domain remained insignificant across all time points of initial exposure to maternal depression, except for the preschool period. However, after further

adjusting for child and family characteristics, there was no significant effect of maternal depression across all the timing points of depression.

Child health through early childhood has a statistically significant effect on the child's communication skills. The odds of being developmentally vulnerable in communication skills are 9-58% greater for children with any health problem through early childhood. According to the fully-adjusted OR, the odds of being vulnerable in communication skills are 16% higher for children with minor illness than those without minor illness and 104% higher for children who were frequently hospitalized during early childhood. Similarly, the odds of being vulnerable in communication skills are 39% greater for low birth weight children than those with normal weight at birth, and 26% greater for children who required longer hospital stays at birth than those who had a normal length of stay at birth. Among child individual-level variables, the odds of being vulnerable in the communication domain are 1.96 times greater for boys than for girls and decline with age (OR =0.95). Among family characteristics, the odds of being not ready on the communication domain are 1.45 times higher for children whose mothers did not have a high school degree than those whose mothers had a high school degree or higher, and 1.19 times higher for children whose mothers were teenagers at birth than those whose mothers were adults. In contrast, children whose mothers initiated breastfeeding at birth are less likely to be vulnerable in the communication domain than those whose mothers did not initiate breastfeeding at birth. Children whose mothers were married are less likely to be not ready in the communication domain than those whose mothers were single. Similarly, the odds of being vulnerable in communication skills are 2.11 times greater for children whose families have more than four children than those whose families have less than four children at birth (family size).

Children in families on income assistance are 1.49 times more likely to be not ready on the communication domain than those in families not on income assistance.

Table 2.8: Unadjusted and Fully-adjusted Odds Ratio of Being Developmentally Vulnerable in the Communication and Knowledge Domain.

	Model 1 Unadjusted OR	Model 2 Partially adj OR	Model 3 Partially adj OR	Model 4 Fully adj OR
Timing of first exposure to maternal depression				
Pregnancy (In uterus)	1.063	1.011	0.961	0.931 [0.820-1.057]
Postnatal (birth -12months)	1.015	0.982	0.948	0.889 [0.788-1.002]
Toddler I (12months-36m)	1.049	1.017	0.987	0.914 [0.833-1.004]
Preschool (36months-60m)	1.188***	1.151***	1.114***	0.968 [0.892-1.051]
Child's health through childhood				
Major illness (2+ major ADGs)		1.179***	1.168***	1.058 [1.000-1.120]
Minor illness (90th+ minor ADGs		1.119***	1.091**	1.159 [1.080-1.245]***
Child hospital admission (>95th)		2.587***	2.549***	2.041 [1.828-2.278]***
Child's health at birth				
Low Birth Weight (< 2500g)		1.296***	1.297***	1.388 [1.198-1.607]***
Preterm (< 37 weeks)		1.153**	1.136**	1.034 [0.913-1.171]
5-minutes Apgar (>=8)		0.978	0.976	0.910 [0.792-1.045]
Length of hospitalization (>6)		1.299***	1.295***	1.260 [1.097-1.447]***
Emergency c-section		0.950	0.952	1.070 [0.921-1.244]
Mother health 5 years before the birth				
Major illness (>2+ ADGs)			1.090***	0.963 [0.910-1.019]
Minor illness (>90th minor ADGs)			1.062*	0.959 [0.895-1.027]
Mother hospital admission (>95th)			1.284***	1.128 [1.017-1.251]***
Child and family characteristics				
Child's age (in months)				0.956 [0.949-0.963]***
Child gender (male=1)				1.962 [1.857-2.072]***
Breastfeeding initiation				0.713 [0.669-0.759]***
Mother married				0.900 [0.853-0. 963]***
Mother has less than HS				1.447 [1.343-1.558]***
Teenage mother				1.191 [1.116-1.270]***
Smoking during pregnancy				0.918 [0.845-0.979]***
Neighbourhood-area SES				0.779 [0.735-0.826]***
Urban residence				0.980 [0.928-1.036]
Family size (>4)				2.117 [1.991-2.251]
The family receive income assistance				1.497 [1.384-1.620]

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) OR is the odds ratios, CI is the confidence interval, (iv) vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

The estimated effect of maternal depression on a child's profile of competencies across the four periods of maternal depression exposure is assumed to be constant across the different child and mother characteristics levels (i.e., health, demographic and socioeconomic status). The effect of maternal depression exposure across the different periods could be more or less pronounced depending on the child's health at birth, child's health through childhood, mother's health prior to birth, and demographic characteristics. I checked for all possible interactions in the model, but I included only statistically significant ones.

The marginal effect of maternal depression exposure is estimated for each time period and each EDI domain. Also, I estimated the average marginal effect for the unadjusted and fully adjusted models. For each timing period, the predicted change in probability of being not ready on the EDI subdomains are estimated for each child based on the child's predicted probability at each level of maternal depression as follows:

$$\hat{P}\ (Y = 1|X) = \ \Lambda\left(X\hat{\beta}\right) = \frac{exp^{(X\hat{\beta})}}{1 + exp^{(X\hat{\beta})}}$$

$$\frac{\Delta\ \Lambda\left(X\hat{\beta}\right)}{\Delta\ X_k} = P(Y = 1|\ X_{k1}) - P(Y = 1|\ X_{k2})$$

The marginal effect for each observation is estimated by calculating the estimated (predicted) probabilities based on the actual values (observed values) of the other variables for each level of the predictor variable (maternal depression) and then the difference between the predicted probabilities across the predictor variable level is averaged across all the observation to get the average marginal effect (AME).

Table 2.9 shows the average marginal effect of maternal depression exposure across the four time periods. In the unadjusted model, children exposed to maternal depression for the first

time during pregnancy and the preschool period are almost 5% and 4% more likely to be vulnerable in the emotional domain at kindergarten age, respectively, compared with those who were not exposed to maternal depression during any time point. Similarly, children exposed to maternal depression during the postnatal period or toddler period were almost 3% more likely to be not ready on the emotional domain than those who were not exposed to maternal depression.

Table 2.9: Average Marginal Effect of Maternal Depression Timing - Vulnerable/Not ready in Emotional Maturity Domain.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Timing of first exposure to maternal depression						
Pregnancy	0.0490***	0.0430***	0.0420***	0.0340***	0.0290***	0.0270***
	(0.0058)	(0.0050)	(0.0050)	(0.0060)	(0.0057)	(0.0050)
Postnatal	0.0317***	0.0261***	0.0260***	0.0200***	0.0125***	0.0113***
	(0.0057)	(0.0057)	(0.0052)	(0.0054)	(0.0056)	(0.0057)
Toddler	0.0278***	0.0235***	0.0232***	0.0180***	0.0098***	0.0098***
	(0.0045)	(0.0045)	(0.0045)	(0.0044)	(0.0045)	(0.0040)
Preschool	0.0424***	0.0378***	0.0377***	0.0318***	0.0206***	0.0199***
	(0.0039)	(0.0039)	(0.0040)	(0.0039)	(0.0039)	(0.0039)
Child health-childhood	NO	YES	YES	YES	YES	YES
Child health-at birth	NO	NO	YES	YES	YES	YES
Mother health prior to birth	NO	NO	NO	YES	YES	YES
Demographic and SES	NO	NO	NO	NO	YES	YES
Interactions	NO	NO	NO	NO	NO	YES

Notes: each model is adjusted for some variables that cover (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic variables; and (v) adjustment for interactions. the Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

After adjusting for child health and family characteristics (models 2 to 6), the marginal effect of maternal depression on children's vulnerability in the emotional domain remained stronger among children exposed to maternal depression during pregnancy and preschool time points. For example, children exposed to maternal depression during pregnancy are 3% to 4% more likely to be not ready in emotional domain, and children exposed to depression during

preschool are 2% to 4% more likely to be not ready compared to children who were not exposed to depression at any time point. Similarly, children exposed to depression during the postnatal and toddler periods are 1% to 3% more likely to be not ready in emotional maturity skills at kindergarten age.

The unadjusted model 1 in table 2.10 shows that the strongest effect of maternal depression exposure on children's physical health was during pregnancy and the preschool period. Children exposed to depression during pregnancy or the preschool period were 4.3 and 5.6 percent more likely to be not ready in the physical health domain compared with those who were not exposed to maternal depression during any time point. On the other hand, children exposed to depression during the postnatal or toddler periods were 3.2 and 3.5 percent more likely to be not ready in the physical health domain. After adjusting for a child's health through childhood health at birth and the mother's health prior to birth, the marginal effect of maternal depression on child physical health was attenuated. However, children exposed to depression during pregnancy or the preschool period still had the highest probability of being not ready in the physical health domain. After further adjusting for the child's and family's characteristics, the marginal effect of maternal depression was attenuated across all the first exposure periods.

Table 2.10: Average Marginal Effect of Maternal Depression Timing – Vulnerable/Not ready in Physical Health Domain

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Timing of first exposure to maternal depression						
Pregnancy	0.0431***	0.0379***	0.0359***	0.0243***	0.0157**	0.0136**
	(0.00057)	(0.0088)	(0.0089)	(0.0065)	(0.0081)	(0.0061)
Postnatal	0.0325***	0.02778***	0.0278***	0.0197***	0.00766	0.01369**
	(0.00043)	(0.0064)	(0.0069)	(0.0053)	(0.0039)	(0.0061)
Toddler	0.0353***	0.0316***	0.0310***	0.0240***	0.0098**	0.0093**
	(0.0047)	(0.0073)	(0.0077)	(0.0065)	(0.0050)	(0.0045)
Preschool	0.0559***	0.0517***	0.0516***	0.0437***	0.0222***	0.009**
	(0.0073)	(0.0012)	(0.0128)	(0.0118)	(0.0115)	(0.0045)
Child health-childhood	NO	YES	YES	YES	YES	YES
Child health-at birth	NO	NO	YES	YES	YES	YES
Mother health prior birth	NO	NO	NO	YES	YES	YES
Demographic and SES	NO	NO	NO	NO	YES	YES
Interactions	NO	NO	NO	NO	NO	YES

Notes: each model is adjusted for some variables that include (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic variables; and (v) adjustment for interactions. the Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively

Table 2.11 shows the marginal effect of maternal depression exposure on the developmental vulnerability in the child's social competence across the four periods. The unadjusted model (1) shows that children exposed to maternal depression during pregnancy are 4.7% more likely to be vulnerable in social competence than those who were not exposed to maternal depression during any period. Similarly, the likelihood of being vulnerable in social competence is 3.9% higher for children exposed to maternal depression during the postnatal period, 2.9% higher for children exposed to depression during the toddler period, and 4.9% higher for children exposed to depression during preschool. Pregnancy and the preschool period have the strongest effect on social competence vulnerability. After adjusting for child health at birth and through the early childhood period, marginal effects of maternal depression exposure

are attenuated across the four periods. Controlling for a mother's health before childbirth was associated with a larger decrease in the value of the marginal effects of maternal depression. Although the negative relationship between maternal depression and child's social competence remained statistically significant after further adjustment for child and family socioeconomic and sociodemographic variables, there was a substantial decrease in the magnitude of the marginal effect of maternal depression exposure across all four periods. After adjustment for interactions among covariates in the model, only the association between maternal depression exposure and child's social competence during pregnancy and the postnatal period remained statistically significant.

Table 2.11: Average Marginal Effect of Maternal Depression Timing – Vulnerable/Not ready in Social Competence Domain.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Timing of first exposure to maternal depression						
Pregnancy	0.0474***	0.0423***	0.0412***	0.0310***	0.0220***	0.0220***
	(0.0066)	(0.0098)	(0.0099)	(0.0081)	(0.0118)	(0.0059)
Postnatal	0.0397***	0.0350***	0.0352***	0.0281***	0.0174***	0.0166***
	(0.0055)	(0.0081)	(0.0085)	(0.0074)	(0.0093)	(0.0056)
Toddler	0.0297***	0.0261***	0.0257***	0.0195***	0.0071	0.0066
	(0.0041)	(0.0060)	(0.0062)	(0.0051)	(0.0038)	(0.0045)
Preschool	0.0492***	0.0452***	0.0451***	0.0381***	0.0206***	0.0066
	(0.0069)	(0.0105)	(0.0109)	(0.0100)	(0.0111)	(0.0045)
Child health-childhood	NO	YES	YES	YES	YES	YES
Child health-at birth	NO	NO	YES	YES	YES	YES
Mother health prior birth	NO	NO	NO	YES	YES	YES
Demographic and SES	NO	NO	NO	NO	YES	YES
Interactions	NO	NO	NO	NO	NO	YES

Notes: each model is adjusted for some variables that include (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic characteristics; and (v) adjustment for interactions. the Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.12 reports the marginal effects of maternal depression exposure on the developmental vulnerability in the child's cognitive and language domain. The unadjusted model indicates that exposure to maternal depression was associated with a negative impact on the child's cognitive and language domain. Compared to children who were not exposed to maternal depression during any period, the likelihood to be not ready in the cognitive domain was 2.7% higher for children who were exposed to depression during pregnancy; 3.2% higher for children exposed to depression in the postnatal period; 2.7% higher for children exposed to maternal depression in the toddler period; and 3.2% higher for children exposed to depression in the preschool period. After adjusting for the child's health at birth and during early childhood, the marginal effect of depression is attenuated across the four periods. Controlling for mother's health prior to birth, resulting in a substantial decrease in the magnitude of the marginal effect of the maternal depression exposure on the cognitive and language domain across all periods, especially during pregnancy, which reflects the importance of maternal health around pregnancy as a potential risk factor. After controlling for child and family socioeconomic and demographic characteristics, the effect of maternal depression on the cognitive and language domain disappeared, which confirms the importance of the socioeconomic status on a child's cognitive skills, over and above the maternal depression exposure effect.

Table 2.12: Average Marginal Effect of Maternal Depression Timing – Vulnerable/Not ready in Cognitive and Language Domain

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Timing of first exposure to maternal depression						
Pregnancy	0.0274***	0.0221***	0.0206***	0.0094	0.0032	0.0021
	(0.0026)	(0.0052)	(0.0051)	(0.0025)	(0.0018)	(0.0061)
Postnatal	0.0322***	0.0274***	0.0274***	0.0196***	0.0094	0.0080
	(0.0030)	(0.0064)	(0.0068)	(0.0053)	(0.0053)	(0.0056)
Toddler	0.0271***	0.0233***	0.0228***	0.0160***	0.0043	0.0080
	(0.0025)	(0.0055)	(0.0057)	(0.0043)	(0.0024)	(0.0056)
Preschool	0.0325***	0.0281***	0.0280***	0.0204***	0.0006	0.0002
	(0.0030)	(0.0066)	(0.0069)	(0.0055)	(0.0003)	(0.0040)
Child health-childhood	NO	YES	YES	YES	YES	YES
Child health-at birth	NO	NO	YES	YES	YES	YES
Mother health prior to birth	NO	NO	NO	YES	YES	YES
Demographic and SES	NO	NO	NO	NO	YES	YES
Interactions	NO	NO	NO	NO	NO	YES

Notes: each model is adjusted for some variables that include (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic characteristics; and (v) adjustment for interactions. The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.13 shows the effect of maternal depression exposure on the developmental vulnerability in the child's communication and knowledge domain. Children exposed to maternal depression during preschool are 1.7% more likely to be not ready on the communication domain compared with those who were not exposed to depression at any time. After controlling for mother health prior to birth, the marginal effect of maternal depression exposure during preschool decreased to be 1.08%. This significant relation disappeared after controlling for the family's socioeconomic characteristics. That is, family socioeconomic characteristics have a strong effect on communication skills.

Table 2.13: Average Marginal Effect of Maternal Depression Timing – Vulnerable/Not ready in Communication Domain.

	Model (1)	Model (2)	Model (3)	Model (4)	Model (5)	Model (6)
Timing of first exposure to maternal depression						
Pregnancy	0.0062	0.0023	0.0011	0.0040	-0.006862	-0.0088
	(0.0003)	(0.0004)	(0.0002)	0.0010	(0.0036)	(0.0061)
Postnatal	0.0015	0.0019	-0.0019	-0.0053	-0.01125*	-0.0127
	(0.0001)	(0.0004)	(0.0004)	0.0013	(0.0059)	(0.0058)
Toddler	0.0049	0.0021	0.0017	-0.0013	-0.008559*	-0.0127
	(0.0002)	(0.0005)	(0.0004)	0.0003	(0.0045)	(0.0058)
Preschool	0.0175***	0.0142***	0.0141***	0.0108***	-0.003092	-0.0040
	(0.0008)	(0.0030)	0.0032	0.0026	(0.0016)	(0.0039)
Child health-childhood	NO	YES	YES	YES	YES	YES
Child health-at birth	NO	NO	YES	YES	YES	YES
Mother health prior birth	NO	NO	NO	YES	YES	YES
Demographic and SES	NO	NO	NO	NO	YES	YES
Interactions	NO	NO	NO	NO	NO	YES

Notes: each model is adjusted for some variables that cover (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic characteristics; and (v) adjustment for interactions. the Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

To sum, maternal depression exposure has a direct and negative effect on a child's emotional, physical, and social competence domains across all the timing points. This significant effect of maternal depression exposure on cognitive and communication domains disappeared after adjusting for child and family socioeconomic and demographic characteristics. Therefore, the influence of maternal depression exposure on child EDI domains may be confounded by family socioeconomic characteristics and this confounding is partial for the emotional, physical, and social domains. In contrast, the effect of maternal depression is fully confounded for both the cognitive and communication domains where the direct effect of maternal depression on these to domains appeared to be not significant after controlling for family socioeconomic characteristics.

Maternal depression during pregnancy is sensitive to the child's school readiness because of the direct and indirect effect of maternal depression during this period compared with the

other timing points. Exposure to maternal depression during pregnancy could be associated with poor health outcomes at birth. Consequently, the effect of maternal depression on a child's EDI domains could be partially confounded by a child's health outcomes at birth, as well as by family socioeconomic characteristics for children exposed to maternal depression during pregnancy. Exposure to maternal depression during pregnancy and the preschool period are the most sensitive periods of the child's development in three developmental areas: emotional, physical, and social skills. However, the preschool period has the strongest effect on the physical health and well-being domain.

To investigate if there is a difference in the developmental vulnerability among the maternal depression exposure periods, I compared the odds of being vulnerable in EDI domains for each pair of children group who were exposed to maternal depression in the four timing periods. Specifically, the log odds of being vulnerable in the EDI domain is modeled for each maternal depression time group and the difference in the log odds, or equivalently the log odds ratio, for each pair of maternal depression timing is reported as

$$log\,(odds_i) - log\,(odds_j) = log\left(\frac{odds_i}{odds_j}\right) = log(\,OR_{ij}\,).$$

Table 2.14 presents the fully-adjusted difference in log odds and ORs of being developmentally vulnerable in the EDI for each pair of maternal depression timing. Results for the emotional domain are shown in the first column of Table 2.14. There is a significant difference in the odds of developmental vulnerability when comparing children exposed to maternal depression during pregnancy and the other three periods (postnatal, toddler, and preschool). There is also a significant difference in the log odds of being vulnerable in the emotional domain when comparing toddler and preschool groups. Children exposed to maternal depression during pregnancy had greater odds of being not ready on the emotional domain at

kindergarten age compared with those exposed to maternal depression during the postnatal period (17% higher odds), the toddler period (20% higher odds), and the preschool period (8% higher odds). That is, exposure to maternal depression during pregnancy has the strongest effect on the child's emotional maturity domain compared with the other periods. In contrast, there is no significant difference in the odds of being not ready on emotional skills for children exposed to maternal depression in the postnatal period compared with those exposed to maternal depression during the toddler period or preschool period. Children exposed to maternal depression during the preschool period are more likely to be not ready on emotional domain compared with children exposed to maternal depression during the toddler period.

The results of the physical health domain are shown in the second column of Table 2.14. Two significant differences across the maternal depression timing points were found. Children exposed to maternal depression during preschool periods are more likely to be vulnerable in the physical domain compared to those exposed to maternal depression during the postnatal or toddler period groups. The difference in the log odds of being vulnerable in the physical domain between postnatal and preschool groups is – 0.1335. Similarly, children exposed to maternal depression during preschool are more likely to be vulnerable in the physical domain compared with those exposed to maternal depression during the toddler period. The difference in the log odds between the toddler period and the preschool period is – 0.1137. In other words, the preschool period is a sensitive period for physical health development. There were two significant differences in the odds of developmental vulnerability in the social competence domain when comparing the different pairs of maternal depression time points. Children exposed to maternal depression during pregnancy had higher odds of being vulnerable compared to those exposed to maternal depression during the toddler period. Similarly, children exposed to

maternal depression during the preschool period had higher odds of being not ready compared with toddler period groups. The difference in log odds of being developmentally vulnerable in social competence domain is 0.146 higher for children exposed to maternal depression during pregnancy compared to those exposed to maternal depression during toddler group, and 0. 1330 lower for children exposed to maternal depression during preschool compared to those exposed to maternal depression during the toddler period.

For both cognitive and communication domains, there was no significant difference in the odds of being vulnerable when comparing each pair of the maternal depression timing points. However, these children were also at risk of poor school readiness irrespective of the timing of maternal depression.

Table 2.14: Fully Adjusted Difference in Log Odds of being vulnerable in EDI Domains.

	Emotional domain		Physical health		Social competence		Lang & Cognitive		Communication	
	Diff in log odds	OR	Diff in log odds	OR	Diff in log odds	OR	Diff in log odds	OR	Diff in log odds	OR
Pregnancy vs Postnatal	0.1551**	1.1678**	0.0738	1.0766	0.0448	1.0450	-0.0613	0.9406	0.0461	1.0472
	(0.0753)	(0.0879)	(0.0745)	(0.0802)	(0.0750)	(0.0784)	(0.0776)	(0.0730)	(0.0850)	(0.0891)
	[4.2406]		[0.9805]		[0.3574]		[0.6229]		[0.2940]	
Pregnancy vs Toddler	0.1820***	1.1997***	0.0540	1.0555	0.1461**	1.1570**	-0.0110	0.9890	0.0178	1.0180
	(0.0672)	(0.0806)	(0.0663)	(0.0700)	(0.0677)	(0.0783)	(0.0699)	(0.0692)	(0.0756)	(0.0770)
	[7.3417]		[0.6643]		[4.6620]		[0.0248]		[0.0554]	
Pregnancy vs Preschool	0.0759**	1.0789**	-0.0597	0.9421	0.0130	1.0130	0.0251	1.0254	-0.0393	0.9614
	(0.0637)	(0.0087)	(0.0629)	(0.0593)	(0.0639)	(0.0648)	(0.0672)	(0.0690)	(0.072)	(0.0693)
	[1.4240]		[0.8996]		[0.0416]		[0.1396]		[0.2979]	
Postnatal vs Toddler	0.027	1.0273	-0.0197	0.9804	0.1013	1.1066	0.0502	1.0515	-0.0283	0.9721
	(0.0665)	(0.0683)	(0.0645)	(0.0632)	(0.0656)	(0.0727)	(0.0662)	(0.0696)	(0.0730)	(0.0709)
	[1.1642]		[0.0938]		[2.3819]		[0.5764]		[0.1505]	
Postnatal vs Preschool	-0.0791	0.9239	-0.1335**	0.8750**	-0.0318	0.9687	0.0864	1.0902	-0.0854	0.9181
	(0.0630)	(0.0582)	(0.0611)	(0.0534)	(0.0618)	(0.0599)	(0.0634)	(0.0691)	(0.0693)	(0.0636)
	[1.5770]		[4.7760]		[0.2641]		[1.8570]		[1.5187]	
Toddler vs Preschool	-0.1061**	0.8990**	-0.1137**	0.8925**	-0.1330**	0.8754**	0.0361	1.0368	0.0571	1.0588
	(0.0529)	(0.0476)	(0.0506)	(0.0452)	(0.0526)	(0.0460)	(0.0536)	(0.0556)	(0.0573)	(0.0607)
	[4.0206]		[5.0467]		[6.3965]		[0.4549]		[0.9938]	

Notes: (i) the null hypothesis test is H_0: $log(odds_i) - log\ (odds_j) = 0$, (ii) the Wald Chi-Square and P-value is used in the test (iii) exponentiating the difference in log odds gives an estimate for the odds ratios. (iv) numbers in parenthesis are the std error. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

2.7. Heterogeneity in Effect of the Timing of Maternal Depression

I explored if there is a heterogeneous effect of maternal depression on children's vulnerability by using child and family characteristics that have been identified in the literature of maternal depression such as child gender, mother's education, neighborhood-area SES ((Goyal et al., 2010; Augustine and Crosnoe, 2010; Quarini et al., 2016; Fairthorne et al., 2018; Wszołek et al., 2018).

Previous research postulated that boys' developmental skills, especially cognitive functioning, are more sensitive to maternal depression than girls' (Sharp et al., 1995; Hay et al., 2001; Kurstjens and Wolke, 2001). Females tend to have better EDI scores than males and children who are younger at the time of EDI assessment had lower EDI scores (Brownell et al., 2012; Janus and Offord, 2007). Literature from brain development indicates that the amygdala[5] reaches its full growth in girls approximately 1.5 years before boys (Uematsu et al., 2012) and elevated maternal cortisol concentration at 15-week gestation is associated with larger right amygdala volumes among girls, but there was no association with right or left hippocampus[6] volume. In contrast, maternal cortisol concentration at 15 weeks was not associated with left or right amygdala volume in boys, but the higher maternal cortisol levels at 15 weeks were associated with smaller left and right hippocampal volume (Buss et al., 2012). It is thus reasonable to postulate that boys and girls may be affected differently by maternal depression exposure.

[5] Amygdala is one of two almond-shaped clusters of nuclei, a cluster of neurons in the central nervous system, located deeply in the brain's medial temporal lobe. It plays a key role in the processing of memory, emotional response (anxiety, sadness, autism, depression, aggression, post-traumatic stress disorder, and phobias), and decision making.

[6] There are two hippocampi (right and left) and they are located in the medial temporal lobe of the brain. Hippocampus encode verbal and visual-spatial memories. They have a role in the cognition domain for children of low birth weight and preterm (Isaacs et al., 2000).

In all categories of maternal depression timing, the effect of depression seemed greater for boys than girls on emotional skills. What can be seen is that the average marginal effect of maternal depression on emotional maturity skills is larger for boys compared to girls. Boys exposed to maternal depression are 1.6% - 3.3% more likely to be vulnerable in emotional skills compared to those who were not exposed to maternal depression. The strongest effect of maternal depression among boys was in pregnancy (3.3%), followed by the preschool period (2.7%) then the postnatal period (1.6%) and toddler period (1.5%).

Table 2.15: Gender Heterogeneity: Marginal effect of Maternal Depression on Children's School Readiness – Vulnerability in EDI Domains

	(1) Emotional	(2) Physical	(3) Social	(4) Cognitive	(5) Communication
Panel (A): Boys					
Pregnancy	0.0330***	0.0230***	0.0324***	0.0115	0.0067
	(0.0095)	(0.0090)	(0.0117)	(0.0054)	(0.0020)
Postnatal	0.0164***	0.0147	0.0235***	0.0182	0.0175
	(0.0047)	(0.0060)	(0.0085)	(0.0084)	(0.0074)
Toddler	0.0157***	0.0159**	0.0086	0.0043	0.0211
	(0.0045)	(0.0065)	(0.0031)	(0.0020)	(0.0091)
Preschool	0.0272***	0.0274***	0.0238***	0.0029	0.0035
	(0.0078)	(0.0114)	(0.0086)	(0.0014)	(0.0015)
Panel (B): Girls					
Pregnancy	0.0232***	0.0080	0.01117	0.0060	0.0070
	(0.0121)	(0.0047)	(0.0065)	(0.0038)	(0.0040)
Postnatal	0.0085	0.0012	0.0012*	0.0002	0.0050
	(0.0044)	(0.0007)	(0.0065)	(0.0001)	(0.0020)
Toddler	0.0036	0.0040	0.0053	0.0038	0.0034
	(0.0018)	(0.0023)	(0.0031)	(0.0024)	(0.0019)
Preschool	0.0137***	0.0174***	0.01742***	0.0014	0.0098
	(0.0071)	(0.0108)	(0.00102)	(0.0009)	(0.0056)

Notes: (i) Panel A shows the effect of maternal depression on boys; (ii) Panel B shows the effect of maternal depression on girls; (iii) All regressions control for child health, mother health, social, economic, and demographic characteristics; (iv) Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

In contrast, girls exposed to maternal depression were 0.3% - 2.3% more likely to be vulnerable in the emotional domain. The strongest effect of depression among girls was in pregnancy (2.3%), followed by the preschool period (1.4%). Maternal depression appears to have

no effect during the postnatal and toddler periods. Regarding the physical health domain, the effect of depression appears to be greater for boys than girls. Boys are 1.4% - 2.7% more likely to be vulnerable in the physical domain compared to those who were not exposed to depression. However, maternal depression did not have a significant effect on boys during the postnatal period. In contrast, maternal depression has a significant effect on the physical domain among girls only during the preschool period. For the social competence domain, maternal depression seemed to have a greater effect on boys than girls. Boys are 0.8% - 3.2% more likely to be vulnerable in the social domain than those who were not exposed to maternal depression. In contrast, girls are 0.5% - 1.7% more likely to be vulnerable in the social domain than girls whose mother did not experience maternal depression. However, exposure to maternal depression among girls appears to have no effect during the pregnancy and toddler periods. On the other hand, although the marginal effect of maternal depression on the cognitive and communication domains are larger for boys than girls, they are not significant. Consistent with the results for the full sample, I find no evidence that maternal depression impacts cognitive and communication ability for boys or girls across different time periods. That is, results in columns 4 and 5 suggest no heterogeneous effect by gender.

I explored whether the effects of exposure to maternal depression during different periods vary depending on maternal characteristics by investigating the relationship separately for families in which the mother has a partner, as well as by maternal schooling levels. I split the sample by married and unmarried mothers to check the role of marital status and then by not having a high school degree, and having a high school degree or higher to check the role of mother's education. These factors may modulate the impact of depression (Shonkoff and Phillips 2000). Table 16 shows that, among children of unmarried mothers, there is a statistically

significant negative relationship between maternal depression exposure during pregnancy and children's emotional, physical and social domains. The relationship was somewhat stronger for emotional and social domains. Children of non-married (no partner) mothers who were exposed to depression during pregnancy were almost 3.8% more likely to be vulnerable in the emotional domain, 3.4% more likely to be vulnerable in the physical health domain, and 3.4% more likely to be vulnerable in social domain compared to children with non-marred and non-depressed mothers.

Table 2.16: Mother's Marital Status Heterogeneity: Marginal effect of Maternal Depression on Children's School Readiness – Vulnerability in EDI domains

	(1) Emotional	(2) Physical	(3) Social	(4) Cognitive	(5) Communication
Panel (A): Children of Married Mothers					
Pregnancy	0.0114	0.0157*	0.0016	0.00098	0.0239*
	(0.0063)	(0.0090)	(0.0009)	(0.0006)	(0.0126)
Postnatal	0.0153*	0.0145*	0.0147*	0.01855**	0.0067
	(0.0084)	(0.0084)	(0.0084)	(0.0114)	(0.0035)
Toddler	0.01335**	0.0103	0.01213**	0.0085	0.0065
	(0.0073)	(0.0069)	(0.0069)	(0.0052)	(0.0034)
Preschool	0.01729***	0.0137**	0.0183***	0.0022	0.00416
	(0.00095)	(0.0079)	(0.0105)	(0.0013)	(0.0022)
Panel (B): Children of Unmarried Mothers					
Pregnancy	0.0382***	0.0158*	0.0344***	0.0040	0.0001
	(0.0179)	(0.0064)	(0.0155)	(0.0020)	(0.00005)
Postnatal	0.01129	0.0036	0.0190**	0.0042	0.0150
	(0.0053)	(0.0014)	(0.0086)	(0.0020)	(0.0067)
Toddler	0.0079	0.0098	0.0042	0.0024	0.0107*
	(0.0037)	(0.0040)	(0.0019)	(0.0011)	(0.0048)
Preschool	0.0238***	0.0286***	0.0238***	0.0030	0.0024
	(0.0112)	(0.0116)	(0.0107)	(0.0014)	(0.0011)

Notes: (i) Panel A shows the effect of maternal depression on children of married mothers; (ii) Panel B shows the effect of maternal depression on children of unmarried mothers; (iii) All regressions control for child health, mother health, social, economic, and demographic characteristics; (iv) Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

On the other hand, among children of married mothers, there is a statistically significant relationship between maternal depression exposure during pregnancy and children's physical health and communication domains only. Children exposed to depression during pregnancy, among married mothers, were almost 1.5% and 2% more likely to be not ready in physical and communication domains, respectively, compared with those whose mother was married and not depressed. When we compare children of married mothers and unmarried mothers during the pregnancy period, the results indicate that children of non-married mothers are more sensitive to depression exposure than children of married mothers in the emotional, physical and social domains. Moreover, there is no difference in the effect of maternal depression exposure during pregnancy on the cognitive domain across the level of mother marital status. Similarly, children of married and unmarried mothers who were exposed to depression during the preschool period are more likely to be vulnerable in emotional, physical, and social domains compared with children of non-depressed mothers. However, the effect of depression is stronger among children whose mothers were unmarried. For example, compared with children whose mothers are unmarried and not depressed, children of unmarried mothers and exposed to depression during the preschool period are 2.3% more likely to be not ready in the emotional domain (almost 1.7% in the married group), 2.8% more likely to be not ready on the physical domain (almost 1.3% in the married group), and 2.3% more likely to be not ready on the social domain (almost 1.8% in the married group). For cognitive and communication domains, there appears to be no difference in the effect of maternal depression across the level of the mother’s marital status. In sum, children of unmarried mothers are more sensitive to depression exposure during pregnancy and the toddler period. Marital status moderates the effect of maternal depression on emotional,

physical, and social health. However, maternal depression appears to have no different effects across marital status levels on cognitive and communication domains.

The results in Table 2.17 indicate that there are no apparent differences in the effect of maternal depression for mothers who have a high school degree and those who do not. However, there are little sizable differences in children's performance on the emotional domain by mother's schooling levels for children exposed to depression in pregnancy and preschool periods. Among children whose mothers have a high school degree, children exposed to maternal depression during pregnancy and preschool are, respectively, 2.6% and 1.8% more likely to be not ready in the emotional domain. In contrast, among children whose mothers did not have a high school degree, children exposed to maternal depression during pregnancy and preschool are 3.1% and 2.1%, respectively, to be not ready in the emotional domain. In summary, the effect of maternal depression does not depend on the education level of the mother for all time periods and EDI domains, except for the emotional domain during pregnancy and the preschool period.

Table 2.17: Mother Education Heterogeneity: Marginal Effect of Maternal Depression on Children's School Readiness – Vulnerability in EDI Domains.

	(1) Emotional	(2) Physical	(3) Social	(4) Cognitive	(5) Communication
Panel (A): Children of mothers with high school					
Pregnancy	0.0267***	0.0134**	0.0212***	0.0025	-0.0052
	(0.0098)	(0.0067)	(0.0096)	(0.0013)	(0.0026)
Postnatal	0.0098	0.0062	0.0182**	0.0060	0.0106
	(0.0036)	(0.0031)	(0.0082)	(0.0033)	(0.0053)
Toddler	0.0096**	0.0092	0.0056	0.0029	0.0104
	(0.0035)	(0.0046)	(0.0025)	(0.0016)	(0.0052)
Preschool	0.0186***	0.0194*	0.0192***	0.0017	0.0023
	(0.0068)	(0.0097)	(0.0087)	(0.0009)	(0.0011)
Panel (B): Children of mothers without a high school					
Pregnancy	0.0314*	0.0209	0.0199	0.000071	0.0275
	(0.0070)	(0.0047)	(0.0051)	(0.00002)	(0.0073)
Postnatal	0.0143	0.0004	0.0005	0.0126	0.0330*
	(0.0032)	(0.0001)	(0.00013)	(0.0037)	(0.0087)
Toddler	0.0008	0.0020	0.0012	0.0010	0.01145
	(0.0001)	(0.0004)	(0.00032)	(0.00030)	(0.0030)
Preschool	0.0212*	0.0267***	0.0172	0.0045	0.0153
	(0.0047)	(0.0061)	(0.0044)	(0.0013)	(0.0040)

Notes: (i) Panel A shows the effect of maternal depression on children of mothers with high school; (ii) Panel B shows the effect of maternal depression on children of mothers who did not have a high school; (iii) All regressions control for child health, mother health, social, economic, and demographic characteristics; (iv) Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

In this subsection, I analyzed whether EDI outcomes vary for children of mothers with different neighborhood-area socioeconomic status (SES) levels. Mothers were stratified into two groups based on their SEFI2 score. The low neighborhood-area SES group includes mothers with SES scores above zero and high neighborhood-area SES group includes mothers with SES scores below zero.

Table 2.18 shows the fully adjusted odds ratios of the effect of maternal depression on the child's EDI domains among children whose mothers have low neighborhood socioeconomic status. The results show some differences in terms of how each timing period is associated with EDI outcomes. Maternal depression exposure has a negative effect on the emotional and social skills domains across all four periods. The effect of maternal depression on children's physical health was significant only during the preschool period. In contrast, the effect of maternal depression exposure on cognitive and communication domains was not significant across all the timing points. Child health through early childhood has a significant effect on all the EDI domains and most latent variables of child's health at birth have a significant effect on a child's EDI outcomes. Further, consistent with the full sample analysis, child and family characteristics have a significant and strong effect on all EDI domains. The odds of being vulnerable in EDI domains are 77% to 192% greater for boys than girls. Furthermore, among family characteristics, mother's education, teenage mother, family size, and receiving income assistance have the strongest effect on all EDI domains. The odds of being not ready in the EDI domains are 16% to 47% greater for children of teenage mothers than those whose mothers were not teenagers at birth; 18% to 101% greater for children in families with more than four children than those in families with less than four children; and 27% to 51% greater for children in families on income assistance than those in families not on income assistance.

Table 2.18: Fully-adjusted Odds Ratios of Vulnerability in EDI Domains – Low Socioeconomic Status Group.

	Model 1 Emotional	Model 2 Physical	Model 3 Social	Model 4 Cognitive	Model 5 Communication
Timing of first exposure to maternal depression					
Pregnancy (In uterus)	1.278***	1.067	1.184***	0.915	0.901
Postnatal (Birth -12 months)	1.180***	1.006	1.185***	1.125*	0.876*
Toddler I (12 months-36 months)	1.142***	1.065	1.064	1.071	0.906
Preschool (36 months-60 months)	1.200***	1.223***	1.194***	1.023	0.943
Child's health through childhood					
Major illness (2+ major ADGs)	1.196***	1.092**	1.114***	1.125***	1.055
Minor illness (90th+ minor ADGs	1.249***	1.093**	1.137***	1.135***	1.181***
Child hospital admission (>95th)	1.371***	1.807***	1.501***	1.908***	1.897***
Child's health at birth					
Low Birth Weight (< 2500g)	1.391***	1.271***	1.208**	1.271***	1.345***
Preterm (< 37 weeks)	0.958	1.090	1.014	1.084	1.054
5-minutes Apgar (>=8)	1.123	0.787***	1.027	0.946	0.994
Length of hospitalization(>6 days)	1.161*	1.370***	1.281***	1.263***	1.257***
Emergency c-section	1.194*	1.164*	1.172*	1.100	1.155
Mother health 5 years before birth					
Major illness (>2+ ADGs)	1.089***	1.083**	1.104**	1.063*	0.991
Minor illness (>90th minor ADGs)	1.095***	1.001	1.057	1.058	0.950
Mother hospital admission (>95th)	1.121***	1.072	1.093	1.127**	1.114*
Child, Mother, family characteristics					
Child's age (in months)	0.975***	0.961***	0.962***	0.922***	0.949***
Child gender (male=1)	2.929***	1.776***	2.457***	1.968***	2.041***
Breast feeding initiation	0.906***	0.883***	0.828***	0.764***	0.716***
Mother married	0.631***	0.603***	0.600***	0.640***	0.861***
Mother less than HS	1.136***	1.094**	1.030	1.170***	1.241***
Teenage mother	1.193***	1.408***	1.326***	1.472***	1.167***
Smoking during pregnancy	0.968	1.118***	1.062	0.936	0.899**
Urban residence	1.010	1.024	1.071*	0.942	0.990
Family size (>4)	1.181***	1.638***	1.282***	1.667***	2.016***
Family receive income assistance	1.277***	1.485***	1.403***	0.516***	1.459***

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii)vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1%, respectively.

Table 2.19: Fully-adjusted Odds Ratios of Vulnerability in EDI Domains – High Socioeconomic Status Group.

	Model 1 Emotional	Model 2 Physical	Model 3 Social	Model 4 Cognitive	Model 5 Communication
Timing of first exposure to maternal depression					
Pregnancy (In uterus)	1.372***	1.314***	1.315***	1.259**	0.964
Postnatal (Birth -12months)	1.036	1.167*	1.151	1.002	0.868
Toddler I (12months-36months)	1.028	1.138*	1.070	0.976	0.915
Preschool (36months-60months)	1.278***	1.226***	1.287***	0.978	1.013
Child's health through childhood					
Major illness (2+ major ADGs)	1.123***	1.125***	1.154***	1.234***	1.070
Minor illness (90th+ minor ADGs	1.293***	1.228***	1.202***	1.206***	1.124*
Child hospital admission (>95th)	1.555***	2.707***	1.832***	2.512***	2.434***
Child's health at birth					
Low Birth Weight (< 2500g)	1.221	1.367**	1.299**	1.594***	1.472***
Preterm (< 37 week)	0.910	1.078	0.898	0.989	1.016
5-minutes Apgar (>=8)	0.830*	0.748***	0.702***	0.718***	0.794***
Length of hospitalization (>6 days)	1.044	1.207	1.185	1.123	1.252*
Emergency c-section	1.202*	1.112	1.178	1.083	0.942
Mother health 5 years before the birth					
Major illness (>2+ ADGs)	1.091*	1.100**	1.127***	0.960	0.921*
Minor illness (>90th minor ADGs)	1.047	1.062**	1.064	1.098	0.969
Mother hospital admission (>95th)	1.086	1.218	0.997	1.097	1.147
Child, Mother, family characteristics					
Child's age (in months)	0.969***	0.964***	0.964***	0.930***	0.970***
Child gender (male=1)	3.540***	1.984***	2.869***	1.940***	1.831***
Breast feeding initiation	0.873**	0.866***	0.764***	0.725***	0.707***
Mother married	0.765***	0.635***	0.727***	0.812***	0.978
Mother less than HS	1.506***	1.580***	1.431***	1.659***	2.102***
Teenage mother	1.196***	1.296***	1.382***	1.668***	1.190***
Smoking during pregnancy	1.311***	1.323***	1.382***	1.207***	0.935
Urban residence	1.199***	1.126***	1.237***	1.170***	1.039
Family size (>4)	1.223***	1.847***	1.401***	1.845***	2.367***
Family receive income assistance	1.450***	1.731***	1.472***	1.620***	1.828***

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.19 shows the fully adjusted odds ratios of the effect of maternal depression on the child's EDI domains among children with mothers in the high neighborhood socioeconomic status group. Maternal depression exposure during pregnancy and the preschool period has a negative effect on the emotional, physical and social skills domains. The effect of maternal depression exposure during the postnatal and toddler periods has a significant effect on the physical health domain only. In contrast, the effect of maternal depression exposure on cognitive and communication domains was not significant for all timing points. Child health through early childhood has a significant effect on all EDI domains and most latent variables of child's health at birth have a significant effect on a child's EDI outcomes. Further, consistent with the full sample analysis, the child and family characteristics has a significant and strong effect on all EDI domains. The odds of being vulnerable in EDI domains are 83% to 254% greater for boys than girls. Furthermore, among family characteristics, mother's education, teenage mother, family size, and receiving income assistance have the strongest effect on all EDI domains. The odds of being not ready in the EDI domains are 19% to 66% greater for children of teenage mothers than those whose mothers were not teenagers at birth; 22% to 136% greater for children in families with more than four children than those in families with less than four children; and 45% to 82% greater for children in families on income assistance than those in families not on income assistance.

When comparing the effect of maternal depression exposure across the two socioeconomic groups, the results indicate that impact of maternal depression exposure during pregnancy and the preschool period on the emotional and social competence domains for children whose mothers are in the low socioeconomic status group is similar to those whose mothers are in the high socioeconomic status group. In contrast, the effect of maternal

depression exposure during the postnatal and toddler periods are not similar across the two socioeconomic groups. Maternal depression exposure does not have a significant effect on emotional and social domains among children whose mothers have high socioeconomic status, while maternal depression exposure does not have a significant effect on physical domain among children whose mothers are in the low socioeconomic status group. The stratified analyses, by the level of socioeconomic status of the child's family, showed that the child's gender effect on the EDI domains is slightly stronger among children whose mothers have a high level of socioeconomic status than those in the low socioeconomic status group. Generally, the effect of maternal depression exposure on children's school readiness is not substantially moderated by the family's socioeconomic status. In other words, the effect of maternal depression on school readiness does not depend on the level of socioeconomic status.

Table 2.20 reports the average marginal effect of maternal depression exposure on children's EDI domains. Panel (A) reports the marginal effect for children of mothers with high neighborhood-area socioeconomic status and panel (B) reports the average marginal effect of maternal depression for children of mothers with low neighborhood-area socioeconomic status. The two panels show some differences in terms of how each timing period is associated with EDI outcomes. Exposure to maternal depression during pregnancy and the preschool period has the same effect on the children's emotional and social domains across the two socioeconomic status groups. Children exposed to maternal depression during pregnancy and the preschool period are almost 2% more likely to be not ready in emotional and social competence domains across the two groups compared to those who were not exposed to maternal depression. In contrast, the effect of maternal depression exposure during the postnatal and toddler periods are

not similar across the two socioeconomic groups. Maternal depression exposure does not have a significant effect on emotional and social domains among children whose mothers have high socioeconomic status. Children who were exposed to depression during the postnatal and toddler periods are 2% and 1.5%, respectively, more likely to be not ready in the emotional domain among children whose mothers have low SES. Children exposed to maternal depression during the postnatal and toddler periods are almost 1% more likely to be not ready among children whose mothers have high socioeconomic status, while there is no significant effect on physical health among the children whose mothers have low SES. The effect of maternal depression exposure was not significant for both the language and cognitive domain and the communication skills domain, for either mother with low SES or high SES.

Table 2.20: Marginal effect of Maternal Depression Exposure on Vulnerability in EDI Domains by Neighborhood-area Socioeconomic Status

	Emotional	Physical	Social	Cognitive	Communication
Mothers with High Socioeconomic Status SES					
Pregnancy	0.0271***	0.0223***	0.0215***	0.0167**	-0.0027
	(0.0157)	(0.0128)	(0.0128)	(0.0104)	(0.0015)
Postnatal	0.0030	0.0126*	0.0110	0.00011	-0.01025
	(0.0017)	(0.0072)	(0.0066)	(0.0001)	(0.0058)
Toddler	0.0023	0.0105*	0.0530	-0.0017	-0.0064
	(0.0014)	(0.0060)	(0.0031)	(0.0011)	(0.0037)
Preschool	0.0210***	0.0167***	0.0198***	-0.0016	0.0010
	(0.0122)	(0.0095)	(0.0119)	(0.0010)	(0.0006)
Mothers with Low Socioeconomic Status SES					
Pregnancy	0.0284***	0.0087	0.0206**	-0.0113	-0.0119
	(0.0135)	(0.0036)	(0.0093)	(0.0052)	(0.0050)
Postnatal	0.0190**	0.00078	0.0206**	0.0149*	-0.0152
	(0.0090)	(0.0003)	(0.0093)	(0.0069)	(0.0069)
Toddler	0.0153**	0.0085	0.0076	0.0087	-0.0113
	(0.0073)	(0.0035)	(0.0034)	(0.0040)	(0.0051)
Preschool	0.0211***	0.0268***	0.0217***	0.0028	-0.0066
	(0.0101)	(0.0112)	(0.0098)	(0.0013)	(0.0030)

Notes: (i) Panel A shows the effect of maternal depression on children of mothers with high socioeconomic status; (ii) Panel B shows the effect of maternal depression on children of mothers with low socioeconomic status; (iii) All regressions control for child health, mother health, social, economic, and demographic characteristics; (iv) Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

2.8. Robustness check

2.8.1 Standard Approach for Identifying Vulnerability

According to the Offord Centre, the standard approach to identify children who fall in the "Vulnerable" or "Not ready" for school category describes the children who score in the bottom 10th percentile cut-off in at least one domain (Janus and Offord, 2007). In this section, I designed vulnerability on the EDI to include children who are not ready in one or more domains, not in just one domain.

Table 2.21 reports the odds ratio of being not ready in at least one EDI domain. The unadjusted model shows a significant association between maternal depression exposure and developmental vulnerability in at least one EDI domain. The odds of being not ready in at least one EDI domain is 38% greater for children who were exposed to maternal depression during pregnancy; 25% greater for children who were exposed to maternal depression during the postnatal period; 29% greater for children exposed to maternal depression during the toddler period; and 45% greater for children exposed to maternal depression during the preschool period compared with children who were not exposed to maternal depression during any time period. After adjusting for child health at birth and throughout childhood, the odds of being not ready in at least one EDI domain are attenuated; however, the negative effect of maternal depression on vulnerability in at least one EDI domain remained significant and strong. Further adjusting for child and family socioeconomic and demographic characteristics were associated with a substantial reduction in the magnitude of the odds ratio of being not ready. According to the fully adjusted model, the odds of being not ready in at least one EDI domain are 14% greater for children who were exposed to maternal depression during pregnancy; 3% greater for children who were exposed to maternal depression during the postnatal period; 8% greater for children

who were exposed to maternal depression during the toddler period; and 15% greater for children who were exposed to maternal depression during the preschool period.

Table 2.21: Unadjusted and Fully adjusted Odds Ratio of Being Developmentally Vulnerable in at Least one EDI Domain

	Model 1 unadjusted	Model 2 Partially adj	Model 3 Partially adj	Model 4 Fully adj
Timing of first exposure to maternal depression				
Pregnancy (In uterus)	1.385***	1.326***	1.210***	1.145***
Postnatal (Birth -12m)	1.251***	1.213***	1.137***	1.033
Toddler I (12m-36m)	1.299***	1.266***	1.198***	1.080**
Preschool (36m-60m)	1.456***	1.420***	1.337***	1.152***
Child's health through childhood				
Major illness (2+ major ADGs)		1.230***	1.209***	1.089***
Minor illness (90th+ minor ADGs)		1.140***	1.085***	1.131***
Child hospital admission (>95th)		2.326***	2.284***	1.883***
Child's health at birth				
Low Birth Weight (< 2500g)		1.226***	1.232***	1.328***
Preterm (< 37 weeks)		1.108***	1.085*	0.993
5-minutes Apgar (>=8)		0.918*	0.916*	0.856
Length of hospital stay(>6)		1.204***	1.196*	1.152***
Emergency c-section		0.955	0.959	1.086
Mother health 5 years before the birth				
Major illness (>2+ ADGs)			1.161***	1.051***
Minor illness (>90th minor ADGs)			1.185***	1.062***
Mother hospital admission (>95th)			1.334***	1.174***
Child, Mother, and family chrematistics				
Child's age (in months)				0.943***
Child gender (male=1)				2.275***
Breastfeeding initiation				0.778***
Mother married				0.732***
Mother has less than HS				1.320***
Teenage mother				1.470***
Smoking during pregnancy				1.112***
Neighbourhood-area SES				0.766***
Urban residence				1.072***
Family size (>4)				1.714***
The family receive income assistance				1.485***

Notes: (i) reference group is developmentally 'ready', (ii) EDI is the Early Development Instrument, (iii) OR is the odds ratios, CI is the confidence interval, (iv) vulnerability is the scoring in the bottom 10th percentile of the EDI domains. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

These results are consistent with the analysis for each domain separately. Child family socioeconomic characteristics have a strong effect on child vulnerability in at least one EDI domain, over and above the disadvantage conferred by maternal depression exposure.

Child health at birth and through early childhood has a significant effect on child vulnerability in at least one EDI domain. The number of hospital admission (accumulated from birth to the child's 5th birthday) has a strong negative effect on the child's school readiness. Children with an accumulated number of hospital admission greater than the 95th percentile of the distribution of the number of hospital admission in the population are 1.88 times more likely to be not ready on at least one EDI domain compared with those who had a number of hospital admission less than the 95th percentile cut-off. Major and minor illness may be a pathway from poor health at birth to vulnerability in EDI domains. Low birth weight is an important risk factor for being not ready for school at kindergarten age. Children whose weight at birth is less than 2500 grams are 1.32 times more likely to be not ready in at least one EDI domain compared to children whose weight above 2500 grams.

For child individual-level variables, boys are more likely to be developmentally vulnerable in at least one EDI domain than girls and young children at the time of EDI assessment are more vulnerable in at least one EDI domain than older ones. The odds of being not ready in at least one EDI domain is 127% greater for boys than girls. Breastfeeding initiation was significantly associated with children's school readiness. Children whose mothers-initiated breastfeeding at birth are less likely to be vulnerable in at least one EDI domain. Similarly, children whose mothers were married or in common-law relationships at birth are less likely to be vulnerable than children whose mothers were single. In contrast, the odds of being not ready in at least one EDI domain is 32% greater for children whose mother did not have a high school

at their birth than children whose mothers had a high school degree, and 47% greater for children whose mothers were teenagers at birth than those whose mother not teenagers. Similarly, children in families with more than four children are more likely to be vulnerable than children in families with less than four children and children in families on income assistance are more vulnerable than those in families not on income assistance. Consequently, the child's family environment still has the strongest effect on the child's school readiness, over and above the disadvantage conferred by maternal depression exposure. These findings are consistent with the individual EDI domain analysis, which confirmed the negative effect of maternal depression exposure on children's school readiness across the four periods.

2.8.2 Sensitive periods of Developmental Vulnerability

The results of the analysis showed that exposure to maternal depression during pregnancy and the preschool period has the strongest effect on the child's developmental vulnerability in emotional, physical, and social domains compared with the postnatal and toddler periods. I compared children who were exposed to maternal depression during pregnancy or the preschool period to children who were exposed to depression during the postnatal or toddler periods. More specifically, I pooled children exposed to depression during pregnancy and preschool as one group and I pooled children exposed to depression in postnatal and toddler periods as another group. The difference in log odds of the two groups and the linear hypothesis test results of this difference are reported in table 2.22.

Table 2.22: Fully adjusted Difference in Log Odds of Being Vulnerable in EDI Domains - Pregnancy & Preschool vs. Postnatal & Toddler.

EDI Domains	Difference in log odds	Odds Ratio OR	95% OR confidence limits	Wald Chi-Square
Emotional	0.2612***	1.298***	1.0841 - 1.555	8.0455
	(0.0921)	(0.1196)		(0.0046)
Physical	0.1875**	1.206**	1.011 - 1.439	4.330
	(0.0901)	(0.109)		(0.0374)
Social	0.1778	1.194	0.998 - 1.429	3.767
	(0.0916)*	(0.0195)*		(0.0522)
Cognitive	-0.0974	0.907	0.7540 - 1.0915	1.0654
	(0.094)	(0.085)		(0.302)
Communication	0.1032	1.108	0.9068 - 1.3557	1.0126
	(0.1026)	(0.1137)		(0.3143)

Notes: (i) the null hypothesis test is $H_0: log(odds_i) - log\ (odds_j) = 0$, (ii) the Wald Chi-Square and P-value is used in the test (iii) exponentiating the difference in log-odds gives estimates for the odds ratios. (iv) numbers in parenthesis are the std error. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

I found that children in the group exposed to depression for the first time during pregnancy or preschool are more likely to be not ready on the emotional domain compared with children who were exposed to depression in the postnatal or toddler periods. The difference in the log odds between the two groups is 0.2612. That is, the odds of being not ready in the emotional maturity domain are 29% greater for children who were exposed to depression in pregnancy or the preschool period than those who were exposed to depression during the postnatal or toddler periods. The difference between these two groups is statistically significant as shown by the Wald chi-square statistic of the linear hypothesis test. The difference in the log odds of being vulnerable in the physical domain between both groups is 0.1875. Children in the group exposed to maternal depression during pregnancy or preschool are more likely to be vulnerable in the physical domain compared with those exposed to maternal depression during the postnatal or toddler periods. The odds of being vulnerable in the physical health domain is 20% greater for children in the pregnancy and preschool period exposed group than those in the postnatal and toddler period exposed group. The difference in log odds of being developmentally

vulnerable in the social competence domain is 0.0916 higher for children exposed to maternal depression during pregnancy or preschool compared to those exposed to maternal depression during the postnatal or toddler periods. That is the odds of being vulnerable in the social competence domain for children in the group exposed during pregnancy or preschool are 19% greater than those exposed in the postnatal or toddler periods. For both cognitive and communication domains, there was no significant difference in the odds of being vulnerable when comparing both groups. This is consistent with the main results that indicate that exposure to maternal depression has the strongest effect during pregnancy, followed by the preschool period for the emotional, physical, and social domains. For the cognitive and communications domains, exposure to depression puts children at risk of poor school readiness irrespective of the timing of maternal depression. However, this impact is not significant after adjusting for the child's health and family socioeconomic and demographic variables.

2.8.3 Analysis of Healthy Children

The literature of early childhood development showed that early chronic illness may have a profound influence on a child's developmental trajectories. Some studies found that children and adolescents with chronic illness have lower academic outcomes compared with healthy ones (Case, Fertig, and Paxson, 2005; Jakson 2009). Children who have poor health before starting school are at greater risk of poorer cognitive and psychosocial outcomes (Quach and Barnett, 2015; Goldfeld et al. 2012; Janus and Duku, 2007). This suggests that ill health in early childhood may influence the relationship between maternal depression and child's developmental vulnerability in the EDI domains. Therefore, I excluded children with poor health status through childhood, relative to the majority of the population in the study. I repeated the analysis on the EDI domains after excluding children with major illnesses that exceeded the 90th

percentile of the major ADGs distribution, children with minor illnesses that exceeded the 90th percentile of the minor ADGs distribution, and children with hospital admissions that exceeded the 95th percentile of the number of hospital admission. I repeated each analysis separately as differential health status criteria.

Table 2.23 indicates the average marginal effect of maternal depression exposure after excluding children with a number of major illnesses greater than the 90th percentile of the major ADGs through early childhood, from birth through the child's 5th birthday.

Table 2.23: Average Marginal Effect - Vulnerability in EDI Domains for Healthy Children (Major ADGs Criterion).

	Pregnancy Period	Postnatal Period	Toddler Period	Preschool Period
Emotional	0.0256***	0.0095	0.0180***	0.0181***
	0.0152	0.0056	0.0107	0.0107
Physical	0.02593***	0.0236***	0.0188***	0.0253***
	(0.0017)	(0.0015)	(0.0012)	(0.0016)
Social	0.0342***	0.0108	0.0149**	0.0225***
	(0.0033)	(0.0010)	(0.0014)	(0.0022)
Cognitive	0.0119	0.0039	0.0110*	-0.0024
	(0.0077)	(0.0025)	(0.0071)	(0.0015)
Communication	0.0030	-0.0070	-0.0024	-0.0060
	(0.0017)	(0.0041)	(0.0014)	(0.0035)

Notes: each model is adjusted for some variables that include (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic variables. The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

I limited the analysis to children who identified as healthy during early childhood based on the number of major illnesses (major ADGs). The results indicate that maternal depression exposure among this group has a negative effect on at least three EDI domains across the four timing points of depression. Exposure during pregnancy has the strongest effect on the emotional domain, followed by the preschool and toddler periods. Children exposed to maternal depression during pregnancy are 2.6% more likely to be not ready in the emotional maturity domain

compared with children who were not exposed to depression. This effect is 0.9% for the postnatal period, and 1.8% for both the toddler and preschool periods. Physical health is negatively affected by exposure to maternal depression across the four periods. There was a small difference between the postnatal period and the other periods. For the social competence domain, maternal depression exposure during the pregnancy has the strongest effect, followed by exposure during the preschool period. On the other hand, maternal depression does not have any significant effect on the language and cognitive domain and the communication domain. These results are in line with the main results of the full sample.

Table 2.24 indicates the average marginal effect of maternal depression exposure after excluding children with a number of minor illnesses greater than the 90th percentile of minor ADGs through early childhood, from birth through the child's 5th birthday.

Table 2.24: Average Marginal Effect – Vulnerability in EDI Domains for Healthy Children (Minor ADGs Criterion).

	Pregnancy Period	Postnatal Period	Toddler Period	Preschool Period
Emotional	0.0270***	0.0117*	0.0108**	0.0227***
	(0.0147)	(0.0064)	(0.0059)	(0.0124)
Physical	0.0164**	0.0046	0.0153***	0.0279***
	(0.0088)	(0.0025)	(0.0083)	(0.0151)
Social	0.0212***	0.0124**	0.0094**	0.0247***
	(0.0118)	(0.0069)	(0.0052)	(0.0138)
Cognitive	0.0019	0.0099	0.0083*	0.0029
	(0.0011)	(0.0058)	(0.0049)	(0.0017)
Communication	-0.0128	-0.0161	-0.0057	-0.0017
	(0.0070)	(0.0088)	(0.0031)	(0.0009)

Notes: each model is adjusted for some variables that include (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic variables. The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Among children who are healthy based on the number of minor illnesses (minor ADGs) solely, maternal depression exposure was negatively associated with three areas of child development. Exposure during pregnancy remained the period with the strongest effect on the emotional, physical, and social competence domains, followed by the preschool period. The findings indicate that maternal depression timing appears to have no significant effect on the cognitive and communication domains. The emotional and social domains are most affected and the preschool period has the strongest effect on the child's physical health domain.

Table 2.25 indicates the average marginal effect of maternal depression exposure after excluding children with the number of hospital admissions greater than 95th percentile of the hospital admission distribution through early childhood, from birth through the child's 5th birthday.

Table 2.25: Average Marginal Effect – Vulnerability in EDI Domains for Healthy Children (hospital admission criterion).

	Pregnancy Period	Postnatal Period	Toddler Period	Preschool Period
Emotional	0.0286***	0.0125**	0.0105**	0.0210***
	(0.0154)	(0.0067)	(0.0057_	(0.0113)
Physical	0.0138**	0.0061	0.0114**	0.0236***
	(0.0021)	(0.0009)	(0.0017)	(0.0036)
Social	0.0237***	0.0179***	0.0084*	0.0228***
	(0.0041)	(0.0031)	(0.0014)	(0.0040)
Cognitive	0.0006	0.0072	0.0027	0.0017
	(0.0004)	(0.0042)	(0.0015)	(0.0009)
Communication	0.0070	-0.0123*	0.0075*	-0.0035
	(0.0037)	(0.0065)	(0.0039)	(0.0018)

Notes: each model is adjusted for some variables that include (i) child health through early childhood (major illness, minor illness, hospital admission); (ii) child health at birth (low birth weight, preterm, Apgar, hospital stay, emergency c-section); (iii) mother health prior to birth (major illness, minor illness, hospital admission); (iv) demographic and socioeconomic variables. The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 2.21 indicates the effect of maternal depression exposure among children who identified as healthy using the number of hospital admission from birth through the child's 5th

birthday criterion. Exposure to maternal depression during pregnancy period has the strongest effect on at least three domains of child development, followed by the preschool period. However, the effect of maternal depression exposure disappears on the cognitive and communication domains across the four timing points of maternal depression exposure. Moreover, Emotional and social domains are more sensitive to maternal depression across the four timing periods.

2.8.4 Cut-off Sensitivity

Some variables were dichotomized based on established cut-off values from previous literature. For these variables, I used their distribution to determine the 90^{th} and 95^{th} percentile values to use as a cut-off point. For all these cases sensitivity analysis was performed using slightly higher or lower cut-off values than reported in the literature. The results did not change after changing the cut-off values. The set of variables that I re-dichotomized includes child major illness (ADG Major), child minor illness (Minor ADGs), number of child hospital admission through childhood, mother major illness (Major ADGs), mother minor illness (Minor ADGs), and mother hospital admission prior to pregnancy.

2.8.5 Stepwise Selection

When I built the main effects model, I applied the traditional and most extensively used procedure, the purposeful selection method, for the variables selection process. Another approach to select variables is a stepwise procedure in which variables are selected in a sequence pattern based only on statistical criteria. Some studies found that the purposeful selection method selects significant variables and includes variables that could be confounders of the other model

variables in a practical manner superior to the stepwise selection method[7] (Bursac et al., 2008). However, I also used the stepwise selection procedure for all variables and interactions in the model and the selected variables were similar to the purposeful selection procedure, except for some variables that are clinically important in previous literature. That is, the stepwise method selected the previously identified variables in the main effects model.

[7] There are some other method of selection which are not used extensively in the logistic regression (e.g., Best Subsets Selection). However, this method, as mechanical selection, has been criticized because it can select irrelevant variables. For more detail see Greenland, 1989; Griffiths and Pope, 1987.

2.9. Discussion and Conclusion:

This study examined the school readiness of children with a history of maternal depression exposure. It significantly contributes to the literature by investigating the impact of the timing of first exposure to maternal depression on a comprehensive measure of children's school readiness that incorporates multidimensional developmental domains that underlie academic school class adaptation and later success. The Early Development Instrument (EDI) scores of 59,413 children were linked to their mothers and followed over time from five years before the child's birth to the child's 5th birthday. Understanding how the timing of maternal depression exposure relates to children's school readiness is important for informing appropriate supportive policies for children of depressed mothers.

One of the unique strengths of the administrative database at Manitoba Center for Health Policy is the ability to follow up mother-child dyads over time and to assign them to one of the initial exposure timing points. The focus on the first exposure to maternal depression is necessary to help isolate the effects resulting from each exposure period and to control for the issue of overlapping repeated depression episodes. Essentially, children exposed to maternal depression in any period were not included as exposed to maternal depression at any later period. I assigned each mother-child dyad to 1 of 4 mutually exclusive exposure timing categories. Specifically, these are: children not exposed to maternal depression at any point (reference group); children first exposed to maternal depression during pregnancy; children first exposed to maternal depression during the postnatal period (birth–12 months); children first exposed to maternal depression during the toddler period (12 months - 36 months); and children first exposed to maternal depression during the preschool period (36 months - 60 months).

The findings suggest that children exposed to maternal depression are at risk for poor school readiness due to maladaptive development of skills and abilities that are the foundation for school adaptation. When controlling for child health at birth, child health through early childhood (as indicated by major ADGs and minor ADGs, and hospital admission frequency), and mother's health prior to pregnancy, however, the effects of exposure to maternal depression on children's abilities in the emotional, physical, and social domains were attenuated across the different timing periods. This effect almost disappeared for cognitive and communication skills. Furthermore, in addition to child health (from birth through childhood), family environment (e.g., unmarried mothers, teenage mothers, a mother with less than high school education, families on social assistance, and families with more than four children) is also a strong risk factor for school readiness, over and above the disadvantage conferred by maternal depression exposure. The effect of family socioeconomic characteristics on the child's EDI domains varies across the developmental domains, but it is very strong, especially for cognitive and communication skills. These findings suggest that programs that target depressed mothers to support disadvantaged families and improve child and mother's health could lead to improvements in school readiness.

Generally, exposure to maternal depression during pregnancy appears to have the strongest effect on children's development, as measured by the EDI, followed by the preschool period. Sensitivity during pregnancy on children's developmental outcomes is supported by studies that found that exposure to maternal depression during the prenatal period is associated with increased infant cortisol levels, low birth weight, preterm birth (Gutteling et al., 2005; Brennan et al., 2008; Field, 2011; Field et al., 2008) less mother-child attachment and decreased breastfeeding initiation, and therefore child development. There is a general consensus among

different disciplines that children who experienced early developmental difficulties may need additional support to increase their chance of success and unproblematic transition to school. Therefore, the prevention of developmental impairments in children of depressed mothers requires more attention to early detection of maternal depression in obstetrician offices, where exposure to elevated cortisol level for fetuses and low birth weight can be identified, which are risk factors for impaired cognitive and social development outcomes (Wadhwa, 2005). Additionally, interventions should commence prior to the start of school because the ongoing challenging school environment may exacerbate the child's developmental difficulties (Bell et al., 2018). Interventions should target very young children and their mothers to mitigate early developmental difficulties due to maternal depression exposure. Programs should give more support to the emotional, physical, and social competence domains for children exposed to maternal depression because these are the most affected domains. Early intervention can moderate negative outcomes of the neurobiological response to the poor environment surrounding the child through the plasticity and resilience[8] of brain development and other biological systems (Thompson, 2014).

The findings show a significant difference in odds of developmental vulnerability depending on the timing of the first exposure to maternal depression. This heterogeneity in the effect of timing was across all EDI domains, though especially in the emotional, physical and social domains, even after controlling for sociodemographic characteristics. Thus, the timing of maternal depression exposure exerts heterogeneous effects on school readiness. I found a significant difference in the odds of developmental vulnerability when comparing children exposed to maternal depression during pregnancy to exposure during the other three periods

[8] Plasticity refers to the capacity of organisms to alter with the surrounding experience. Plasticity is highest in early life and declines with age. This explains why early shocks have a bigger impact on younger children than older ones (for more detail see Thompson, 2014).

(postnatal, toddler, and preschool). There is also a significant difference in the log odds of being vulnerable in the emotional domain when comparing toddler and preschool groups. Children exposed to maternal depression during pregnancy had greater odds of being not ready on the emotional domain at kindergarten age compared with those exposed to maternal depression during the postnatal period (17% higher odds), the toddler period (20% higher odds), and in preschool (8% higher odds). That is, exposure to maternal depression during pregnancy has the strongest effect on the child's emotional maturity domain compared with the other timing points. Children exposed to maternal depression during the preschool period are more likely to be not ready on the emotional domain than children exposed to maternal depression during the toddler period. Two significant differences across the different time periods were found for physical health. Children exposed to maternal depression during preschool are more likely to be vulnerable in the physical domain compared to those exposed to maternal depression during the postnatal or toddler periods. Similarly, children exposed to maternal depression during preschool are more likely to be vulnerable in the physical domain compared with those exposed to maternal depression during the toddler period. Thus, the preschool period is a sensitive period for physical health development. There were two significant differences in the odds of developmental vulnerability in the social competence domain. Children exposed to maternal depression during pregnancy had higher odds of being vulnerable compared to those exposed to maternal depression during the toddler period. Similarly, children exposed to maternal depression during the preschool period had higher odds of being not ready compared with those exposed during the toddler period. For both cognitive and communication domains, there was no significant difference in the odds of being vulnerable when comparing each pair of maternal depression

timing points. However, these children were also at risk of poor school readiness irrespective of the timing of maternal depression.

A child's emotional domain is the most sensitive domain to maternal depression exposure during pregnancy or the preschool period. Compared with findings from the area of a child's emotional development (especially during the preschool period), the latter is surprising. Exposure to maternal depression during the preschool period (36 months – 60 months) comes after the recognized child-mother attachment period. In the analyses, children in this group were not exposed to maternal depression at any other time point (since these include only first exposure) and hence they should have been able to form a healthy mother attachment bond before the child's 3rd birthday. However, psychologists assume (Bowlby, 1980; Mancini et al., 2009) that children, in the 2 – 5 year age, who lose the healthy attachment bond during a depression episode may experience a loss similar to the actual loss of a parent This can explain the stronger effect of maternal depression during preschool than the postnatal period, where the emotional cost could be stronger after the establishment of the attachment bond rather than during the attachment period itself (Naicker et al., 2012).

The results of this study suggest that a child's family environment has a strong effect on their cognitive and communication domains compared to other domains. Family size has the strongest effect on the communication and knowledge domain, followed by the language and cognitive domains. Income assistance has the strongest negative effect on children's language and cognitive domain, followed by the communication domain. These results align with another MCHP study that found that socioeconomics status and family disadvantage have a strong effect on those domains (Santos et al., 2012).

The literature reports different channels through which maternal depression exposure can affect early child development. In psychology, healthy parent interactions have a pivotal role in how children respond to and interact with their surrounding environment. These interactions help children learn the expression of positive and negative emotions that can help them build their own coping methods, called 'emotion regulation and knowledge' (Kujawa et al., 2014). Depressed mothers represent an obstacle to the development of these skills and serve as maladaptive models of emotion regulation (Blandon et al., 2008). Early impairments in children's 'emotion knowledge' can extensively contribute to developmental vulnerability in social competence (Kujawa et al., 2014; Wang and Dix, 2015). Depressed mothers are less sensitive to their children's needs and engage in more negative social communication (Feldman et al., 2009), which disrupts the needed skills during this period to build healthy social relationships. Children who experience maladaptive development of these skills are less socially involved and excluded from their peers; moreover, they usually develop aggressive communication and attribute blame on others (Wang and Dix, 2015). Exposure to depression may also disrupt the interaction between mothers and their children through decreasing the stimulated-engaging activities that develop skills such as memory, language, reading ability, exploring the surrounding world (Stein et al., 2008; Keim et al., 2011), which results in poor cognitive performance (Keim et al., 2011; Perra et al., 2015). In addition to the psychological mechanism, maternal depression can impact children's skills formation through children's brain physiological mechanisms. Some studies in neuroscience have been working on brain biological vulnerabilities as a mechanism through which maternal depression can affect children's cognitive and language development.

The mother's ill health (as indicated by Major ADGs, Minor ADGs, and the number of hospital admissions) prior to child's birth has a significant negative effect on children's school readiness. Among the mother's latent health variables, the number of hospital admissions has the strongest effect on a child's development across all domains. The cognitive and language domain is the most affected by the mother's poor health prior to birth, followed by the child's physical health domains. However, the effect of the mother's health on a child's school readiness was attenuated after adjusting for child family environment factors.

The results showed that boys are more likely to be developmentally vulnerable on all EDI domains than girls, and younger children at the time of EDI assessment are more vulnerable than older children. There is also gender heterogeneity in the effect of maternal depression exposure on children's school readiness. In all categories of maternal depression timing, the effect of depression seemed greater for boys than girls on the emotional skills. Children of unmarried mothers are more sensitive to depression exposure during pregnancy and the toddler period. Marital status moderates the effect of maternal depression on emotional, physical, and social health. However, maternal depression appears to have no effect by marital status on cognitive and communication domains. The effect of maternal depression does not depend on the education level of the children's mother across all periods and EDI domains, except for the emotional domain during pregnancy and the preschool period.

Further research that elaborates the overlapping impact of timing and severity of maternal depression could be informative in designing the appropriate programs that support children who are exposed to the different severity levels of maternal depression.

2.10. Appendices

Appendix A: Variables Descriptions

Variable	Explanation
Pregnancy	=1 if child exposure to maternal depression during pregnancy
postnatal	=1 if child exposure to maternal depression for the first time during the postnatal period (from birth to 12 months)
Toddler	=1 if child exposure t maternal depression for the first time during toddler period (12 months – 36 months)
Preschool	=1 if child exposure to maternal depression for the first time during the preschool period (36 months – 60 months)
Major illness	=1 if a child has a number of Major ADGs greater than the 90^{th} percentile of the distribution of the number of major ADGs from birth to the child's 5^{th} birthday.
Minor illness	=1 if a child has a number of Minor ADGs greater than the 90^{th} percentile of the distribution of the number of minor ADGs from birth to the child's 5^{th} birthday.
Child hospital admission	=1 if the number of hospital admission greater than the 95^{th} of the distribution of the number of hospital admission from birth to child's 5^{th} birthday.
Low Birth Weight	=1 if a child's birth weight is less than 2500 gm
Preterm	=1 if the gestation
5-minutes Apgar	=1 if child's score on the 5minutes APGAR greater than or equal 8
Length of hospital stay	=1 if a child stay at the hospital more than 6 days at birth
Emergency c-section	=1 if a child is delivered by emergency c-section
Mother Major illness	=1 if a mother has a number of Major ADGs greater than the 90^{th} percentile of the distribution of the number of major ADGs through 5 years prior to birth.
Mother Minor illness	=1 if a mother has a number of Minor ADGs greater than the 90^{th} percentile of the distribution of the number of minor ADGs through 5 years prior to birth.
Mother hospital admission	=1 if a mother has a number of hospital admission greater than the 95^{th} percentile of the distribution of the number of hospital admission through 5 years prior to birth.
Child's age (in months)	Continuous variable from birth to the time of EDI

	assessment
Child gender	= 1 if child is male
Breastfeeding initiation	=1 if the mother-initiated breastfeeding at birth or mixed breastfeeding with artificial milk
Mother married	=1 if a mother is married or in common law
Mother has less than HS	=1 if a mother does not have a high school degree at birth
Teenage mother	=1 if a mother is teenage at birth (less than 19 years)
Smoking during pregnancy	=1 if a mother is smoking during pregnancy
Neighbourhood-area SES	=1 if child family living in favorable socioeconomic
Urban residence	= if child family leaving in Winnipeg or Brandon
Family size	=1 if the child has more than 4 siblings
Family receive income assistance	=1 if child family receive income assistance

Appendix: B

Variables Description and Database Source

Variable	Variable description	Source
Cognitive & Language Development (domain), Physical Health (domain), Emotional maturity (domain), Social Competence (domain), Communication skills & Knowledge (domain)	Children at kindergarten age can be classified as "not ready", "ready", "very ready", based on the score distribution inside each domain. That is, children who score in the bottom 10th percentile cut-off score are classified as "not ready" or "vulnerable" for school, based on skills in that domain. In contrast, children who score in the top 30th percentile cut-off score are classified as "very ready" for school. Children who score between 11th and 69th percentile are classified as "ready" to school. (outcome variables)	Early Development Instrument (EDI) Database
Maternal depression (Binary)	Depressed mothers are identified as follow: (i) at least one hospitalization with any of ICD-9-CM diagnosis codes 296.1-296.8, 300.0, 300.2-300.4, 300.7, 309, 311or ICD-10-CA codes F31, F32, F33, F34.1, F38.0, F38.1, F40, F41.0-F41.3, F41.8, F41.9, F42, F43.1, F43.2, F43.8, F45.2, F53.0, F93.0; OR (ii) at least one hospitalization (any dx code) with ICD-9-CM code 300 or ICD-10-CA codes F32, F34.1, F40, F41, F42, F44, F45.0, F45.1, F48, F68.0, F99 AND one or more prescription for antidepressant or mode stabilizer. (iii) at least one physician visits (prefix=7) with ICD-9-CM codes 296, 311 (iv) at least one physician visits (prefix=7) with ICD-9-CM code 300 AND one or more prescription for antidepressant or mode stabilizer; OR (v) at least three physician visits (prefix=7) with ICD-9-CM codes 300 or 309 (must be 3 of same dx code). Drugs to treat mood and anxiety disorders include (i) Antidepressants, ATC code N06A; (ii) Benzodiazepine Derivatives Anxiolytics, ATC code N05BA and (iii) Lithium, ATC code N05AN01. (Martens et al., 2004; Fransoo et al., 2009; Martens et al., 2010; Chartier et al., 2012).	Hospital Abstracts data, Medical Services data and DPIN data
Timing of first exposure to maternal depression (categorical)	This study identifies four sensitive periods in early child's life which have some developmental characteristics that make them independent of other periods. these sensitive periods cover the period from pregnancy onset to the child's kindergarten age as follows: (i) pregnancy period, which includes the nine months prior to the child's date of birth. (ii) postnatal period, which includes the time from the child's date of birth and extends up to the day before the child's first birthday (birth – 12 months. (iii) Toddler period: This period begins from the child's first birthday and extend up to the extends up to the child's third birthday (12months – 36 months); (v) preschool period: this period begins from the child's third birthday and extends up to the child's kindergarten age (36months – 60 months).	Hospital Abstracts data, Manitoba Health Insurance Registry, Medical Services data
Child's age (continuous)	Age in months (continuous). age in months was calculated using birthdate from the population registry and date of assessment from the EDI database and entered as a continuous variable.	Manitoba Health Insurance Registry
Child's sex	The biological sex of the child was identified using the population registry.	Manitoba Health Insurance Registry
Marital status (Binary)	Marital status (married/unmarried) is defined using the Manitoba Health Insurance Registry data and supplemented with data from the Families First Screen (Baby First screen available 2000 – 2003;	Manitoba Health Insurance Registry;

	Families First available between 2003 and 2014; maintained by Healthy Child Manitoba).	Families First Screen; Baby First Screen
Neonatal Intensive Care Unit (NICU) admission at birth (binary).	NICU is a binary indicator of whether the child was admitted to an intensive or intermediate care unit during their birth hospitalization episode (including transfers)	hospital abstracts database.
Length of stay at birth (Binary).	It measures the number of days stayed at the hospital at birth. It is calculated as the number of days of care from the child's date of birth to the discharge date. This also is initially meant to serve as a proxy of the severity of health at birth. It is a binary variable. Long birth staying at hospital equals 1 if the number of days at hospital at birth is more than 6 days; 0 otherwise.	hospital abstracts database.
Breastfeeding initiation (Binary)	Breastfeeding variable was created from data in each child's birth hospital record to indicate whether breastfeeding was initiated in the hospital at the time of birth. Exclusive breastfeeding and breastfeeding mixed with artificial feeding are combined in one category, and exclusive artificial feeding is the other category.	hospital abstracts database.
APGAR score (continuous)	Apgar score at 5 minutes after birth (scaled score: 0-10). Five-minute Apgar score, which measures the physiological well being of new babies at five minutes after birth. A score of zero, one, or two is given for each of five vital signs which include infant's heart rate, respiration, muscle tone, reflex, and color, and a total score less than 8 out of 10 reflect problem (Oreopoulos et al., 2008; Jutte et al., 2010; Roos et al., 2011).	hospital abstracts database.
Low Birthweight (Binary)	Birthweight (grams) is a binary variable where low birth weight defined as less than 2500gm	hospital abstracts database.
Small for gestational age (SGA). (Binary)	Gestational age (maternal self-reported number of weeks since last menstrual period). Small for gestational age (SGA), which is identified at or below the 10th percentile in birthweight from an infant population of the same sex and gestational age (Kramer et al., 2001; Jutte et al., 2010; Chen et al., 2014)	hospital abstracts database.
Preterm birth (Binary)	Preterm birth is a dichotomous measure of whether the child was born 'preterm' (before 37 complete weeks of gestation)	hospital abstracts database.
C-Section delivery (Binary)	The child is delivered by emergency Caesarean section (indicator variable: 0/1)	hospital abstracts database.
The Aggregated Diagnostic Grouping (ADG) -Major ADG -Minor ADG They are dichotomized based on the 90th percentile of the distribution of the number of ADGs	The Aggregated Diagnostic Grouping (ADG) variable is used as a measure of a child's health status from birth to the child's 5th birthday (Child's major and minor illness in early childhood from birth until age 5 years. Based on the Johns Hopkins ACG® Case-Mix System (computer program), patients are assigned an ADG if they have one or more of the ADG's constituent diagnoses coded on at least one physician claim or hospital separation record (Johns Hopkins School of Public Health, 2009). According to MCHP (2013), ICD-9 and ICD-10 codes are assigned to one of 32 different ADGs, of which 8 are considered Major for children, and 24 are considered Minor. Each year, every child could be assigned to any number of ADGs, based on diagnoses attributed to them from physician visits or hospital episodes. For each child, the number of major and minor ADGs to which they are assigned each year are counted, and then summed over the entire preschool period (birth to 5th birthday).	Hospital abstract, physician visits, and John Hopkins ACG system at MCHP

Family size 4+ (Binary)	Family size 4+ children: a dichotomous measure of whether the child's mother had 4 or more children, as of the child's 5th birthday. Family size was first measured based on counting the number of children born to a mother and/or counting the number of children with the same Manitoba Health Registration Number at the same date(s) by matching the mother and baby records. These values were then dichotomized into two categories: whether the family had 4 or more children or not.	Manitoba Health Insurance Registry
Receiving income assistance (Binary)	using the Social Assistance Management Information Network (SAMIN) data; this dataset is maintained by Manitoba Jobs and the Economy and has information for individuals starting on April 1, 1995. Receiving employment and income assistance (EIA) will be defined as a binary measure; mothers receiving at least two consecutive months of EIA in a given time period are identified as receiving EIA.	Social Assistance Management Information Network (SAMIN) data.
Location of Neighborhood	The location of the neighborhood (urban, rural south, rural mid, rural north) at the time of the birth of the child will also be included as women in rural neighborhoods (particularly those in the rural north) may be more isolated and have less access to social services. In Manitoba, urban neighborhoods include those located in Winnipeg and Brandon, the rural south includes neighborhoods in the South Eastman, Central, and Assiniboine Regional Health Authorities, neighborhoods in rural mid/north include those in the North Eastman, Interlake, and Parkland Regional Health Authorities and neighborhoods in northern Manitoba include those in Nor-Man, Churchill, and Burntwood Regional Health Authorities.	Manitoba Health Insurance Registry
Mother is teenage at birth	The mother's age at the time of childbirth is extracted from the population registry and hospital abstract. The mother identifies as a teenage mom if her age at birth is less than 19 years.	hospital abstracts database, and Manitoba Health Insurance Registry
Mother has High School at Time of Child's Birth (Binary).	A binary variable of whether the mother graduated high school by the birth of her child.	Enrollment, Marks, and Assessments data; Families First Screen FF; Baby First Screen BF
Substance Use During Pregnancy (Binary).	A binary variable of whether the mother uses alcohol and drug during pregnancy. It includes mothers who self-reported consuming alcohol and the use of substances during pregnancy on the BF/FF screening form.	Families First Screen FF; Baby First Screen BF
Smoking During Pregnancy (Binary)	A binary variable of whether the mother is Smoking During Pregnancy. This indicator includes women who self-reported smoking during pregnancy on the BF/FF Screening Form	Families First Screen FF; Baby First Screen BF
Neighborhood-area Socioeconomic status	A child's family socioeconomic status will be captured by using the Socioeconomics Factor Index II(SEFI-2) at MCHP. The SEFI-2 will be linked to the child's mother postal code to determine the socioeconomic status of the child's family.	Canada Census (Statistics Canada) at MCHP.

2.10. References:

1. Aizer, A. & Stroud, L. & Buka, S. (2016). Maternal Stress and Child Outcomes: Evidence from Siblings. *Journal of Human Resources* 51(3), 523-555. University of Wisconsin Press. Retrieved August 20, 2019, from Project MUSE database.
2. Aktar, E1., Qu, J., Lawrence, PJ., Tollenaar, MS., Elzinga, BM1., Bögels, SM5. (2019). Fetal and Infant Outcomes in the Offspring of Parents With Perinatal Mental Disorders: Earliest Influences. *Front Psychiatry*. 27, 10:391. DOI: 10.3389/fpsyt.2019.00391. eCollection 2019.
3. American Academy of Pediatrics. (1997). Breastfeeding and the Use of Human Milk. Work Group on Breastfeeding, *Pediatrics*, 100(6), 1035-9. DOI: 10.1542/peds.100.6.1035
4. Augustine, J. M., & Crosnoe, R. (2010). Mothers' depression and educational attainment and their children's academic trajectories. *Journal of health and social behavior*, *51*(3), 274–290. DOI:10.1177/0022146510377757
5. Baron, I.S., Erickson, K., Ahronovich, M.D., Baker, R., & Litman, F.R. (2011). Cognitive deficit in preschoolers born late-preterm. *Early human Development*, 87, 115-119.
6. Baron, I.S., Erickson, K., Ahronovich, M.D., Kelly Coulehan, K., Baker, R., Litman, F.R. (2009) Visuospatial and verbal fluency relative deficits in 'complicated' late-preterm preschool children. *Early Human Development*, 85(12), 751-754.
7. Bell, M.A., Wolfe, C. D. (2004). Emotion and Cognition: An Intricately Bound Developmental Process, *Child Development*, 75(2), 366-370.
8. Bhutta, A.T., Cleves, M.A., Casey, P.H., Cradock, M.M., & Anand, K.J. (2002). Cognitive and Behavioral Outcomes of School-Aged Children Who Were Born Preterm A Meta-

analysis. *Journal of the American Medical Association*, 288(6), 728-737. DOI:10.1001/jama.288.6.728.

9. Blandon, AY., Calkins, SD., Keane, SP., O'Brien, M. (2008). Individual differences in trajectories of emotion regulation processes: the effects of maternal depressive symptomatology and children's physiological regulation. *Dev*elopmental *Psychol*ogy. 44 (4), 1110-1123. DOI: 10.1037/0012-1649.44.4.1110.
10. Bowlby, J. (1970). Disruption of affectional bonds and its effects on behavior. Journal of Contemporary Psychotherapy, 2(2), 75-82. DOI:10.1007/BF02118173.
11. Bowlby, J. (1980). Attachment and loss: Vol. 3. Loss: Sadness and depression. New York: Basic Books.
12. Bowlby, J. (1982). *Attachment and Loss*. Volume I Attachment, Second Edition, New York: Basic Books.
13. Brennan, P. A., Pargas, R., Walker, E. F., Green, P., Newport, D. J., & Stowe, Z. (2008). Maternal depression and infant cortisol: influences of timing, comorbidity and treatment. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *49*(10), 1099–1107. http://doi.org/10.1111/j.1469-7610.2008.01914.x
14. Bretherton, I. (1985). Attachment Theory: Retrospect and Prospect. *Monographs of the Society for Research in Child Development, 50*(1/2), 3-35. doi:1. Retrieved from http://www.jstor.org/stable/3333824 doi:1
15. Brinkman. S., Gregory, T., Harris, J., Hart, B, Blackmore, S., Janus. M. (2013). Associations between the early development instrument at age 5, and reading and numeracy skills at ages 8, 10 and 12: a prospective linked data study. *Child Indicators Research*. 6(4), 695–708.

16. Britto, P. R., Boller, K., & Yoshikawa, H. (2011). Quality of early childhood development programs in global contexts: Rationale for investment, conceptual framework, and implications for equity. *Social Policy Report*, *22* (2), 1-31. Retrieved from: http://www.srcd.org/sites/default/files/documents/spr_v252rev.pdf

17. Brownell, M., Chartie,r M., Santos, R., Ekuma, O., Au, W., Sarkar, J., MacWilliam, L., Burland E., Koseva, I., Guenette, W. (2012). How Are Manitoba's Children Doing? Winnipeg, MB. Manitoba Centre for Health Policy, October 2012.

18. Bursac, Z., Gauss, C.H., Williams, D.K. et al. (2008). Purposeful selection of variables in logistic regression. Source Code Biol Med. 3, 17. DOI:10.1186/1751-0473-3-17

19. Buss C., Davis, E.P., Shahbaba, B., Pruessner, J.C., Head, K., Sandman, C.A. (2012). Maternal cortisol over the course of pregnancy and subsequent child amygdala and hippocampus volumes and affective problems. Proceedings of the National Academy of Sciences of the United States of America (PNAS). May 2012, 109 (20) E1312-E1319; DOI: 10.1073/pnas.1201295109

20. Case, A., Fertig, A., Paxson, C. (2005). The lasting impact of childhood health and circumstance. *Journal of Health Economics*. 24(2), 365–389.

21. CDCHU- Center on the Developing Child at Harvard University (2009). Maternal Depression Can Undermine the Development of Young Children: Working Paper No. 8. http://www.developingchild.harvard.edu

22. Center on the Developing Child at Harvard University. (2009). Maternal Depression Can Undermine the Development of Young Children: Working Paper No. 8. http://www.developingchild.harvard.edu

23. Center on the Developing Child. (2010). The Foundations of Lifelong Health Are Built in Early Childhood. Retrieved from www.developingchild.harvard.edu.

24. Chartier, M., Finlayson, G., Prior, H., McGowan, K., Chen, H., de Rocquigny, J., Walld, R., Gousseau, M. (2012). Health and Healthcare Utilization of Francophones in Manitoba. Winnipeg, MB: Manitoba Centre for Health Policy.

25. Chen, J. H., Claessens, A., & Msall, M. E. (2014). Prematurity and school readiness in a nationally representative sample of Australian children: does typically occurring preschool moderate the relationship? *Early Human Development*, 90(2), 73-79.

26. Chyi, L.J., Lee, H.C., Hintz, S.R., Gould, J.B., & Sutcliffe, T.L. (2008). School Outcomes of Late Preterm Infants: Special Needs and Challenges for Infants Born at 32 to 36 Weeks Gestation. *Journal of Pediatrics*, 153(1), 25-31.

27. Cicchetti, D., Rogosch, F. A., & Toth, S.L. (2000). The efficacy of toddler-parent psychotherapy for fostering cognitive development in offspring of depressed mothers. Journal of Abnormal Psychology. 28(2), 135-148.

28. Claessens, A., Engel, M., & Curran, F.C. (2015). The Effects of maternal depression on child outcomes during the first years of formal schooling. *Early Childhood Research Quarterly*, 32, 80-93.

29. Cohen, M.M., MacWilliam, L. (1994). Population Health: Health Status Indicators. Volume II: Tables and Figures. Winnipeg, MB: Manitoba Centre for Health Policy and Evaluation.

30. Couperus, J.W., & Nelson, C.A. (2006). Early brain development and plasticity. In K.McCartney & D. A. Phillips (Eds.), *Handbook of early childhood development* (pp. 85-105). Oxford: Blackwell Publishing.

31. Cunha, F., and Heckman, J. (2007). The technolgy of skill formation. *American Economic Review*. 97(2), 31-47.

32. Cunha, F., Heckman, J. J. and Schennach, S. M. (2010), Estimating the Technology of Cognitive and Non-Cognitive Skill Formation. *Econometrica*, 78: 883–931. DOI:10.3982/ECTA6551

33. Cunha, F., Heckman, J., Lochner, L. and Masterov, D. (2005). Interpreting the evidence on life cycle skill formation. In F. W. E. Hanusheck (Ed.), *Handbook of the economics of education* (pp. 697-812). Amsterdam, The Netherlands: NorthHolland.

34. Dahlen, H. M. (2016). The Impact of Maternal Depression on Child Academic and Socioemotional Outcomes. *Economics of Education Review, 52,* 77-90. doi:http://dx.doi.org.uml.idm.oclc.org/10.1016/j.econedurev.2016.01.006.

35. Davidson, R., & Fox, N. (1989). Frontal brain asymmetry predicts infants' response to maternal separation. *Journal of Abnormal Psychology, 98,* 127–131.

36. Davies S, Janus M, Duku E, Gaskin A (2016). Using the early development instrument to examine cognitive and noncognitive school readiness and elementary student achievement. *Early Childhood Research Quarterly*. 35,63–75.

37. Dawson, G., Ashman, S.B., Hessl, D., Spieker, S., Frey, K., Panagiotides, H., et al. (2001). Autonomic and brain electrical activity in securely- and insecurely-attached infants of depressed mothers. *Infant Behavior & Development*, 24(2), 135–149.

38. Dawson, G., Panagiotides, H., Grofer Klinger, L., & Hill, D. (1992). The role of frontal lobe functioning in the development of self-regulatory behavior in infancy. *Brain and Cognition, 20,* 152–175.

39. De Jong, M., Verhoeven, M., & van Baar, A. L. (2012). School outcome, cognitive functioning, and behavior problems in moderate and late preterm children and adults: a review. Seminars in Fetal and Neonatal Medicine, 17(3), 163–169. http://dx.doi.org/10.1016/j.siny.2012.02.003

40. Deave, T., Heron, J., Evans, J., Emond, A. (2008). The impact of maternal depression in pregnancy on early child development. An International Journal of Obstetrics and Gynaecology. 115(8), 1043-51. doi: 10.1111/j.1471 0528.2008.01752.x.

41. Dennis, C.L., McQueen, K. (2009). The Relationship Between Infant-Feeding Outcomes and Postpartum Depression: A Qualitative Systematic Review. *Pediatrics*, 123(4). e736-51. DOI: 10.1542/peds.2008-1629

42. Diego, M. A., Field, T., Hernandez-Reif, M., Schanberg, S., Kuhn, C., & Gonzalez-Quintero, V. H. (2009). Prenatal depression restricts fetal growth. *Early Human Development*. 85, 65-70.

43. Diego, M.A., Field, T., Jones, N.A., Hernandez-Reif, M. (2006). Withdrawn and intrusive maternal interaction style and infant frontal EEG asymmetry shifts in infants of depressed and non-depressed mothers. *Infant Behavior & Development*, 29, 220–229.

44. Downey, G., Coyne, J.C., (1990). Children of depressed parents: an integrative review. *Psychological Bulletin*, 108 (1), 50-76.

45. Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A.C., ….. Japel, C. (2007). *Developmental Psychology*, 43(6), 1428-1446. http://dx.doi.org/10.1037/0012-1649.43.6.1428

46. Elder, G. H. Jr. (1998). The life course and human development. In W. Damon and R. M. Lerner (Eds.), *Handbook of Child Psychology: Volume 1: Theoretical Models of Human Development.* New York: Wiley.

47. Emory, E. K., & Dieter, J. N. (2006). Maternal depression and psychotropic medication effects on the human fetus. Annals of the New York Academy of Sciences, 1094, 287–291. DOI:10.1196/annals.1376.036

48. Essex, M.J., Klein, M.H., Miech, R., and Smider, N.A. (2001). Timing of initial exposure to maternal major depression and children's mental health symptoms in kindergarten, *The British Journal of Psychiatry*, 179 (2), 151-156. DOI: 10.1192/bjp.179.2.151

49. Estrada, P., Arsenio, W.F., Hess, R.D., Holloway, S.D. (1987). Affective Quality of the Mother-Child Relationship: Longitudinal Consequences for Children's School-Relevant Cognitive Functioning. *Developmental Psychology*, 23(2), 210–215.

50. Fairthorne, J., Hanley, G. E., & Oberlander, T. F. (2018). Depressed Women of Low Socioeconomic Status Have High Numbers of Physician Visits in the Year Before Pregnancy: Implications for Care. *Journal of clinical medicine research*, *10*(6), 516–522. DOI:10.14740/jocmr3377w

51. Feldman, R., Granat, A., Pariente, C., Kanety, H., Kuint, J., Gilboa-Schechtman, E. (2009). Maternal depression and anxiety across the postpartum year and infant social engagement, fear regulation, and stress reactivity. *Journal of American Academy of Child* and *Adolescent Psychiatry*.;48 (9), 919-927. DOI:10.1097/ CHI.0b013e3181b21651.

52. Field, T & Diego, M. (2008). Maternal Depression Effects on Infant Frontal EGG Asymmetry, *International Journal of Neuroscience*, 118(8), 1081-1108. DOI: 10.1080/00207450701769067

53. Field, T. (2002). Prenatal effects of maternal depression. In S. H.Goodman & I. H. Gotlib (Eds.), Children of depressed parents: Mechanisms of risk and implications for treatment (pp. 59-88). Washington, DC: American Psychological Association.

54. Field, T. (2011). Prenatal depression effects on early development: A review. *infant behavior and development*, 34(1), 1-14.

55. Field, T., Diego, M., Dieter, J., Hernandez-Reif, M., Schanberg, S., Kuhn, C., Yando, R., & Bendell, D. (2004). Prenatal depression effects on the fetus and the newborn. *Infant Behavior and Development*, 27(2), 216-229

56. Field, T., Diego, M., Hernandez-Reif, M., Deeds, Schanberg, S., Kuhn, C., et al. (2008). Chronic prenatal depression and neonate outcome. *International Journal of Neuroscience*, 118, 95–103.

57. Field, T., Fox, N.A., Pickens, J., Na wrocki, T. (1995). Relative right frontal EEG activation in 3- to 6-month-old infants of "depressed" mothers. *Developmental Psychology*, 31(3), 358-363. http://dx.doi.org/10.1037/0012-1649.31.3.358

58. Forer, B., & Zumbo, B. D. (2011). Validation of multilevel constructs: validation methods and empirical findings for the EDI. Social Indicators Research, 103(2), 231–265. DOI:10.1007/s11205 011-9844-3

59. Forget-Dubois, N., Lemelin, J.P., Boivin, M., Dionne, G., Séguin, J.R., Vitaro, F., & Tremblay, R.E. (2007). Predicting Early School Achievement with the EDI: A Longitudinal Population-Based Study. *Early Education and Development*, 18(3), 405-426. DOI: 10.1080/10409280701610796

60. Fransoo, R., Martens, P., Burland, E., The Need to Know Team, Prior, H., Burchill, C. (2009). Manitoba RHA Indicators Atlas 2009. Winnipeg, MB: Manitoba Centre for Health Policy.

61. Fransoo, R.R., Roos, N.P., Martens, P.J., Heaman, M., Levin, B., Chateau, D. (2008). How Health Status Affects Progress and Performance in School: A Population-based Study. Canadian Journal of Public Health, 99(4), 344-349.

62. Georgiadis, A., Benny, L., Crookston, B., Duc, L., Hermida, P., Mani, S., … Behrman, J. (2016). Growth trajectories from conception through middle childhood and cognitive achievement at age 8 years: Evidence from four low- and middle-income countries. SSM - Population Health, 2, 43- 54.

63. Gitau, R., Cameron, A., Fisk, N., & GIover, V. (1998). Fetal exposure to maternal cortisol. *The Lancet*, 352(9129), 707-708. DOI: http://dx.doi.org/10.1016/S0140-6736(05)60824-0

64. Glover, V., O'Connor, T.G., Heron, J., & Golding, J. (2004). Antenatal maternal anxiety is linked with atypical handedness in the child. *Early Human Development*, 79(2), 107–118.

65. Glover, V., Teixeira, J., Gitau, R., & Fisk, N.M. (1999). Mechanisms by which maternal mood in pregnancy may affect the fetus. *Contemporary Reviews in Obstetrics and Gynecology*, *11*, 1–6.

66. Goldfeld, S., O'Connor, M., Sayers, M., Moore, T., Oberklaid F. (2012). Prevalence and correlates of special health care needs in a population cohort of Australian children at school entry. *Journal of Developmental & Behavioral Pediatrics*. 33(4), 319-327.

67. Goodman, J.H., & Tyer-Viola, L. (2010). Detection, Treatment, and Referral of Perinatal Depression and Anxiety by Obstetrical Providers. *Journal of Women's Health*, 19(3).

68. Goyal, D., Gay, C., Lee, K.A. (2010). How much does low socioeconomic status increase the risk of prenatal and postpartum depressive symptoms in first-time mothers. Women Health Issues. 20(2), 96-104. DOI: 10.1016/j.whi.2009.11.003.

69. Greenland, S. (1989). Modeling variable selection in epidemiologic analysis. American Journal of Public Health. 79, 340-349.

70. Greenough, W.T. (1991). Experience as a component of normal development: Evolutionary considerations. Developmental Psychology, 27:14-17.

71. Greenough, W.T., and Black, J.E. (1992). Induction of brain structure by experience: Substrates for cognitive development. Pp. 155-200 in Developmental Behavior Neuroscience, Volume 24. M.R. Gunnar, editor; and C.A. Nelson, editor. , eds. Hillsdale, NJ: Erlbaum.

72. Griffiths, W.E., Pope, P.J. (1987). Small sample properties of probit models. Journal of the American Statistical Association, 82, 929-937.

73. Grimm, K. J., Steele, J.S., Mashburn, A.J., Burchinal, M., & Pianta, R.C. (2010). *Developmental Psychology*, 46(5), 976-983. http://dx.doi.org/10.1037/a0018878

74. Guhn M, Gadermann AM, Almas A, Schonert-Reichl KA, Hertzman C. (2016). Associations of teacher-rated social, emotional, and cognitive development in kindergarten to self-reported wellbeing, peer relations, and academic test scores in middle childhood. *Early Childhood Research Quarterly*. 35,76–84.

75. Guhn, M., Zumbo, B., Janus, M., & Hertzman, C. (2011). Validation theory and research for a population-level measure of children's development, wellbeing, and school readiness. *Social Indicators Research*, 103(2), 183–191. DOI:10.1007/s11205-011-9841-6.

76. Gutteling, B.M., de Weerth, C., Buitelaar, J.K. (2005). Prenatal stress and children's cortisol reaction to the first day of school. *Psychoneuroendocrinology*. 2005 Jul, 30(6), 541-9.

77. Hammen, C., Brennan, P.A. (2003). Severity, Chronicity, and Timing of Maternal Depression and Risk for Adolescent Offspring Diagnoses in a Community Sample. *Archives of General Psychiatry*, 60(3):253-258. DOI:10.1001/archpsyc.60.3.253.

78. Hay, D.F., Pawlby, S., Sharp, D., Asten, P., Mills, A., Kumar, R. 2001. Intellectual problems shown 11-year-old children whose mothers had postnatal depression. *Journal of Child Psychology and Psychiatry*. 42, 871–890.

79. Heckman, J. (2007). The economics, technology, and neuroscience of human capital formation. *Proceedings of the National Academy of Sciences of the USA.* 104(33), 13250-13255.

80. Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics, 24 (3), 411-482.

81. Heckman, J., and Krueger, A. (2003). Inequality in America: What Role for Human Capital Policies. Cambridge: MIT Press.

82. Heckman, J., Pinto, R., and Savelyev, P. (2013). Understanding the mechanisms through which an influential early childhood program boosted adult outcomes. *American Economic Review*. 103(6), 2052-86.

83. Heckman, J., Seong, H., Pinto, R., Savelyev P., and Yavitz, A. (2010). Analyzing social experiments as implemented: A reexamination of the evidence from the High Scope Perry preschool program. *Quantitative Economics.* 1(1), 1-46.

84. Heckman, J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. Journal of Labor Economics. (24), 411-482.

85. Houston, S. M., Herting, M. M., & Sowell, E. R. (2014). The neurobiology of childhood structural brain development: conception through adulthood. Current topics in behavioral neurosciences, 16, 3–17. doi:10.1007/7854_2013_265

86. Huizink, A., Robles de Medina, P.G., Mulder, E.J.H., Visser, G.H.A., & Buitelaar, J.K. (2003). Stress during pregnancy is associated with a developmental outcome in infancy. *Journal of Child Psychology and Psychiatry*, 44, 810–818.

87. Isaacs, E.B., Lucas, A., Wood, S.J., Hohnson, C.L., Vargha-Khadem, F., Gadian, D.G. (2000). Hippocampal Volume and Everyday Memory in Children of Very Low Birth Weight. *Pediatric Research*. 47, 713–720.

88. Jackson, MI. (2009). Understanding links between adolescent health and educational attainment. *Demography*. 46(4), 671–694.

89. Jaffee, S. R. (2002). Pathways to adversity in young adulthood among early childbearers. *Journal of Family Psychology*,16(1), 38–49.

90. Janus, M., & Offord, D. (2007). Development and psychometric properties of the early development instrument (EDI): a measure of children's school readiness. *Canadian Journal of Behavioural Science*, 39(1), 1–22.

91. Janus, M., Brinkman, S., & Duku, E. (2011). Validity and psychometric properties of the Early Development Instrument in Canada, Australia, the United States, and Jamaica. *Social Indicators Research*, 103(2), 283–297. DOI:10.1007/s11205-011-9846-1.

92. Janus, M., Duku, E. (2007). The school entry gap: socioeconomic, family, and health factors associated with children's school readiness to learn. *Early Education and Development*. 18(3), 375-403.

93. Janus, M., Reid-Westoby, C. (2016). Monitoring the development of all children: the Early development instrument. Early Childhood Matters, Bernard van Leer Foundation's journal, 125, 40-45.

94. Jefferis, B.J.M.H., Power, C., Hertzman, C. (2002). Birth weight, childhood socioeconomic environment, and cognitive development in the 1958 British birth cohort study. *British Medical Journal*, 325,305. DOI: http://dx.doi.org/10.1136/bmj.325.7359.305

95. Johns Hopkins ACG (2014). The Johns Hopkins ACG® System. Excerpt from Version 11.0 Technical Reference Guide November 2014

96. Johnson, M. (2005). Sensitive periods in functional brain development: Problems and prospects. *Dev Psychobiology*, *46*(3), 287-92.

97. Jones, N., Field, T., Fox, N.A., & Gomez, C. (2001). EEG during different emotions in 10-month-old infants of depressed mothers. Journal of Reproductive and Infant Psychology, 19(4), 295-312.

98. Jones, N.A., Field, T., & Davalos, M. (200). Right Frontal EEG Asymmetry and Lack of Empathy in Preschool Children of Depressed Mothers. *Child Psychiatry and Human Development*, 30(3), 189–204.

99. Jutte, D., Roos, N., Brownell, M., Briggs, G., MacWilliam, L., & Roos, L. L. (2010). The ripples of adolescent motherhood: social, educational, and medical outcomes for children of teen and prior teen mothers. *Academic Pediatrics*, 10, 293–301.

100. Keim, SA., Daniels, JL., Dole, N., Herring, AH., Siega-Riz, AM., Scheidt, PC. (2011). A prospective study of maternal anxiety, perceived stress, and depressive symptoms in relation to infant cognitive development. *Early Human Development*. 87(5), 373- 380.

101. Kleinbaum, D.G., Klein, M. (2010). Logistic Regression A Self Learning Text. New York, Springer New York, Second edition.

102. Knudsen, E. (2004). Sensitive periods in the development of the brain and behavior. *Journal of Cognitive Neuroscience, 16*(8), 1412-1425.

103. Knudsen, E., J Heckman, J. Cameron, J. Shankoff. (2006). Economic, neurobiological, and behavioral perspectives on building America's future workforce. *Proceedings of the National Academy of Sciences of the USA, 103*(27), 10155-10162.

104. Koponen H, Jokelainen J, Keinanen-Kiukaanniemi S, Kumpusalo E, Vanhala M. (2008). Metabolic syndrome predisposes to depressive symptoms: a population-based 7-year follow-up study. The Journal of clinical psychiatry. 69, 178–182.

105. Kujawa, A., Dougherty, L., Durbin, CE., Laptook, R., Torpey, D., Klein, DN. (2014). Emotion recognition in preschool children: associations with maternal depression and early parenting. *Developmental Psychopathol*. 26(1), 159-170. DOI:10.1017/S0954579413000928.

106. Kurstjens, S. and Wolke, D. (2001). Maternal depression on cognitive development of children over the first seven years of life. *Journal of Child Psychology and Psychiatry*. 43(5), 623-636.

107. La Paro, K. M., & Pianta, R. C. (2000). Predicting children's competence in the early school years: A meta-analytic review. *Review of Educational Research*, 70, 443–484.

108. Laplante, D.P., Barr, R.G., Brunet, A., Du Fort, G.G., Meaney, M.L., Saucier, J.F., Zelazo, P.R., & King, S. (2004). Stress During Pregnancy Affects General Intellectual and Language Functioning in Human Toddlers. *Pediatric Research*, 56(3), 400–410. DOI:10.1203/01.PDR.0000136281.34035.44

109. Latendresse, G., Wong, B., Dyer, J., Wilson, B., Baksh, L., & Hogue, C. (2015). Duration of Maternal Stress and Depression: Predictors of Newborn Admission to Neonatal Intensive Care Unit and Postpartum Depression, *Nursing Research*, 64(5), 331-34.

110. Lundy, B. L., Jones, N. A., Field, T., Nearing, G., Davalos, M., Pietro, P., Schanberg, S., & Kuhn, C. (1999). Prenatal depression effects on neonates. *Infant Behavior and Development*, *22*, 121–137.

111. Luoma, I., Tamminen, T, Kaukonen, P., Laippala, P., Puura, K., Salmelin, R., Almqvist, F. (2001). Longitudinal Study of Maternal Depressive Symptoms and Child Well-Being, *Journal of the American Academy of Child & Adolescent Psychiatry*, 40(12), 1367-1374.

112. Mancini, A.D., Robinaugh, D., Shear, K., Bonanno, GA. (2009). Does attachment avoidance help people cope with loss? the moderating effects of relationship quality. *Journal of Clinical Psychology*. 65(10), 1127–1136.

113. Martens PJ, Derksen S, Gupta S. Predictors Of hospital readmission of Manitoba newborns within six weeks post-birth discharge: a population-based study. Pediatrics 2004; 14(3):708-713.

114. Martens, P., Brownell, M., Au, W., MacWiliam, L., Prior, H., Schultz, J., Guenette, W., Elliott, L., Buchan, S., Anderson, M., Caetano, P., Metge, C., Santos, R., Serwonka, K. (2010). Health Inequities in Manitoba: Is the Socioeconomic Gap in Health Widening or Narrowing Over Time? Winnipeg, MB: Manitoba Centre for Health Policy.

115. Martens, PJ., Fransoo, R., McKeen, N., The Need to Know Team, Burland, E., Jebamani, L., Burchill, C., De Coster, C., Ekuma, O., Prior, H., Chateau, D., Robinson, R., Metge, C. (2004). Patterns of Regional Mental Illness Disorder Diagnoses and Service Use in Manitoba: A Population-Based Study. Winnipeg, MB: Manitoba Centre for Health Policy.

116. Mattes, E., McCarthy, S., Gong, G., van Eekelen, J. A., Dunstan, J., & Foster, J. (2009). Maternal mood scores in mid-pregnancy are related to aspects of neonatal immune function. *Brain Behavior & Immunity*. 23(3), 380-388.

117. McCormick, M.P., O'Connor, E.E., Barnes, S.P. (2016). Mother-child attachment styles and math and reading skills in middle childhood: The mediating role of children's exploration and engagement. *Early Childhood Research Quarterly*, 36, 295-306.

118. McCormick, MC., Gortmaker, SL., Sobol, AM. (1990). Very low birth weight children: behavior problems and school difficulty in a national sample. Journal of Pediatrics. 117(5), 687 -693.

119. Mukherjee, S. (2011). Three Essays on Child Health and Skill Formation (Doctoral Dissertation). Retrieved from ProQuest Dissertations and Theses. UMI Number: 3456901.

120. Murray, L. (1992). The Impact of Postnatal Depression on Infant Development. *Journal of Child Psychology and Psychiatry*, 33, 543–561. DOI:10.1111/j.1469-7610.1992.tb00890.x

121. Naicker, K., Wickham, M., and Colman, I. (2012). Timing of First Exposure to Maternal Depression and Adolescent Emotional Disorder in a National Canadian Cohort. PLoS ONE, 7(3): e33422. doi:10.1371/journal.pone.0033422

122. Naicker, K., Wickham, M., Colman, I. (2012). Timing of First Exposure to Maternal Depression and Adolescent Emotional Disorder in a National Canadian Cohort. PLoS ONE. 7(3): e33422. DOI:10.1371/journal.pone.0033422

123. National Research Council and Institute of Medicine. (2000). In Jack P. Shonkoff and Deborah A. Phillips (Ed.), *From neurons to neighborhoods: The science of early childhood development.* Washington, DC: The National Academies Press.

124. National Scientific Council on the Developing Child (2005/2014). Excessive Stress Disrupts the Architecture of the Developing Brain: Working Paper No. 3. Updated Edition. Retrieved from www.developingchild.harvard.edu.

125. National Scientific Council on the Developing Child (2007). The Science of Early Childhood Development. Closing the Gap Between What We Know and What We Do. Center on the Developing Child at Harvard University. http://www.developingchild.net

126. National Scientific Council on the Developing Child (2010). Early Experiences Can Alter Gene Expression and Affect Long-Term Development: Working Paper No. 10. Retrieved from www.developingchild.harvard.edu.

127. Niederhofer, H., & Reiter, A. (2004). Prenatal maternal stress, prenatal fetal movements, and perinatal temperament factors influence behavior and school marks at the age of 6 years. *Fetal Diagnosis and Therapy*, 19, 160–162.

128. NIHCM, National Institute for Health Care Management. (2010). Identifying and Treating Maternal Depression: Strategies & Considerations for Health Plans. NIHCM Foundation Issue Brief. Retrieved from http://www.nihcm.org/pdf/FINAL_MaternalDepression6-7.pdf

129. NSCDC-National Scientific Council on the Developing Child. (2008/2012). *Establishing a Level Foundation for Life: Mental Health Begins in Early Childhood: Working Paper 6*. Updated Edition. http://www.developingchild.harvard.edu

130. O'Connor, TG., Heron, J., Golding, J., Beveridge, M., Glover, V. (2002) Maternal antenatal anxiety and children's behavioural/emotional problems at 4 years. Report from

the Avon Longitudinal Study of Parents and Children. *British Journal of Psychiatry*. 180(6), 502-508.

131. Oreopoulos, P., Stabile, M., Walld, R., Roos, L.L. (2008). Short, medium, and long term consequences of poor infant health: An analysis using siblings and twins. *Journal of Human Resources*, 2008, 43(1), 88-138.
132. Panzer, A. (2008). The neuroendocrinological sequelae of stress during brain development: The impact of child abuse and neglect. African Journal of Psychiatry, 11(1), 29–34.
133. Perra, O., Phillips, R., Fyfield, R., Waters, C., Hay, DF. (2015). Does mothers' postnatal depression influence the development of imitation? *Journal of Child Psychol Psychiatry*. 56(11),1231-8
134. Petterson, S. M., & Albers, A. B. (2001). Effects of poverty and maternal depression on early child development. Child Development, 72, 1794 – 1813.
135. Philipps, L. H., &O'Hara, M. W. (1991). Prospective study of postpartum depression: 4 1/2-year follow-ups of women and children. *Journal of Abnormal Psychology*, *100*, 151–155.
136. Pieterse, D., (2015). Childhood Maltreatment and Educational Outcomes: Evidence from South Africa. *Health Economics*, 24, 876–894 (2015)
137. Plant, D. T., Pariante, C. M., Sharp, D., & Pawlby, S. (2015). Maternal depression during pregnancy and offspring depression in adulthood: role of child maltreatment. *The British journal of psychiatry : the journal of mental science*, *207*(3), 213–220. doi:10.1192/bjp.bp.114.156620
138. Public Health Agency of Canada. (2014). Pregnancy and Women's Mental Health in Canada. Results From The Canadian Maternity Experiences Survey, No:140282. Retrieved

from http://healthycanadians.gc.ca/publications/healthy-living-vie-saine/pregnancy-mental-health-grossesse-sante-mentale/alt/pregnancy-mental-health-grossesse-sante-mentale-eng.pdf

139. Quach. J., and Barnett, T. (2015). Impact of chronic illness timing and persistence at school entry on child and parent outcomes: Australian Longitudinal Study. *Academic Pediatric*. 15(1), 89-95.

140. Quarini, C., Pearson, R.M., Stein, A., Ramchandani, P.G., Lewis, G., Evans, J . (2016). Are female children more vulnerable to the long-term effects of maternal depression during pregnancy?. *Journal of Affective Disorder*. 189, 329-35.

141. Raposa, E., Hammen, C., Brennan, P., & Najman, J. (2014). The long-term effects of maternal depression: early childhood physical health as a pathway to offspring depression. *Journal of Adolescent Health*, 54(1), 88–93.

142. Robinson M; Zubrick SR; Pennell CE; Van Lieshout RJ; Jacoby P; Beilin LJ; Mori TA; Stanley FJ; Newnham JP; Oddy WH. (2013). Pre-pregnancy maternal overweight and obesity increase the risk for affective disorders in offspring. *Journal of Developmental Origins of Health and Disease.* 4(1), 42-48.

143. Romano, E., Babchishin, L., Marquis, R., & Frechette, S. (2015). Childhood maltreatment and educational outcomes. Trauma, Violence & Abuse, 16(4), 418–437. http://dx.doi.org/10.1177/1524838014537908.

144. Romano, E., Babchishin, L., Pagani, L.S., & Kohen, D. (2010). School readiness and later achievement: replication and extension using a nationwide Canadian survey. *Developmental Psychology*, 46(5):995–1007. http://dx.doi.org/10.1037/a0018880

145. Romeo, D.M., Di Stefano, A., Conversano, M., Ricci, D., Mazzone, D., Romeo, M.G., & Mercuri, E. (2010). Neurodevelopmental outcome at 12 and 18 months in late preterm infants, *European Journal of Paediatric Neurology*, 14, 503-507.

146. Roos, L.L., Hiebert, B., Manivong, P., Edgerton, J., Walld, R., MacWilliam, L., de Rocquigny, J. (2013). What is most important: Social factors, health selection, and adolescent educational achievement. *Social Indicators Research*, 110(1), 385-414.

147. Santos, R., Brownell, M., Ekuma, O., Mayer, T., Soodeen R. (2012). The Early Development Instrument (EDI) in Manitoba: Linking Socioeconomic Adversity and Biological Vulnerability at Birth to Children's Outcomes at Age 5. Winnipeg, MB: Manitoba Centre for Health Policy. http://mchp-appserv.cpe.umanitoba.ca/deliverablesList.html

148. Sharp, D., Hay, DF., Pawlby, S., Schmücker, G., Allen, H., Kumar, R. (1995). The impact of postnatal depression on boys' intellectual development. *Journal of Child Psychology and Psychiatry*. 36, 1315–36.

149. Shonkoff, JP., Phillips, DA. (2000). National Research Council (US) and Institute of Medicine (US) Committee on Integrating the Science of Early Childhood Development; Shonkoff JP, Phillips DA, editors. From Neurons to Neighborhoods: The Science of Early Childhood Development. Washington (DC): National Academies Press (US); 2000. Available from: https://www.ncbi.nlm.nih.gov/books/NBK225557/ DOI: 10.17226/9824

150. Snow, K.L. (2006). Measuring School Readiness: Conceptual and Practical Considerations. *Early Education and Development*, 17(1), 7-41. DOI: 10.1207/s15566935eed1701_2.

151. Sohr-Preston, S.L. & Scaramella, L.V. (2006). Implications of Timing of Maternal Depressive Symptoms for Early Cognitive and Language Development. *Clinical Child and Family Psychology Review, 9(1), 65-83.* DOI: 10.1007/s10567-006-0004-2

152. Spieker, S. J., Nelson, D. C., Petras, A., Jolley, S. N., & Barnard, K. E. (2003). Joint influence of childcare and infant attachment security for cognitive and language outcomes of low-income toddlers. *Infant Behavior and Development*, 26(3), 326–344. http://dx.doi.org/10.1016/S0163-6383(03)00034-1.

153. Stein, A., Malmberg, LE., Sylva, K., Barnes, J., Leach, P; FCCC team. (2008). The influence of maternal depression, caregiving, and socioeconomic status in the post-natal year on children's language development. *Child Care Health Dev*. 34(5), 603-612.

154. Talge, N. M., Neal, C., Glover, V., & the Early Stress, Translational Research and Prevention Science Network: Fetal and Neonatal Experience on Child and Adolescent Mental Health. (2007). Antenatal maternal stress and long-term effects on child neurodevelopment: How and why? *Journal of Child Psychology and Psychiatry*, 48, 245–261.

155. Turney, K. (2010). Chronic and Proximate Depression among Mothers: Implications for Child Well-being, *Fragile Families Working Paper:10-07-FF, Robert Wood Johnson Health & Society Scholar. The University of Michigan.*

156. Uematsu A, Matsui M, Tanaka C, Takahashi T, Noguchi K, Suzuki M, Nishijo H (2012). Developmental trajectories of amygdala and hippocampus from infancy to early adulthood in healthy individuals". PLOS ONE. 7 (10): e46970. doi:10.1371/journal.pone.0046970. PMC 3467280. PMID 23056545.

157. Wadhwa, P. (2005). Psychoneuroendocrine processes in human pregnancy influence fetal development and health. *Psychoneuroendocrinology,* 30(8), 724–743.

158. Wang, Y., Dix, T. (2015). Mothers' early depressive symptoms predict children's low social competence in first grade: mediation by children's social cognition. *Journal of Child Psychol and Psychiatry*. 56(2), 183-192. DOI:10.1111/jcpp.12297.

159. Weissman, M.M., Wickramaratne, P., Nomura, Y., Warner, V., Pilowsky, D., Verdeli., H. (2006). Offspring of depressed parents: 20 years later. *American Journal of Psychiatry,*163(6), 1001-1008.

160. West, K.K., Mathews, B.L., & Kerns, K.A. (2013). Mother-child attachment and cognitive performance in middle childhood: An examination of mediating mechanisms. *Early Childhood Research Quarterly,* 28, 259-270.

161. Williams, J. M. G., Watts, F. N., MacLeod, C., & Mathews, A. (1988). The Wiley series in clinical psychology. Cognitive psychology and emotional disorders. Oxford, England: John Wiley & Sons.

162. Winters, K. C., Stinchfield, R. D., Latimer, W. W., & Stone, A. (2008). Internalizing and externalizing behaviors and their association with the treatment of adolescents with substance use disorder. Journal of substance abuse treatment, 35(3), 269–278. DOI:10.1016/j.jsat.2007.11.002.

163. Wszołek, K., Żak, E., Żurawska, J., Olszewska, J., Pięta, B., Bojar, I. (2018). Influence of socio-economic factors on emotional changes during the postnatal period. Annals of Agricultural and Environmental Medicine, 25(1), 41-45. https://doi.org/10.26444/aaem/74486.

Chapter 3

The Impact of Childhood-Onset Type I Diabetes on Educational Attainment and Labor Market Outcomes in Early Adulthood

Abstract

This study examines the impact of Type I Diabetes Mellitus (TIDM) during childhood on educational attainment and labor market outcomes in adulthood using the last ten waves of the National Health Interview Survey (NHIS). Individuals are identified as Type I diabetics during childhood if they were diagnosed for the first time with diabetes at age less than 15 years and have been taking insulin. The logistic regression model is used for the binary outcome variables and the Tobit model is used for continuous outcome variables to avoid the selection bias that may arise from the zero working hours and earnings of unemployed individuals. The results show that individuals who developed Type I diabetes early in life are 7 to 17 percentage points less likely to be employed, work fewer hours (3 to 11 hours less per week) and are 5 to 10 percentage points more likely to receive social welfare assistance than non-diabetic individuals. In addition, Type I diabetics experience less educational attainment than non-diabetics. They are 3 to 7 percentage points more likely to drop out of high school and 5 to 9 percentage points less likely to get a university degree. I also found lower wages for type I diabetic persons. They can conservatively expect to lose more than $3,050 annually compared to their peers without Type I diabetes. The results show that there is socioeconomic heterogeneity in the impact of Type I diabetes on educational attainment and labor market outcomes. Individuals of parents with less than high school and who are in a low-income group have the worst educational attainment and labor market outcomes in adulthood. The socioeconomic status of the family has a positive impact on the long-run consequences of Type I diabetes, mitigating the negative effects of diabetes on their children.

Chapter 3

The Impact of Childhood-Onset Type I Diabetes on Educational Attainment and Labor Market Outcomes in Early Adulthood

3.1. Introduction

A growing body of literature in Economics and Medicine argues that childhood health has a vital impact on outcomes in adulthood. Empirical studies show that health and socioeconomic status are highly correlated, regardless of how health status is measured (Deaton, 2003; Gerdtham and Johannesson, 2000; Gerdtham et al., 2004). Moreover, some studies assume that the causal relationship between health and socioeconomic status runs in both directions. Kuh and Wadsworth (1993) argue that health problems early in life affect the stock of health a lifetime, which negatively impacts socioeconomic outcomes such as education and labor market outcomes.

Studies that investigate the impact of childhood health on outcomes in adulthood link different measures of early life health to various outcome measures, such as earnings and education (Currie, 2009, Almond and Currie, 2011). Measures of early life health include birthweight (Figlio et al., 2013; Almond and Currie, 2011; Black et al., 2007; Behrman and Rosenzweig, 2004), overall health status (Hass et al., 2012; Smith, 2009; Smith, 2007), and height (Lundborg et al., 2014; Case and Paxson, 2008). For example, Black et al., 2005 found that a 10 percent increase in the birth weight of Norwegian twins was associated with a 1 percent increase in total earnings. Similarly, Hass (2006) found that poor childhood health was associated with a 22 percent decrease in earnings. Johnson and Schoeni (2007) found that low

birth weight children are 5 percent less likely to be employed, have lower earnings by 10 percent than the healthy group at age 25, and 15 percent lower earnings at age 35. Currie et al (2010), using data from public health insurance records for 50,000 children born in Manitoba, Canada, also found a negative effect on school grades and on the probability of being on social assistance. There are also studies that assess the impact of different health issues during childhood. For instance, Lundborg et al (2014) assess the impact of mental health conditions, diabetes, asthma, and injury on labor market outcomes among young adults. They found that individuals with mental health conditions earn on average 20% less and diabetes appears to be almost as severe as these mental problems.

A growing literature suggests that early heath is important for success in the labor market in adulthood. This relationship investigated by linking various measures of early childhood health to various outcomes in adulthood, such as earning and educational attainment (Currie, 2009, Almond and Currie, 2011). However, the impact of Type I diabetes, as an early health shock, on the labor market participation and probability of receiving social assistance is not clear. Also, little is known about its impact on educational attainment. Most of the previous literature on the socioeconomic consequences of diabetes does not distinguish between Type I diabetes (T1DM) and Type II diabetes (T2DM), which are very different in many aspects, especially the timing of onset (Fletcher and Richards, 2012; Hass et al., 2011; Brown et al., 2011; Minor, 2011; Case et al., 2005).

This study aims to fill this gap by estimating the impact of the health shock of a T1DM diagnosis in childhood on socioeconomic status in adulthood, specifically on educational attainment and labor market outcomes in early adulthood. This paper contributes to the literature on the relationship between early childhood health shock and socioeconomic status

(SES) in early adulthood by exploring labor market consequences (such as labor force participation, average hours worked per week, earnings, and the probability of receiving social assistance) and educational attainment (such as the probability of dropping out of high school, and the probability of having a university degree) of early childhood onset Type I diabetes using a unique dataset from the National Health Interview Survey (NHIS) that allows for data from more than one year to be pooled to increase the sample size for analytic purposes.

Specifically, the study addresses the following research questions:

- *What is the impact of having Type I diabetes during childhood on labor force participation?*
- *Are individuals with Type I diabetes more likely to receive social assistance?*
- *What is the impact on hours worked and earnings among individuals with T1DM who participate in the labor force?*
- *What is the impact of T1DM on the probability of dropping out of high school and having a university degree?*
- *Is there socioeconomic heterogeneity in the impact of Type I diabetes on the labor market outcomes and educational attainment?*

The remainder of this chapter is organized as follows. Sections 2 and 3 present an overview of diabetes, and a review of the relevant economics literature; section 4 introduces the dynamic capabilities formation model to show how diabetes can affect adulthood outcomes; section 5 describes the data and presents summary statistics for the sample; section 6 provides the empirical econometric specification and methodology; sections 7 and 8 present the empirical results and conclusion.

3.2. Diabetes Background

Diabetes is the seventh leading cause of death in the U.S., and approximately 29.1 million individuals or 9.3% of the U.S. population have diabetes (American Diabetes Association, 2014). In 2011, about 282,000 emergency room visits for adults aged 18 years or older listed hypoglycemia as the main diagnosis and diabetes as a secondary diagnosis, and about 175,000 emergency room visits were by people of all ages who had a hyperglycemic crisis. Type 1 diabetes mellitus (T1DM) is one of the most common chronic health problems in school-aged children. According to the U.S. Centers for Disease Control and Prevention (CDCP), about 208,000 people younger than 20 years have been diagnosed with diabetes. During the period 2008- 2009, an estimated 18,436 people younger than 20 years were newly diagnosed with type 1 diabetes and 5,089 people with type 2 diabetes (NDSR, 2014).

Narayan et al (2003) estimated the lifetime risk of any diabetes mellitus diagnosis for an individual born in the U.S. in 2000 to be one out of three for males and two out of five for females over a lifetime. The incidence rate of T1DM in 2002-2003 peaked in age groups 5-9 years and 10-14 years. The incidence rates per 100,000 persons by age group are as follows: 0-4 years, 14.3; 5-9 years, 22.1; 10-14 years 25.9; 15-19 years, 13.1. Although disease onset can occur at any age, the peak age for diagnosis is in the mid-teens, between 10 and 14 years (NDSR, 2014).

Type I diabetes mellitus (T1DM) also known as “juvenile diabetes” or “insulin-dependent diabetes” is an autoimmune disorder in which the beta cells that produce the insulin hormone in the pancreas are destroyed and not able to produce the insulin to control the amount of glucose in the body (Daneman, 2006). People with type I diabetes produce no insulin, so glucose cannot get into the body’s cells for use as energy. People with type 1 diabetes must use

insulin injections or pump infusions to control their blood glucose levels. The causes of type I diabetes are not fully understood, it has not been possible to identify a single underlying cause of T1DM, and there is currently no cure (Dahlquist 1998; Dahlquist et al., 1999; Daneman 2006; Akerblom et al. 2002). However, the exogenous injected insulin is not able to simulate the natural functions of the pancreas in producing insulin. Individuals with type 1 diabetes face some fluctuations in the bloodstream glucose level throughout the day depending on their diet, physical exercise, amount of insulin injected during the day (Homes et al, 1999). For instance, hypoglycemia (low blood glucose) in most cases occurs when there is too much insulin and not enough glucose in the body. Warning signs for hypoglycemia include dizziness, weakness, lightheadedness, sweaty/clammy skin, irritability, unusual behavior and lack of concentration. In contrast, hyperglycemia (high blood glucose) occurs when the body does not have enough insulin or cannot use the insulin it does have. Both hypo and hyperglycemia can be dangerous.

The brain is one of the most metabolically active organs in the body. It uses approximately 20% to 23% of the body's total energy requirements, despite accounting for only 2% of the body's mass (Cunnane et al., 2011; Prins, 2008; Mosconi et al., 2008). Since 90% of the brain's metabolic fuel energy is provided by glucose and cerebral cells can only save glucose for few minutes (Prins, 2008), the acute fluctuations in blood glucose (e.g., hypoglycemia and hyperglycemia) can have negative physiological, neurological, and cognitive outcomes (Mooredian, 1988). For instance, acute hypoglycemia can cause a decrease in metabolic activity of the brain (neuroglycopenic manifestations), increasing activity of the autonomic nervous system (adrenergic response), and hormonal responses. Such dysfunctions can cause different physical, social, and psychological problems such as headache, coma, seizures, sweating, hunger, nervousness, irritability (Meo et al., 2013; Cheach and Amiel, 2012;

James et al, 2002). Also, children with T1DM experience mild to moderate cognitive ability and neuropsychological difficulties such as impaired memory, attention, motor skills when compared to children without diabetes (Bade-White and Obrzut, 2009; Homes et al., 1999; Kucera and Sullivan 2011; Northam et al., 2001; Ryan, 1988). Since academic performance in schools depends strongly on many neuropsychological functions and cognitive ability, children with type I diabetes may experience problems with their academic performance and educational attainment (Desrocher and Rovet, 2004; Meo et al., 2013). Many studies have shown that diabetes in children has a significant effect on the neurocognitive function of the brain and consequently low academic achievement (Gaudieri et al., 2008; Abozaid, 2014).

3.3. Literature Review

There is a growing literature that links early life health to future outcomes. Several studies in the U.S., Canada, and Europe show a link between low birth weight and lower educational attainment and labor market outcomes, even among siblings or twins. Lawlor et al (2006) used Scottish siblings born between1950 and 1956 and found that lower birth weight siblings have lower scores on a test of intelligence at age 7. Black et al (2007) used a sample of Norwegian twins and found that a 10 percentage point increase in birth weight is associated with a one percentage point increase in the likelihood of graduating from high school and a one percentage point increase in earnings. Oreopoulos et al. (2005) used data from the province of Manitoba and found that children whose weight fell into the range of 1,500 to 2,500 grams are 8 percentage points less likely to complete grade twelve by age 17 than siblings who weighed over 3,500 grams. Royer (2005) found that an increase in birth weight by 1,000 grams is associated with a gain of 0.16 years of education. Currie and Hyson (1999) used the 1958 British birth cohort and found that low birth weight children have lower test scores, educational attainment, wages,

and the probability of being employed by age 33. Case et al (2005) extended the Currie and Hyson (1999) study and found that children who developed chronic health conditions during early life have lower educational attainment, wages, and the probability of being employed in adulthood by age 42.

The impacts of early mental health problems in childhood have been increasingly investigated in economics, epidemiology, and psychology. Almost 1 in 5 children in the U.S. have some impairment from a mental or behavioral disorder; 11% have significant functional impairments; 5% suffer extreme functional impairment. Currie and Madrian (1999) found that mental health problems are one of the leading causes of days missed at work. Farmer (1993, 1995) used data from the 1958 British birth cohort and found that children who are in the top decile of an aggregate behavior problems score at ages 7, 11 and 16 have lower educational attainment, earnings, and the probability of being employed at age 23. Gregg and Machin (1998) used the same data and found that behavioral problems at age 7 are associated with lower educational attainment at age 16, which in turn is associated with poor labor market outcomes at ages 22 and 23. Kessler et al. (1995) used data from the U.S. National Comorbidity Study and found that children who developed early-onset psychiatric problems were less likely to graduate from high school or attend college. Miech et al. (1999) used a cohort of New Zealand children and found that youths with hyperactivity and conduct disorders have less schooling, while anxiety and depression have little impact on schooling levels. Similarly, McLeod and Kaiser (2004), using data from the National Longitudinal Survey of Youth (NLSY), found that children who have behavior problems at ages 6 to 8 are less likely to graduate from high school or to attend college.

Health shocks in early life due to wars and other crises can have lasting effects on health (Doblhammer, 2004). For instance, Almond (2006) examines the direct impact of the influenza

epidemic on education and labor market outcomes of individuals affected by the disease in utero. He found that children of infected mothers were 15% less likely to have a high school degree, are more likely to receive social assistance, and the wages of affected men were 5 to 9 percentage point less than unaffected ones. Salm and Schunk (2011) used a sample of German children to examine the impact of childhood health on cognitive and verbal development at age six. They found that mental health problems capture a big portion of the variation in cognitive and verbal test scores.

The health economics literature on diabetes is based on two streams of research. The first focuses on estimating the direct medical cost of diabetes at both the macro (i.e., total expenditure) and the micro-level (i.e., health care providers and employers) (see: Daggan et al., 2006; Gilmer et al., 2005; Olivia et al., 2004; Brown et al., 1999; Selby et al., 1997). The second stream focuses on estimating the socioeconomic consequences of diabetes on individuals who experience diabetes.

Some studies find that individuals who were diagnosed with Type I diabetes during their childhood are less likely to be employed than non-diabetic individuals (Songer et al., 1989; Masato et al., 1993; Ingberg et al., 1996; Matsushimal et al 1993). For instance, Ingberg et al (1996) compared young adults who experience a childhood-onset type I diabetes to a control group of healthy young adults with matching characteristics on sex and age. They found that the probability of employment was lower among people with diabetes (71% versus 85%, $p<0.05$). They also found that diabetic individuals are more likely to depend on social welfare benefits than individuals without diabetes(15% versus 3%, $p<0.01$).

To explain the low rate of employability of individuals with diabetes, Songer et al (1989) used data of children with insulin-dependent diabetes mellitus diagnosed during the period 1950

– 1964 from the patient registry of the Children's Hospital of Pittsburgh. Using a case-control design they found that children diagnosed at age less than 17 years with insulin-dependent diabetes were more likely to report job-refusal than their non-diabetic siblings (56% versus 42%, $p<0.02$). Moreover, individuals who told the job interviewers in advance about their diabetes were more likely to report job refusal than their siblings (64% versus 42%, $p<0.005$) whereas respondents who did not tell the job interviewers about their diabetes experienced a job refusal rate similar to their siblings (44% versus 41%). They also found that insulin-dependent diabetes respondents are more likely to report a work disability relative to their siblings (32.4% versus 4.6%, $p<0.001$). Respondents with insulin-dependent diabetes report that diabetic complication is the main factor behind work disability and the most common disabling conditions were kidney disorder, blindness, severe retinopathy, heart disorder, and eye disorders. Along the same lines, Matsushimal et al. (1993) conducted a case-control study on diabetic patients in Japan whose age at onset was 19.5 years. Using sex and age-matched siblings as a control group, they found that patients with diabetes are more likely to be refused a job in their lifetime than their sibling controls (20% versus 0%). Moreover, they found that the majority of diabetic patients who were refused a job told the interviewers in advance about their diabetes.

Steen et al. (2010) estimated the long term detrimental consequence of Type I diabetes on annual earnings, using the national Diabetes Incidence Study in Sweden (DISS) database, which is a register of persons with diabetes onset between the ages 14-34 year since 1983. They compared the progression of annual earnings of all persons registered with Type I diabetes onset during 1983-2005 (n=3650 cases) with a control group (n=14629). They found that there is no difference in median earnings between individuals before their diabetes onset and the control group. In contrast, the annual earnings of individuals after the onset of diabetes gradually

decreased below that of the control group. That is, their annual earnings were less than the control group in each year from 1995-2005. Moreover, the onset of more than 10 years ago was associated with a decline in annual earnings by 4.2% for men and 8.1% for women.

Overall, this growing literature establishes that early-life health has a vital impact on outcomes in adulthood. However, prior research has not yet investigated the effect of specific diagnoses early in life, such as Type I diabetes in childhood, on labor market outcomes and educational attainment in adulthood. This paper contributes to this literature by estimating the effect of type I diabetes in childhood on labor market outcomes and educational attainment, using a unique data set from NHIS.

3.4. Theoretical Framework

Literature has shown that health shock during childhood can have lasting effects on adulthood health and labor market outcomes. Theoretically, the dynamic of the capabilities formation model can be used to explain the impact of an early health shock on early adulthood and later periods through its impact on cognitive and non-cognitive abilities, and stock of mental and physical health. According to this model, developed by Heckman (2007), individuals are born with heterogeneous endowments of capabilities (skills & abilities)

$$\Omega_i = (w_1^c,\ w_1^N,\ w_1^H)$$

where w_1^c is a vector of cognitive abilities at birth (e.g., fluid intelligence, which reflects the rate at which people learn and is measured by IQ; crystallized intelligence, which reflects acquired knowledge and is captured by achievements test (Kautz et al., 2014), w_1^N is a vector of non-cognitive abilities at birth (e.g., motivation, perseverance, time preference, risk aversion, self-control, preference for leisure, patience, self-esteem, neuroticism), and w_1^H is a vector of mental and physical

health stock at birth. Assuming the child lives multiple periods $t \in \{1, \cdots \cdots, T\}$, the accumulated stock of a child's abilities at age (t) is

$$\Omega_t = f_t\left(\sum_t w_t^C,\ \sum_t w_t^N,\ \sum_t w_t^H\right)$$

The technology of skill (k) formation at any period (t) of a child's lifetime can be given by

$$\Omega_{k,t+1} = f_{k,t}\left(\Omega_{k,t},\ \theta_p, I_{k,t}\right)$$

where $\theta_p = (\theta_p^C, \theta_p^N, \theta_p^H)$ denotes parental capabilities (e.g., education, income, genes, IQ, psychological factors) and $I_{k,t}$ is parental and government investment in child skills at a period (t). The function $f_{k,t}(\cdot)$ is increasing in its argument and is a twice continuously differentiable function. Capability production exhibits self-productivity, where skill attainment at one stage in a child's life raises skill attainment at later stages in the life cycle and it is also Dynamic Complementarity, where skills produced at one stage increase productivity of investment in subsequent stages in a child's life cycle (early investment increases the productivity of later investment). In other words, skills beget skills and abilities beget abilities. By recursive backward substitution for Ωt, $\Omega t-1$, $\cdots \cdots$ in the skill formation function, the stock of a child's capability at the period $(t+1)$ can be expressed as a function of all past investments during the past period,

$$\Omega_{t+1} = g_t\left(\Omega_1,\ \theta_p,\ \mathrm{I}_1,\ \cdots\cdots,\ I_t\right)$$

Consequently, an adult's stock of capabilities (H) at the period $(T+1)$ can be expressed as a function of the initial endowment of capabilities, parental endowment of capabilities, and all past investments in capabilities during childhood,

$$\mathrm{H}_{T+1} = m_{T+1}\left(\Omega_1,\ \theta_p,\ \mathrm{I}_1,\ \cdots\cdots,\ I_t\right)$$

Adult's outcome, j, $at\ period$ $(T+1)$ is based on the accumulated skills and abilities during her/his childhood period,

$$Y_{j,T+1} = n_j(H_{T+1}), \qquad J \in \{1, \cdots\cdots, J\}$$

$$Y_{j,T+1} = \lambda_j \left(\underbrace{\sum_{t=1}^{T} \omega_t^C, \ \sum_{t=1}^{T} \omega_t^N, \ \sum_{t=1}^{T} \omega_t^H}_{\Omega_{j,T+1}}, \ e_{j,T+1} \right)$$

where (λ_j) is a multiplier that can take any positive value, and $(e_{j,T+1})$ is the level of the adult's effort at the period $(T+1)$. These abilities are used with different weights and importance to explain the differences in labor market outcomes: employment probability, the difference in income, worked hours, days missed.

It is assumed that there are two stages of childhood development, $s \in \{1,2\}$, followed by adulthood working lifetime. Let ε_t denote a shock during early childhood (Type I diabetes), which can affect the accumulation of capabilities during the rest of the lifetime. Then, the technology of producing capabilities at stage s is as follows:

$$\Omega_{k,t+1,s} = f_{k,s}\left(\Omega_{t,s}, \ I_{k,t,s}, \ \theta_p, \ \varepsilon_{k,t}\right),$$

For $k \in \{C,N\}$, $t \in \{1,2, \cdots\cdots, T\}$, and $s \in \{1,2\}$. Parental investment decisions in every period of childhood are now endogenous in the model and affected by the health shock (Type I diabetes) at age (t),

$$\mathrm{I}_{k,t,s} = n_{k,s}\left(\Omega_{k,t,s}, \ \theta_p, \ \varepsilon_{k,t}\right).$$

When a child is first diagnosed with Type I diabetes at any time $t \in \{1,\cdots\cdots,T\}$ in this two-stage model, her/his skills and abilities can be affected in the next period $s2$, and the final effect on the accumulated capabilities in adulthood will depend on two effects:

$$\frac{d\Omega_{k,s_2}}{d\varepsilon_{s_1}} = \underbrace{\frac{\partial \Omega_{s_2}}{\partial \varepsilon_{s_1}}}_{Medical\ effect} + \underbrace{\frac{\partial \Omega_{s_2}}{\partial I_{s_2}} \frac{\partial I_{s_2}}{\partial \varepsilon_{s_1}}}_{parental\ investment\ effect}$$

The first term of the right-hand side of the equation captures the biological (medical) effect of a Type I diabetes shock at the early life and is expected to be negative (i.e., reduce the capabilities of the child). The second term captures the effect of parental investment responses to the child's health shock in the first period when the child is diagnosed with Type I diabetes. The final effect on the child's capabilities and hence early adulthood labor market outcomes and educational attainment will depend on the net effect of biological and parental investments.

3.5. Data and Variables

This study uses the last ten waves (2004 – 2014) of the National Health Interview Survey (NHIS), which is conducted by the National Center for Health Statistics (NCHS) in the U.S. since 1957. The survey includes approximately 35,000 households containing about 87,500 persons. There are two sample restrictions for this study. First, the sample includes only individuals who are between 20 and 45 years, so that the impact of type I diabetes on labor market outcomes in early adulthood can be estimated. Second, the sample includes only individuals who are diagnosed with Type I diabetes during early childhood and excludes any individual who experiences diabetes during any point in time in adulthood. The final complete sample consists of 240,980 individuals with full information for the study variables.

Although NHIS is not a panel data that follows individuals over their life course and does not distinguish between Type I diabetes and Type II diabetes, it is one of the few surveys that includes specific questions about diabetes that can be used to identify Type I and type II. Specifically, this study uses three questions from the survey that cover different dimensions of the timing of diabetes: "Have you EVER been told by a doctor or health professional that you have diabetes or sugar diabetes*?"*, "*How old were you when a doctor FIRST told you that you had diabetes or sugar diabetes?", "Are you NOW taking insulin?"*. One of the most important

distinctions between Type I and Type II diabetes is the timing of the diagnosis. Type I diabetes, often referred to as juvenile or insulin-dependent diabetes, is diagnosed early in childhood. Therefore, for this study, individuals with type 1 diabetes are captured by a dummy variable (*Diabetes*) that takes the value "1" if an individual was diagnosed with diabetes before age 15 and he/she is taking insulin at the time of onset.

3.5.1. Outcome Variables

The individual's employment status is captured by a dummy variable *(employed)* that takes the value "1" if the respondent is working for pay and "0" if not working. "*Total annual earnings"* is total earnings from employment during the last year, which is measured on a discrete scale ranging from zero to eleven (Minor, 2011). Individuals who choose not to participate in the labor market are assigned "0", and the values from 1 to 11 are used as continuous variables for individuals who choose to participate in the labor market. On this scale, a value of "1" means earnings range from 1$ to $4999, and each successive value captures a $5,000 increment in the level of earnings up to level "5", then $10,000 increments for levels "6" to "10", and level 11 captures all with income above $75,000. "*Worked Hours"* are the average hours worked per week during the last year. "*receiving social assistance*" is a dummy variable that takes the value "1" if the respondent receives social assistance and "0" if not. Educational attainment is measured by two binary outcome measures. "*dropping out of high school*" which takes the value "1" if the respondent reported completing 9 to 12 years of schooling but had no credentials, or did not have a high school degree, and "0" otherwise. "*university degree*" which takes value "1" if the respondent has completed a university degree and "0" otherwise.

3.5.2. Control Variables

Control variables were added to capture heterogeneity effects, e.g. education, age, an individual's family income, marital status, and gender. Marital status is captured by a dummy variable which equals "1" if the individual is married or in a common-law relationship and "0" otherwise. The race has four categories: white (reference group), black, Asian, and others. Respondent's mother or father's educational attainment has four levels: less than high school (reference category), high school, some post-secondary degree (i.e., college or any certificate), and university degree. An individual's family income level is divided into three levels: low income (reference category), middle income, and high income. An individual's weight is captured by one of four levels: healthy weight (reference group), underweight, overweight, and obese. A complete list of other covariates, which include comorbidities and family health, are shown in Table 3.1.

Table 3.1: Summary Statistics and Variables Description

	Diabetic		Non-diabetic	
	Mean	St Dev	Mean	St.Dev
	(1)	(2)	(3)	(4)
Employed	0.7257	0.4472	0.841	0.366
On Social Assistance	0.1416	0.3494	0.038	0.191
Hours worked per week	27.153	21.589	36.703	16.231
Earnings	6.39	3.446	7.414	3.309
Full time work	0.114	0.323	0.255	0.436
Dropout out of high school	0.169	0.375	0.096	0.294
Has a university degree	0.195	0.396	0.315	0.464
Age (in years)	31.806	7.463	32.552	7.467
Female	0.565	0.497	0.52	0.5
Married	0.383	0.487	0.601	0.49
White	0.748	0.435	0.74	0.439
Black	0.159	0.366	0.142	0.349
Asian	0.028	0.167	0.081	0.273
Others races	0.065	0.247	0.037	0.19

Low-income family	0.529	0.501	0.314	0.464
Middle income family	0.289	0.455	0.3	0.458
High income family	0.174	0.38	0.366	0.482
Mother has no high school	0.202	0.402	0.171	0.377
Mother has a high school	0.206	0.405	0.217	0.412
Mother has some post secondary	0.367	0.483	0.326	0.469
Mother has a university degree	0.226	0.419	0.286	0.452
Functional limitation	0.403	0.492	0.079	0.27
Underweight	0.020	0.141	0.019	0.136
Healthy weight	0.270	0.445	0.368	0.482
Overweight	0.331	0.471	0.311	0.463
Obesity	0.379	0.486	0.302	0.459
Live with both parents	0.060	0.239	0.115	0.319
ADHD during childhood	0.085	0.279	0.006	0.075
Hypertension in early childhood	0.379	0.486	0.132	0.339
Cholesterol	0.278	0.449	0.109	0.312
Asthma in early childhood	0.024	0.154	0.003	0.054
Coronary Heart Diesese	0.020	0.141	0.005	0.073
Mother has Diabetes	0.367	0.483	0.005	0.068
Father has Diabetes	0.254	0.436	0.003	0.054
Work at private sector	0.815	0.389	0.802	0.398
Self-employee	0.050	0.218	0.057	0.231
Work at public sector	0.131	0.338	0.138	0.345
Work at family business	0.005	0.067	0.003	0.052
Live in Northeast region	0.141	0.349	0.155	0.362
Live in Midwest region	0.226	0.419	0.193	0.395
Live in South region	0.359	0.481	0.353	0.478
Live in West region	0.274	0.447	0.299	0.458
Age_group0 (20-25)	0.245	0.431	0.192	0.394
Age_group1 (26-30)	0.224	0.418	0.199	0.399
Age_group2 (31-35)	0.198	0.4	0.195	0.396
Age_group3 (36-40)	0.122	0.328	0.184	0.388

Table 3.1 reports the summary statistics for the working sample. Columns 3 and 4 show the mean and standard deviation for the treatment group (respondents with type I diabetes) and columns 5 and 6 show the mean and standard deviation of the control group (respondents without diabetes). Comparing the two groups, parents' diabetic history is significantly different between the treatment and control groups. About 37% of diabetic respondents report that their mothers have been diagnosed with diabetes, compared to less than 1% of non-diabetics. In addition, almost 25% of diabetic respondents report that their fathers have been diagnosed with diabetes compared to 0.3% of non-diabetics. This may reflect the effect of genetic factors.

Labour market outcomes among the type I diabetes population are different from the non-diabetic group. Around 72% of type I diabetes respondents have been employed during the last year, while almost 84% of non-diabetics have a job. This reflects lower labor force participation for people with Type I diabetes. Non-diabetics who are working reported working an average of 36 hours per week while type I diabetics reported 27 hours. Around 14 percent of type I diabetics report that they are on social assistance while around 4 percent of the non-diabetic control group are on social assistance.

The average annual earnings of non-diabetics are $37,000 compared to $34,000 for diabetics. On average, this would imply that individuals who develop Type I diabetes early in life can expect to lose more than $135,000 during their working lifetime. Individuals with type 1 diabetes report lower educational outcomes compared to non-diabetics. Almost 17% of individuals with type 1 diabetes do not complete their high school degree while 10% of non-diabetics drop out of high school. In addition, 19% of diabetics have a university degree compared with 31% of non- diabetics. Finally, the results show that 40% of individuals who

experienced Type I diabetes during childhood have a functional limitation compared with almost 8% for non-diabetics.

3.6. Empirical Methodology

The goal of this study is to estimate the impact of a Type I diabetes diagnosis during early childhood on the educational attainment and various labor market outcomes in early adulthood. The baseline model controls for basic demographic variables, which include age, sex, marital status, race, and region, and estimates the impact of type 1 diabetes on labor market outcomes and educational attainment. Different sets of covariates will subsequently be added to the baseline model to measure the accumulating effect of each (Fletcher and Richards, 2012). Three sets of control variables are used: *Family socioeconomic background controls* is an indicator of family socioeconomic status that includes family income, mother's and father's education, and whether the child is born into a family with both parents present; *Comorbidities controls* are indicators of child health, specifically the presence of certain chronic diseases during early childhood such as Attention Deficit Hypertension Disorder (ADHD), depression, cholesterol, obesity, asthma, and functional limitation. *Family health controls* are indicators of family health status, such as whether the mother and/or father have diabetes. I use two econometric models to estimate the effect of Type I diabetes on labor market outcomes and educational attainment. I use the logistic regression model for binary outcome variables and the Tobit model for continuous outcome variables.

A logistic regression model is used to empirically estimate the impact of Type I diabetes on the probability of being employed, the probability of receiving social assistance, the probability of dropping out of high school, and the probability of having a university degree. The conditional probability of $Y = 1$ is specified as

$$Pr(Y_i = 1 \,|X_i) \;\equiv\; \pi(X_i) = \frac{e^{\beta_0 + \sum \beta X}}{1 + \; e^{\beta_0 + \sum \beta X}} \qquad (1)$$

and log transformation gives the following model

$$ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \; \beta_0 + \sum \beta_i \, X \quad (2)$$

$$g(x)_i = \; \alpha_0 + \; \alpha_1 Diabetes_i + \; \alpha_2 Z_i \; + \; \alpha_3 X_i + \; \alpha_4 G_i + \; \varepsilon_i \quad (3)$$

where the dependent variable $g(x)_i$ (either labor market outcomes or educational attainment in adulthood) for the individual (i) is a function of time-invariant $Diabetes_i$ and other covariates. Labor market outcomes include the probability of being employed and the probability of receiving social welfare. Educational attainment is measured by the probability of dropping out of high school, and the probability of having a university degree. $Diabetes_i$ is a dummy variable that takes value "1" if respondent i has type I diabetes and "0" otherwise. Z_i is a vector of controls. X_i is a vector of individual characteristics such as age, age squared, gender, marital status, region, and race. G_i is a vector of dummy variables that represent the age categories: 20-25 (reference category), 26-30, 31-35, 36-40, 41-45.

Estimating the impact of type I diabetes on other labor market outcomes, such as working hours and earnings, will include only the individuals who participate in the labor market, which means that working hours and earnings are censored variables. Hence to avoid selection bias, a Tobit model is used to estimate the impact of type I diabetes on working hours and earnings:

$$y_i^* = \; \beta_0 + \; \beta_1 TypeI_i + \; \beta_2 Z_i \; + \; \beta_3 X_i + \; \beta_4 G_i + \; \varepsilon_i \quad (4)$$

$$y_i = \begin{cases} y_i^* & if \; y_i^* > 0 \\ 0 & if \; y_i^* \leq 0 \end{cases}$$

where $y_i^* \; and \; y_i$ are latent and observed, respectively, working hours and earnings. Maximum likelihood estimation is used to estimate the Tobit model, and the likelihood function consists of

two parts: the first part captures censored working hours and earnings observations and the second part captures the linear uncensored property of observed working hours and earnings. Hence, the likelihood function is given by

$$L = \prod_{y_i=0} \left[1 - \Phi\left(\frac{X_i'\beta}{\sigma}\right)\right] \prod_{y_i>0} \frac{1}{\sigma}\phi\left(\frac{y_{i-X_i'\beta}}{\sigma}\right) \quad (5)$$

Where $\Phi(\cdot)$ is the cumulative distribution function (CDF) of the standard normal distribution, and $\phi(\cdot)$ is the corresponding density function (pdf).

3.7. Estimated results

3.7.1. Labor force participation

Table 3.2 reports the odds ratio of the impact of a Type I diabetes diagnosis during early life on the probability of being employed during early adulthood. When controlling for basic demographic covariates, we found that individuals who experience early onset Type I diabetes are less likely to be employed relative to non-diabetics when they are adults. The odds of being employed are 60% lower for individuals who experienced type I diabetes than non-diabetics. Moreover, the odds ratio of diabetes increased from 0.40 to 0.47 after controlling for family socioeconomic status and to 0.63 and 0.74 after controlling for early comorbidities and family health during childhood, respectively. The increase in the odds ratio of diabetes by almost 85 percent, from 0.40 to 0.74, after adding family socioeconomic status, comorbidities, and family health status to the baseline model reflects the significant importance of these factors during early childhood on the likelihood of participating in the labor market in adulthood. For instance, controlling for family socioeconomic status, we can see that children who were born and raised in middle- and high-income families are 3 and 4 times more likely to be employed relative to those in a low-income family.

Table 3.2: The Odds Ration of Impact of Type I Diabetes on Employment in Adulthood

	Model (1)		Model (2)		Model (3)		Model (4)	
Type I diabetes	0.409***	(0.037)	0.473***	(0.065))	0.630***	(0.090)	0.745**	(0.108))
Age	1.134***	(0.003)	1.094***	(0.005)	1.117***	(0.005)	-	-
Age-squared	0.998***	(0.0001))	0.998***	(0.001)	0.998***	(0.001)	-	-
Female	0.647***	(0.010)	0.651***	(0.015)	0.689***	(0.017)	0.672***	(0.017)
Married	1.112***	(0.017)	0.652***	(0.017)	0.620***	(0.017)	0.626***	(0.017)
White (reference group)								
Black	0.741***	(0.016)	0.930***	(0.031)	0.908***	(0.031)	0.921***	(0.032)
Asian	1.161***	(0.041)	1.115**	(0.063)	1.031	(0.060)	1.046	(0.061)
Others	0.637***	(0.026)	0.786***	(0.049)	0.844***	(0.054)	0.847***	(0.055)
(West region) reference group)								
Northeast	1.195***	(0.028)	1.051	(0.038)	1.080**	(0.040)	1.076**	(0.040)
Midwest	1.210***	(0.027)	1.214***	(0.041)	1.273***	(0.044)	1.265***	(0.044)
South	0.957***	(0.019)	0.933***	(0.028)	0.959	(0.029)	0.955	(0.029)
Self employed	0.422***	(0.064)	0.346***	(0.083)	0.318***	(0.077)	0.314	(0.077)
Gov emp	0.221***	(0.033)	0.128***	(0.030)	0.122***	(0.029)	0.119	(0.029)
Private emp	0.188***	(0.028)	0.152***	(0.036)	0.142***	(0.034)	0.139***	(0.034)
Mother has less than high school (reference group)								
High school			1.397***	(0.049)	1.350***	(0.049)	1.318***	(0.048)
Some post-secondary			1.421***	(0.048)	1.402***	(0.049)	1.370***	(0.048)
University degree			1.820***	(0.069)	1.672***	(0.066)	1.668***	(0.066)
Low-income family (reference group)								
Middle income			3.112***	(0.091)	2.836***	(0.086)	2.866***	(0.087)
High income			4.860***	(0.168)	4.248***	(0.151)	4.213***	(0.150)
Live with both parents			1.313***	(0.028)	1.303***	(0.028)	1.340***	(0.029)
Health weight - BMI (reference group)								
Underweight					0.741***	(0.068)	0.724***	(0.067)
Overweight					1.304***	(0.039)	1.322***	(0.040)
Obesity					1.269***	(0.039)	1.314***	(0.041)
ADHD					0.429***	(0.029)	0.433***	(0.030)
Asthma					0.475***	(0.035)	0.470***	(0.035)
Coronary heart					0.704***	(0.038)	0.730***	(0.040)
Cholesterol					0.836***	(0.023)	0.843***	(0.023)
Hypertension					0.799***	(0.022)	0.810***	(0.022)
Function limitation					0.459***	(0.012)	0.460***	(0.012)
Mother diabetes							0.830***	(0.048)
Father diabetes							0.626***	(0.037)
Age group 20 - 25 (reference group)								
age_group1 (26-30)							0.631***	(0.035)
age_group2 (31-35)							0.590***	(0.031)
age_group3 (36-40)							0.635***	(0.033)
age_group4 (41-45)							0.756***	(0.038)

Notes: Standard errors in parentheses. **, **, *** indicate significance level at 10%, 5%, 1% respectively

Also, children whose mothers achieved post-secondary education or a university degree are almost respectively 1.3 and 1.6 times more likely to be employed, relative to children of the non-educated mother. Therefore, the socioeconomic status of the family has a positive impact on the likelihood of being employed and moderates the negative effect of diabetes, apparent from the increase in the value of the odds ratio. When controlling for comorbidities during early childhood, the results show that individuals who experienced Attention Deficit Hypertension Disorder (ADHD) during early childhood are less likely to be employed than those without ADHD. The odds of being employed are 57% lower for an individual with ADHD during childhood than those who did not experience ADHD. This result is consistent with other studies, which found that the impact of ADHD on labor market outcomes depends on the timing of the disorder (Fletcher, 2014 & Currie et al., 2010). Individuals who developed asthma during early childhood are 53 percentage points less likely to be employed as adults.

The estimated odds ratios of model 3 assume implicitly that the effect of TIDM in early life on labor market participation in adulthood is constant across age groups, family socioeconomic status, and gender. However, the impact of type I diabetes could be more or less pronounced depending on the child's family's educational attainment and income, age, and gender of the individuals with diabetes, and these could be important for policy implications to mitigate the negative impacts of diabetes on labor market participation. Therefore, to capture this heterogeneity, interaction terms are included. Estimated odds ratios are reported in table 3.3, and the first column shows the odds ratio of the interaction terms.

Table 3.3: Odds Ratios of the Heterogeneity effect of Type I diabetes

Comparisons	Labor market participation
	Odds Ratio
Diabetic group (Females versus Males)	0.694***
Female group (diabetics versus non-diabetics)	0.484***
Male group (diabetics versus non-diabetics)	0.569***
Diabetic group (Middle income versus Low income)	7.968***
Diabetic group (High income versus Low income)	10.318***
Low income group (diabetics versus non-diabetics)	0.320***
Middle income group (diabetics versus non-diabetics)	0.556***
High-income group (diabetics versus non-diabetics)	1.035***
Mother dropped out of high school (diabetics versus no-diabetics)	0.230***
Mother has a high school degree (diabetics versus non-diabetics)	0.356***
mother has some post-secondary certificates (diabetics versus non-diabetics)	0.622***
mother has a university degree (diabetics versus non-diabetics)	0.768***
Diabetic groups (high school degree versus have no high school degree)	1.860***
Diabetic group (some post-secondary degrees versus have no high school degree)	3.993***
diabetic group (university degree versus have no high school)	6.757***
mother has no high school degree and family in a low-income group (diabetic versus nondiabetics)	0.147***
mother has some post-secondary certificates and family in the middle-income group (diabetic versus nondiabetics)	0.550***
mother has a university degree and family in the high-income group (diabetics versus nondiabetics)	0.665***

Notes: *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Type I diabetic females are less likely to be employed relative to non-diabetic females, while diabetic males less likely to be employed relative to non-diabetic males. Odds of being employed are 52% lower for type I diabetic females than non-diabetic females, and 44% lower for diabetic males than non-diabetic males. In the diabetic subgroup, diabetic females are less likely to be employed relative to diabetic males.

The results also show that there is heterogeneity with respect to parent's educational attainment, i.e. individuals with Type I diabetes born to and raised by parents with less than high school are less likely to be employed relative to non-diabetic individuals in the same group. Odds of being employed are 77% lower for individuals whose parents have less than a high school degree than non-diabetic individuals. Odds of being employed for diabetic individuals of parents with high school, some post-secondary degrees (i.e college or certificate), or a university degree are, respectively, 65%, 48% and 24% lower relative to non-diabetics in each group. Within the subgroup of individuals with diabetes, those of parents with high school degrees are 2 times more likely to be employed relative to the diabetic individual whose parents have less than high school and diabetic individuals whose parents have a post-secondary degree or a university degree are 4 and 7 times more likely to be employed relative to diabetic individuals whose parents have less than high school.

Controlling for heterogeneity across family income, individuals from low income and middle-income families are less likely to be employed relative to non- diabetics in those same income groups. In contrast, individuals from high-income families are more likely to be employed relative to nondiabetic individuals, but the result is not significant. Within the diabetes subgroup, children with early-onset diabetes in middle and high-income families are 7 and 10 times more likely to be employed than Type I diabetics who grew up in low-income families.

Table 3.4: Average Marginal Effect of Type I Diabetes on Labor Market Participation and social assistance.

	Model (1)	Model (2)	Model (3)	Model (4)
Employed	-0.1740***	-0.1383***	-0.0981***	-0.0664***
	(0.01273)	(0.01231)	(0.0123)	(0.0134)
On Social Assistance	0.1020***	0.093***	0.06624***	0.0488***
	(0.0060)	(0.006)	(0.0061)	(0.0066)
Demographic	YES	YES	YES	YES
Family socioeconomic	NO	YES	YES	YES
Childhood health	NO	NO	YES	YES
Family health & age group	NO	NO	NO	YES

Notes: each model is adjusted for some variables that include (i) age, gender, race, marital status, geographic area; (ii) living with two parents, mother and father education, and family income; (iii) obesity, cholesterol, hypertension, depression, asthma, ADHD; and (v) father and mother diabetes, and age groups. the Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 3.4 shows the average marginal effect of Type I diabetes on the probability of being employed and on the probability of receiving income assistance. In the unadjusted model (1), individuals who experience type I diabetes in childhood are almost 17 percentage points less likely to be employed, compared with those who did not experience diabetes during early childhood. After controlling for family socioeconomic status, the effect of type I diabetes on the likelihood of being employed is attenuated. Individuals with type I diabetes are 13 percentage points less likely to be employed. After further controlling for childhood health and family health, individuals who developed type I diabetes are almost 7 percentage points less likely to be employed than non-diabetic individuals.

In summary, individuals who are diagnosed with Type I diabetes during childhood are less likely to be employed in adulthood and there is heterogeneity in gender and family socioeconomic status. Educational achievement of parents and their income group have a significant mitigating role.

3.7. 2. Social assistance

Table 3.5 results show that Type I diabetes in early childhood has a significant effect on receiving social assistance in early adulthood. Controlling for basic demographic variables, type I diabetic individuals are 3 times more likely to receive social assistance in early adulthood relative to non-diabetics. The odds ratio of receiving social assistance decreased from 3.23 to 2.46 after controlling for family socioeconomic status and to 1.66 and 1.44 after controlling for early comorbidities and family health during childhood, respectively. The decrease in the odds ratio of diabetes by almost 55 percent, from 3.23 to 1.44, after adding the family socioeconomic status, comorbidities, and family health status to the baseline model reflects the significant importance of these factors during early childhood in explaining the variation in receiving social welfare in early adulthood. The results also show that the odds of depending on social welfare in adulthood for individuals who experience ADHD or asthma during early childhood are 59% and 44% greater than individuals who did not experience ADHD or asthma likely to depend on social welfare in adulthood, respectively.

Controlling for heterogeneity across gender and family socioeconomic status, Table 3.6 shows that there is no substantial gender heterogeneity in the impact of diabetes on social assistance in adulthood. Diabetic females and males are almost 1.98 and 2.18 times more likely to receive social assistance relative to non-diabetic females and males respectively. Within the subgroup of individuals with diabetes, the odds of depending on social welfare in adulthood for type I diabetic females are 3% greater relative to diabetic males. Controlling for family income heterogeneity, the odds of depending on social welfare in adulthood for diabetic individuals in middle- and high-income families are 33% and 62% lower relative to diabetic individuals in low-income families.

Table 3.5: The Odds Ratio of the Impact of Type I Diabetes on Receiving Social Assistance

	Model (1)		Model (2)		Model (3)		Model (4)	
Type I Diabetes	3.230***	(0.363)	2.467***	(0.416)	1.662***	(0.296)	1.446**	-0.261
Age	1.080***	(0.005)	1.126***	(0.008)	1.081***	(0.008)	-	-
Female	1.380***	(0.038)	1.358***	(0.057)	1.248***	(0.054)	1.325***	(0.063)
Married	0.288***	(0.009)	0.526***	(0.026)	0.545***	(0.028)	0.553***	(0.028)
White (reference group)								
Black	2.364***	(0.075)	1.954***	(0.094)	2.041***	(0.103)	2.022***	(0.102)
Asian	0.858**	(0.060)	1.071	(0.117)	1.235*	(0.141)	1.223*	(0.140)
Others	2.300***	(0.137)	1.814***	(0.160)	1.699***	(0.154)	1.678***	(0.153)
West region (reference group)								
Northeast	1.046	(0.041)	1.190***	(0.070)	1.128***	(0.068)	1.131**	(0.069)
Midwest	0.751***	(0.030)	0.756***	(0.045)	0.694***	(0.043)	0.696***	(0.043)
South	0.679***	(0.024)	0.624***	(0.034)	0.578***	(0.032)	0.580***	(0.032)
Self employed	0.423***	(0.077)	0.525**	(0.146)	0.558**	(0.161)	0.559***	(0.162)
Gov employed	0.278***	(0.049)	0.519**	(0.143)	0.493***	(0.140)	0.489***	(0.140)
Private employed	0.532***	(0.093)	0.623*	(0.168)	0.617***	(0.172)	0.613*	(0.171)
Mother has Less than high school (reference group)								
High school			0.570***	(0.030)	0.589***	(0.032)	0.591***	(0.032)
Some post-secondary			0.523***	(0.026)	0.532***	(0.028)	0.533***	(0.028)
University			0.302	(0.023)	0.333***	(0.026)	0.329***	(0.026)
Low-income family (reference group)								
Middle income			0.179	(0.012)	0.209***	(0.015)	0.210***	(0.015)
High income			0.099	(0.010)	0.125***	(0.013)	0.127***	(0.013)
Live with both parents			0.878	(0.030)	0.899***	(0.031)	0.892***	(0.031)
Normal weight - BMI (reference group)								
Underweight					1.093	(0.162)	1.094	(0.162)
Overweight					0.845***	(0.047)	0.839	(0.047)
Obesity					0.937	(0.049)	0.914*	(0.048)
ADHD					1.578***	(0.136)	1.593***	(0.138)
Asthma					1.447***	(0.119)	1.442***	(0.119)
Coronary heart disease					1.402***	(0.106)	1.359***	(0.103)
Cholesterol					1.183***	(0.057)	1.145***	(0.056)
Hypertension					1.273***	(0.063)	1.245***	(0.062)
Function limitation					3.381***	(0.173)	3.352***	(0.171)
Mother diabetes							1.141*	(0.085)
Father diabetes							1.622***	(0.149)
Age group 20 - 25 (reference group)								
Age_group1 (26-30)							1.328***	(0.119)
Age_group2 (31-35)							0.964	(0.086)
Age_group3 (36-40)							0.976	(0.086)
Age_group4 (41-45)							0.853*	(0.073)

Notes: The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Table 3.6: Odds Ratios of the Heterogeneity effect of Type I diabetes

Comparisons	Social Assistance
	Odds Ratio
Diabetic group (Females versus Males)	1.033***
Female group (diabetics versus non-diabetics)	1.989***
Male group (diabetics versus non-diabetics)	2.188***
Diabetic group (Middle income versus Low income	0.383***
Diabetic group (High income versus Low income	0.778***
Low-income group (diabetics versus non-diabetics)	1.67***
Middle income group (diabetics versus non-diabetics)	1.2***
High-income group (diabetics versus non-diabetics)	1.07***
Mother does not have a high school (diabetics versus no-diabetics)	5.16***
Mother has a high school degree (Diabetics versus non-diabetics)	2.19***
mother has some post-secondary certificates (diabetics versus non-diabetics)	1.49***
mother has a university degree (diabetics versus non-diabetics)	1.19***
Diabetic groups (high school degree versus have no high school degree)	0.95***
Diabetic group (some post-secondary degree versus have no high school degree)	0.868***
diabetic group (university degree versus have no high school)	0.599***
mother has no high school degree and family in the low-income group (diabetic versus nondiabetics)	5.34***
mother has some post-secondary certificates and family in the middle-income group (diabetic versus nondiabetics)	1.32***
mother has a university degree and family in high-income group (diabetics versus nondiabetics)	1.09***

Notes: *, **, *** indicate significance level at 10%, 5%, 1% respectively.

Type I diabetic individuals' parental educational attainment has a heterogeneous effect on the probability of receiving social assistance in adulthood. The odds of depending on social welfare in adulthood for diabetic individuals whose parent has completed high school are 5% lower relative to diabetic children whose parents did not complete high school. Similarly, the odds of depending on social welfare in adulthood for diabetic individuals whose parents have some post-secondary degree or university degree are 14% and 44% lower relative to diabetic individuals whose parents did not complete high school.

For the probability of receiving social assistance in adulthood, the unadjusted model 1 in table 3.4 indicates that individuals experience Type I diabetes in childhood are almost 10 percentage points more likely to receive social assistance, compared with those who did not experience diabetes during early childhood. After controlling for family socioeconomic status, the effect of Type I diabetes on the likelihood of being receiving social assistance is attenuated. Individuals with Type I diabetes are 9 percentage points more likely to depend on social assistance. After further controlling for childhood health and family health, individuals who developed type I diabetes are almost 5 percentage points more likely to be employed than non-diabetic individuals.

3.7.3 Earnings

Table 3.7 reports the impact of Type I diabetes during early childhood on earnings in adulthood. Controlling for basic socioeconomic factors, the initial estimate shows a loss of $15,000 per year by having Type I diabetes in childhood for individuals who participate in the labor force. Controlling for family socioeconomic status reduces the coefficient on diabetes to 1.799, which reflects a loss of about $8,950 per year in earnings and this may be consistent with the intergenerational transmission of socioeconomic status. That is, family education and income

during early life have a significant impact on adults' labor market earnings. Moreover, parental income and education moderate the negative effect of diabetes on earnings.

Table 3.7: The Impact of Type I Diabetes on Earnings (Tobit Model)

	Model (1)		Model (2)		Model (3)		Model (4)	
Type I diabetes	-3.195***	(0.177)	-1.799***	(0.202)	-0.997***	(0.253)	-0.610**	(0.257)
Age	0.671***	(0.003)	0.530***	(0.004)	0.562***	(0.007)	-	-
Age squared	-0.008***	(0.000)	-0.006***	(0.000)	-0.007***	(0.000)	-	-
Female	-1.585***	(0.018)	-1.533***	(0.020)	-1.596***	(0.035)	-1.618***	(0.036)
Married	1.295***	(0.020)	-0.268***	(0.024)	-1.481***	(0.040)	-1.479***	(0.040)
Races (white is a reference group)								
Black	-0.801***	(0.027)	-0.119***	(0.032)	-0.239***	(0.052)	-0.223***	(0.052)
Asian	0.126***	(0.036)	-0.204***	(0.042)	-0.451***	(0.076)	-0.429***	(0.075)
Others	-0.877***	(0.051)	-0.276***	(0.058)	-0.318***	(0.099)	-0.309***	(0.099)
Mother education (less than high school is reference group)								
High school			0.700***	(0.035)	0.931***	(0.060)	0.895***	(0.060)
Some post-secondary			0.964***	(0.033)	1.276***	(0.056)	1.242***	(0.056)
University degree			1.971***	(0.035)	2.177***	(0.060)	2.146***	(0.060)
Family income (low income is a reference group)								
Middle income			2.877***	(0.028)	3.533***	(0.045)	3.536***	(0.045)
High income			4.851***	(0.029)	5.698***	(0.051)	5.706***	(0.051)
Live with both parents			0.529***	(0.017)	1.165***	(0.035)	1.178***	(0.035)
Individuals BMI (healthy weight is a reference group)								
Underweight					-0.862***	(0.144)	-0.868***	(0.144)
Overweight					0.432***	(0.044)	0.445***	(0.044)
Obesity					0.288***	(0.045)	0.337***	(0.045)
ADHD					-2.788***	(0.135)	-2.774***	(0.135)
Asthma					-2.457***	(0.149)	-2.462***	(0.149)
Coronary heart disease					-1.035***	(0.102)	-0.942***	(0.103)
Cholesterol					-0.301***	(0.044)	-0.274***	(0.044)
Hypertension					-0.471***	(0.044)	-0.440***	(0.044)
Mother diabetes							-0.502***	(0.105)
Father diabetes							-0.939***	(0.107)
Age groups (reference group 20 – 25)								
Age_group1 (26-30)							0.147**	(0.067)
Age_group2 (31-35)							-0.334***	(0.065)
Age_group3 (36-40)							-0.665***	(0.066)
Age_group4 (41-45)							-0.670***	(0.066)

Notes: The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1%

Controlling for comorbidities reduces the coefficient of Type I diabetes on earnings to 0.997 less than non-diabetics. The decline in the predictive power of diabetes to 0.997 once comorbidities are added to the model shows that chronic health problems during childhood, especially attention deficit hypertension disorder (ADHD) and asthma, may capture part of the negative effect of diabetes on the labor market earnings. Individuals who developed ADHD early in life earn, on average, $13,900 less than individuals without ADHD, and individuals who develop asthma earn, on average, $12,250 less than individuals without asthma.

Controlling for family health reduces the coefficient of diabetes to 0.610, which means that the earnings of individuals with type I diabetes are $3,050 less annually than non-diabetics' earnings. The decline in the predictive power of diabetes when family health is introduced into the model shows that parents' diabetes and obesity capture some of the negative impacts on labor market earnings. That is, parents' health and whether or not they have diabetes have significant impacts on their children's long-run earnings. This may reflect the effect of biological and environmental factors during childhood on labor market earnings in adulthood. The results show that the loss in earnings is higher for older age groups. Individuals' earnings in age category 31-35, 36-40, 41–45 are, respectively, $1,600, $3,300, $3,350 less than individuals' earnings in age category 20 – 25.

3.7.4. Hours Worked

Table 3.8 reports the effect of an early life Type I diabetes diagnosis on the average hours worked. Change in hours worked is statistically significantly associated with type I diabetes. Individuals with type I diabetes report less working hours (3 to 11 hours) compared to non-diabetics. This may be due to complications of diabetes and the inability to work a full day. On

the other hand, diabetic individuals may have less capacity or preference to work than healthy individuals.

Controlling for basic demographic factors, we can see that individuals who experienced Type I diabetes in childhood work 11 hours less per week than their peers who were not diagnosed with type I diabetes during childhood. The difference in hours worked declined by another 24 percent after controlling for family socioeconomic status. It further decreases by 37 percent when comorbidities during childhood were included as controls, and another 46 percent after controlling for family health. The reductions in hours worked after controlling for parents' socioeconomic and health status reflect the significant importance of these factors in estimating the effect of diabetes on working hours for individuals who choose to participate in the labor market. The results also show that individuals who were diagnosed with asthma or ADHD during childhood report the lowest hours worked, 12 and 11 hours less.

Overall, the estimated results show that type I diabetes may reduce the likelihood of being employed, average hours worked, total income earnings, and increase the probability of social welfare dependency. These results can be explained by different factors. Type I diabetic complications, especially acute hypoglycemia or hyperglycemia may prevent individuals from working a full day or may increase absenteeism (Julius et al., 1993). Individuals with type I diabetes may also experience discrimination. For instance, employers may discriminate against individuals with diabetes, who may be at risk of hypoglycemia, and not allow them to work on jobs that require high levels of safety and concentration (Matsuuhima et al., 1993; Songer et al., 1989; Kraut et al., 2001).

Table 3.8: The Impact of Type I Diabetes on Worked Hours (Tobit Model)

	Model (1)		Model (2)		Model (3)		Model (4)	
Type I diabetes	-11.180***	(1.121)	-8.486***	(1.526)	-5.287***	(1.500)	-2.864**	(1.521)
Age	3.057***	(0.030)	2.418***	(0.041)	2.452***	(0.041)	-	-
Female	-7.819***	(0.153)	-7.338***	(0.209)	-6.530***	(0.209)	-	-
Married	0.886***	(0.157)	-5.076***	(0.237)	-5.212***	(0.232)	-4.994***	(0.232)
Races (white reference)								
Black	-3.889***	(0.227)	-1.043***	(0.320)	-1.305***	(0.315)	-1.140***	(0.315)
Asian	0.581*	(0.325)	-0.429	(0.460)	-0.801*	(0.452)	-0.562	(0.452)
Others	-5.266***	(0.448)	-2.301***	(0.599)	-1.481***	(0.587)	-1.455**	(0.586)
West region reference								
Northeast	1.791***	(0.240)	0.195	(0.330)	0.360	(0.323)	0.287	(0.323)
Midwest	2.335***	(0.222)	1.888***	(0.299)	2.129***	(0.293)	2.033***	(0.293)
South	0.512***	(0.200)	0.290	(0.272)	0.456*	(0.267)	0.412	(0.266)
Self employed	-3.472***	(1.264)	-4.253**	(1.776)	-4.951***	(1.737)	-5.092***	(1.733)
Gov employed	-8.307***	(1.250)	-11.87***	(1.757)	-11.85***	(1.719)	-12.08***	(1.715)
Private employed	-9.307***	(1.240)	-9.652***	(1.743)	-9.816***	(1.705)	-10.044***	(1.701)
Mother has less than HS								
High school			4.066***	(0.360)	3.563***	(0.353)	3.211***	(0.353)
Some post-secondary			3.932***	(0.339)	3.561***	(0.333)	3.236***	(0.333)
University degree			5.876***	(0.363)	4.999***	(0.357)	4.943***	(0.357)
Family income (low-income reference)								
Middle income			14.204***	(0.272)	12.821***	(0.267)	12.962***	(0.267)
High income			18.714***	(0.305)	16.834***	(0.300)	16.860***	(0.300)
Live with both parents			3.376***	(0.205)	3.169***	(0.200)	3.523***	(0.201)
Healthy weight reference								
Underweight					-2.623***	(0.859)	-2.837***	(0.859)
Overweight					2.645***	(0.257)	2.817***	(0.257)
Obesity					3.071***	(0.269)	3.473***	(0.269)
ADHD					-12.468***	(0.801)	-12.446***	(0.800)
Asthma					-11.593***	(0.860)	-11.663***	(0.862)
Coronary health					-6.007***	(0.602)	-5.360***	(0.607)
Cholesterol					-1.825***	(0.259)	-1.762***	(0.260)
Hypertension					-2.233***	(0.259)	-2.058***	(0.260)
Function limitation					-8.308***	(0.248)	-8.256***	(0.248)
Mother diabetes							-3.468***	(0.623)
Father diabetes							-5.671***	(0.634)
Age group 20-25 reference								
Age_Group1 (26-30)							-3.057***	(0.400)
Age_Group2 (31-35)							-5.787***	(0.383)
Age_Group3 (36-40)							-6.208***	(0.388)
Age_Group4 (41-45)							-5.413***	(0.383)

Notes: The Standard errors are in parentheses *, **, *** indicate significance level at 10%, 5%, 1% respectively.

3.7.5. High School Dropout

Table 3.9 reports the results of the effect of Type I diabetes on the probability of dropping out of high school. Individuals who experienced Type I diabetes during childhood are at high risk of dropping out of high school. Controlling for basic demographic factors, Type I diabetics are more likely to drop out of high school than non-diabetics. The odds of dropping out the high school for Type I diabetics are 67% greater than non-diabetics.

Controlling for family socioeconomic status and health reduced the odds ratio by 60 percent, which reflects the importance of these factors in mitigating the negative impacts of diabetes on school attendance. The odds of dropping out the high school are 10% lower for individuals of mothers with a high school degree, 13% lower for individuals of mothers with some post-secondary certificate, 29% lower for individuals whose mothers have a university degree relative to individuals of mothers without a high school diploma, respectively. Moreover, individuals in middle- and high-income families are less likely to drop out of high school relative to children in low-income families. The odds of dropping out of high school are 62% and 89% less for individuals in middle- and high-income families, respectively. Controlling for comorbidities, diagnosis with early chronic conditions and functional limitations are associated with much more high school dropout. The individual with ADHD or asthma diagnoses during childhood are 43 and 32 percentage points more likely to drop out of high school. Individuals with functional limitations are around 16 percentage points to drop out of high school.

Table 3.9: Odds Ratio of the Impact of Type I Diabetes on Dropping Out of High School

	Model (1)		Model (2)		Model (3)		Model (4)	
Type I Diabetes	1.678***	(0.174)	1.239***	(0.103)	1.270***	(0.092)	1.170***	(0.032)
Age	0.980***	(0.002)	1.016***	(0.003)	1.016***	(0.005)	1.015***	(0.005)
Female	0.895***	(0.012)	0.849***	(0.017)	0.857***	(0.026)	0.836***	(0.026)
Married	0.796***	(0.011)	1.326***	(0.029)	1.531***	(0.049)	1.542***	(0.050)
White reference								
Black	1.425***	(0.025)	1.151***	(0.031)	1.268***	(0.049)	1.261***	(0.049)
Asian	0.581***	(0.019)	0.660***	(0.033)	0.612***	(0.051)	0.610***	(0.051)
Others	1.387***	(0.047)	1.247***	(0.064)	1.071	(0.084)	1.069	(0.084)
West region reference								
Northeast	0.779***	(0.016)	0.860***	(0.028)	0.906**	(0.043)	0.907**	(0.043)
Midwest	0.748***	(0.015)	0.767***	(0.023)	0.808	(0.036)	0.808***	(0.036)
South	1.059***	(0.017)	1.045*	(0.027)	1.076*	(0.041)	1.075*	(0.041)
Mother without HS reference								
High school			0.901***	(0.032)	0.892***	(0.092)	0.890***	(0.043)
Some post-secondary			0.871***	(0.040)	0.869***	(0.053)	0.868***	(0.035)
University degree			0.718***	(0.028)	0.692***	(0.027)	0.690***	(0.036)
Family income (low-income reference)								
Middle income			0.384***	(0.009)	0.333***	(0.012)	0.335***	(0.012)
High income			0.115***	(0.004)	0.101***	(0.006)	0.102***	(0.006)
Live with both parents			0.864***	(0.012)	0.806***	(0.020)	0.805***	(0.020)
Healthy weight reference								
Underweight			0.245***	(0.021)	1.203*	(0.133)	1.211*	(0.134)
Overweight					1.114***	(0.042)	1.110***	(0.042)
Obesity					1.101***	(0.042)	1.089**	(0.042)
ADHD					1.432***	(0.082)	1.331***	(0.082)
Asthma					1.328***	(0.099)	1.317***	(0.098)
Coronary Health					1.119*	(0.068)	1.109*	(0.068)
Cholesterol					0.973	(0.035)	0.960	(0.035)
Hypertension					1.106***	(0.039)	1.095***	(0.039)
Function limitation					1.167***	(0.040)	1.163***	(0.040)
Mother diabetes							1.263***	(0.080)
Father diabetes							1.007	(0.076)

Notes: The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1%

Table 3.10 shows the average marginal effect of Type I diabetes on the probability of dropping out of high school and on the probability of having a university degree. In the unadjusted model (1), individuals who experience type I diabetes in childhood are almost 7 percentage points more likely to drop out of high school, compared with those who did not experience diabetes during early childhood. After controlling for family socioeconomic status, the effect of type I diabetes on the likelihood of being employed is attenuated. Individuals with type I diabetes are 6 percentage points more likely to drop out of high school. After further controlling for childhood health and family health, individuals who developed type I diabetes are almost 3 percentage points more likely to drop out of high school than non-diabetic individuals.

Table 3.10: Average Marginal Effect of Type I Diabetes on Dropping out of High School and Have a University Degree

	Model (1)	Model (2)	Model (3)	Model (4)
Dropping out of high school	0.0730***	0.0602***	0.0553***	0.0265*
	(0.01327)	(0.0131)	(0.01337)	(0.0145)
Have a university degree	-0.0932***	-0.070***	-0.06218***	-0.0484***
	(0.0116)	(0.0112)	(0.0115)	(0.0125)
Demographic	YES	YES	YES	YES
Family socioeconomic	NO	YES	YES	YES
Childhood health	NO	NO	YES	YES
Family health & age groups	NO	NO	NO	YES

Notes: each model is adjusted for some variables that include (i) age, gender, race, marital status, geographic area; (ii) living with two parents, mother and father education, and family income; (iii) obesity, cholesterol, hypertension, depression, asthma, ADHD; and (v) father and mother diabetes, and age groups. the Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1% respectively.

3.7.6. University Degree

Table 3.11 shows the effect of the Type I diabetes on the probability of having a university degree. Individuals with type I diabetes are 10 to 38 percentage points less likely to have a university degree relative to non-diabetics. Type I diabetes in children as a chronic health shock may increase uncertainty about future health and the negative impacts of complications on productivity, which in turn may decrease the incentive to invest in higher education, especially after getting a high school degree. Controlling for basic demographic factors, individuals who experienced Type I diabetes during childhood are 38 percentage points less likely to have a university degree. The probability of not having a university degree decreased to 10 percentage points after controlling for family socioeconomic status and comorbidities during childhood. Controlling for family socioeconomic status, the results also show that individuals of mothers with a high school diploma are almost 9 percentage points more likely to have a university degree relative to individuals of mothers with less than high school. In addition, individuals of mothers with a post-secondary certificate or university degree are 10 and 20 percentage points more likely to have a university degree than individuals of mothers who did not complete high school. Controlling for comorbidities, individuals diagnosed with chronic conditions during early childhood have worse educational outcomes: individuals diagnosed with ADHD and asthma are almost 30 percentage points less likely to have a university degree. This evidence is consistent with the previous literature on the relationship between early health shocks and educational attainment. For instance, Kessler et al. (1995) found that children who developed early ADHD or mental health shock are less likely to graduate from high school or to attend college.

Table 3.10 reports the average marginal effect of having Type I diabetes in childhood on the likelihood of having a university degree. The unadjusted model (1) shows that individuals experience type I diabetes in childhood are almost 9 percentage points less likely to attend and get a university degree, compared with those who did not experience diabetes during early childhood. After controlling for family socioeconomic status, the effect of type I diabetes on the likelihood of being employed is attenuated. Individuals with type I diabetes are 7 percentage points less likely to have a university degree. After further controlling for childhood health and family health, individuals who developed type I diabetes are almost 5 percentage points less likely to have a university degree than non-diabetic individuals.

In summary, a childhood Type I diabetes diagnosis has a negative effect on an individual's educational attainment and the size of the effect is especially large for dropping out of high school. Moreover, other health conditions and functional limitation have negative effects on both dropping out of high school and having a university degree; specifically, adding comorbidities to the basic model has a significant impact on the other coefficient estimates, which implies that parents' education and income can be a pathway for children's health which, in turn, can serve as a channel to improve long-run outcomes.

Table 3.11: Odds Ratio of the Impact of Type I Diabetes on Having a University Degree.

	Model (1)		Model (2)		Model (3)		Model (4)	
Type I diabetes	0.623***	(0.062)	0.904***	(0.031)	0.847***	(0.026)	0.901***	(0.036)
Age	1.062***	(0.002)	1.019***	(0.002)	1.009***	(0.004)	1.009***	(0.004)
Female	1.037***	(0.009)	1.103***	(0.015)	1.054***	(0.021)	1.062***	(0.022)
Married	1.436***	(0.014)	0.792***	(0.013)	0.619***	(0.014)	0.616***	(0.014)
White reference								
Black	0.591***	(0.008)	0.750***	(0.017)	0.748***	(0.024)	0.750***	(0.024)
Asian	2.509***	(0.039)	2.612***	(0.066)	2.613***	(0.106)	2.618***	(0.106)
Others	0.633***	(0.018)	0.708***	(0.030)	0.779***	(0.048)	0.779***	(0.048)
West region reference								
Northeast	1.342***	(0.017)	1.375***	(0.028)	1.332***	(0.040)	1.331***	(0.040)
Midwest	1.043***	(0.013)	1.074***	(0.021)	1.064***	(0.030)	1.064***	(0.030)
South	1.109***	(0.012)	1.207***	(0.021)	1.228***	(0.032)	1.228***	(0.032)
Mother does not have HS reference								
High school			1.103***	(0.089)	1.091***	(0.092)	1.086***	(0.087)
Some post-secondary			1.122***	(0.099)	1.101***	(0.091)	1.100***	(0.001)
University degree			1.371***	(0.011)	1.322***	(0.001)	1.202***	(0.019)
Family income (low-income reference)								
Middle income			2.388***	(0.047)	2.833***	(0.075)	2.824***	(0.075)
High income			7.210***	(0.138)	8.215***	(0.232)	8.180***	(0.232)
Live with both parents			1.538***	(0.020)	1.747***	(0.046)	1.748***	(0.047)
Healthy weight reference								
Underweight					0.829***	(0.064)	0.827***	(0.064)
Overweight					0.787***	(0.019)	0.788***	(0.019)
Obesity					0.619***	(0.016)	0.622***	(0.016)
ADHD					0.725	(0.014)	0.712	(0.015)
Asthma					0.656***	(0.049)	0.699***	(0.049)
Coronary heart					1.107**	(0.056)	1.114***	(0.057)
Cholesterol					1.089***	(0.026)	1.096***	(0.027)
Hypertension					0.788***	(0.019)	0.792***	(0.020)
Function limitation					0.806***	(0.019)	0.808***	(0.019)
Mother diabetes							0.848***	(0.051)
Father diabetes							0.966	(0.058)

Notes: The Standard errors are in parentheses. *, **, *** indicate significance level at 10%, 5%, 1%

3.8. Conclusion

This study estimates the impact of type I diabetes diagnosis in early life on educational attainment and labor market outcomes in early adulthood. Type I diabetes is used as an exogenous health shock during early life to investigate one type of health shock on long-run outcomes. This study extends the current literature on the impacts of diabetes by explicitly including only type I diabetes. Furthermore, this study also addresses selection bias, which can appear when type I diabetics decide not to participate in the labor force. The results show negative impacts on both educational attainment and labor market outcomes during early adulthood (20 – 45) for those who developed type I diabetes in childhood (before the age of 15). The negative impact on labor market outcomes is clearer after age 26 when benefits from investment in education exercise their role. On average, individuals who were diagnosed with type I diabetes in childhood have a lower probability of being employed, work fewer hours per week on average, have lower total earnings, and a higher probability of receiving social assistance. There are several explanations for these results. Complications from type 1 diabetes, especially acute hypoglycemia, may prevent individuals from working a full day or may increase absenteeism and this explanation is consistent with other literature (Julius et al., 1993). Individuals with type I diabetes may experience discrimination. For instance, employers may discriminate against individuals with diabetes who are at risk of hypoglycemia, by not allowing them to work on jobs that require high levels of safety and concentration (Matsuuhima et al., 1993; Songer et al., 1989; Kraut et al., 2001). Overall, Type I diabetes as a health shock in early life may affect cognitive and non-cognitive and mental and physical health stock in early adulthood, which may, in turn, affect health, productivity, labor force participation and

consequently earnings. In addition, diabetic individuals may have weaker preferences for work and miss more days due to poor health conditions.

The results also show that there is a socioeconomic heterogeneity in the impact of type I diabetes on labor market outcomes. That is, the impact of type I diabetes in childhood on labor force outcomes in adulthood varies with parental income and educational attainment. Type I diabetic individuals whose parents did not complete high school and who are low income are at higher risk with respect to different labor market outcomes.

The results also suggest that individuals with type 1 diabetes have lower educational attainment than non-diabetics. They are more likely to drop out of high school and to get a university degree. The socioeconomic status of the family has a positive impact on the long-run consequences of diabetes, mitigating the negative effects. Since type I diabetes management is costly, a more negative impact on educational attainment can be expected among children of parents with low education and income. According to the theoretical model, parental investment in children in their early stages in life is more productive – in terms of cognitive and non-cognitive abilities - than later investment, when accumulated skills increase the success probability of getting more abilities and skills in the future.

One major limitation of this study is the cross-sectional nature of the NHIS survey. Panel data would provide a more comprehensive analysis, by allowing follow up over the years in question and a more complete picture of the evolution of skills and abilities from the health shock during childhood to outcomes in adulthood.

3.9. References

1. Abozaid, H., (2014). Cognitive function and academic achievement in children with type 1 diabetes. American Journal of Research Communication. 2014: Vol 2(12).
2. Åkerblom, H. K., Vaarala, O., Hyöty, J., Ilonen, M. Knip, M. (2002). Environmental Factors in the Etiology of Type 1 Diabetes." American Journal of Medical Genetics 115: 18–29.
3. Almond, D., & Currie, J. (2011). Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, *25*(3), 153–172.
4. Almond, D., Currie, J. (2011). Human capital development before age five. O. Ashenfelter, D. Card (Eds.), Handbook of Labor Economics, vol. 4, North Holland, Amsterdam, Netherlands
5. American Diabetes Association. (2009). Diabetes care in the school and daycare setting. Diabetes Care, 32, 568 – 572.
6. Bade-White, P., & Obrzut, J. (2009). The neurocognitive effects of type 1 diabetes mellitus in children and young adults with and without hypoglycemia. Journal of Developmental and Physical Disabilities, 21, 425 – 440.
7. Behrman, J. R., & Rosenzweig, M. R. (2004). Returns to birthweight. The Review of Economics and Statistics, 86(2), 586–601.
8. Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. The Quarterly Journal of Economics, 122(1), 409-439.
9. Brown, I. I. I. H., Perez, A., Yarnell, L. M., Pagan, J. A., Hanis, C. L., Fischer-Hoch, S. P. & Mccormick, J. B. (2011). Diabetes and employment productivity: does diabetes management matter? *The American journal of managed care, 17, 569-76.*

10. Brown, JB., Nichols, CA., Glauber, HS., Bakst, AW., (1999). Type 2 diabetes: incremental medical care costs during the first 8 years after diagnosis. *Diabetes Care,* 22, 1116-1124.
11. Case, A., and Paxson, C. (2008). Stature and status: Height, ability, and labor market outcomes. Journal of Political Economy, 116(3).
12. Case, A., Fertig, A., Paxson, C., (2005). The lasting impact of childhood health and *circumstance. Journal of Health Economics, 24, 365-389.*
13. Centers for Disease Control and Prevention: National diabetes statistics report, 2014: estimates of diabetes and its burden in the United States A, GA: U.S. Department of Health and Human Services; 2014. http://www.cdc.gov/diabetes/pubs/statsreport14/national-diabetes-report-web.pdf
14. Cheah, YS., Amiel, SA., (2012). Metabolic neuroimaging of the brain in diabetes mellitus and hypoglycemia. Nat Rev Endocrinol. 8(10): 588–597.
15. Currie, J. (2009). Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. Journal of Economic Literature. 47 (1), 87-122.
16. Currie, J., and Rosemary, H., (1999). Is the impact of health shocks cushioned by socioeconomic status? The case of low birth weight. American Economic Review 2, 19-22.
17. Currie, J., Stabile, M., Manivong, P., Roos, L. L. (2010). Child health and young adult outcomes. Journal of Human Resources, 45(3), 517–548.
18. Dahlquist, G. (1998). The Aetiology of Type 1 Diabetes: An Epidemiological Perspective. Acta Paediatrica Supplement. 425: 5–10.
19. Dahlquist, G. G., C. Patterson, and G. Soltesz. (1999). Perinatal Risk Factors for Childhood Type I Diabetes in Europe. The EURODIAB Substudy 2 Study Group. Diabetes Care. 22, 1698–1702.

20. Daneman, D. (2006). Type 1 Diabetes. *The Lancet*. 367, 847–858.
21. Deaton, A. (2003). Health, inequality, and economic development. *Journal of Economic Literature XLI*, 113–158.
22. Desrocher, M., Rovet, J., (2004). Neurocognitive correlates of type 1 diabetes mellitus in childhood. Child Neuropsychol 2004; 10(1): 36–52
23. Farmer, E. (1993). Externalizing Behavior in the Life Course: The Transition from School to Work. *Journal of Emotional and Behavioral Disorders*, 1, 179-188.
24. Farmer, E., (1995). Extremity of Externalizing Behavior and Young Adult Outcomes. *Journal of Child Psychology and Psychiatry*, 36, 617-632.
25. Figlio, D. N., Guryan, J., Karbownik, K., & Roth, J. (2013). The effects of poor neonatal health on children's cognitive development (Tech. Rep.). National Bureau of Economic Research.
26. Fletcher, J.M., Richards, M.R., (2012). Diabetes's 'Health Shock' To Schooling and Earnings: Increased Dropout Rates and Lower Wages and Employment In Young Adults. Health Affairs, 31(1), 27-34.
27. Gaudieri PA, Chen R, Greer TF, Holmes CS. (2008). Cognitive Function in Children With Type 1 Diabetes. Diabetes Care 2008;31(9):1892-97.
28. Gerdtham, U. G. and Johannesson, M. (2000). Income-related inequality in life-years and quality-adjusted life-years. *Journal of Health Economics*, 19, 1007-26.
29. Gerdtham, U.-G., and Johannesson, M. A. I., (2004). Absolute Income, Relative Income, Income Inequality and Mortality? *Journal of Human Resources*, 39, 228-247

30. Gilmer, TP., O'Connor, PJ., Rush, WA., Crain, AL., Whitebird, RR., Hanson, AM., Solberg, LI., (2005). Predictors of health care costs in adults with diabetes. Diabetes Care, 28(1), 59–64.
31. Haas, S. A., Glymour, M. M., & Berkman, L. F. (2011). Childhood health and labor market inequality over the life course. Journal of Health and Social Behavior, 52(3), 298–313.
32. Haas, S.A. (2006). Health Selection and the Process of Social Stratification: The Effect of Childhood Health on Socioeconomic Attainment. *Journal of Health and Social Behavior*. 47(4), 339–54.
33. Hannonen, R., Komulainen, J., Eklund, K., Tolvanen, A., Riikonen, R., Ahonen, T., (2010). Verbal and academic skills in children with early-onset type 1 diabetes. Dev Med Child Neurol 52:143 147.
34. Heckman, J., (2007). The Economics, Technology and Neuroscience of Human Capability Formation. IZA DP No. 2875.
35. Holmes, C. S., Cant, M. C., Fox, M. A., Lampert, N. L., Greer, T. (1999). Disease and demographic risk factors for disrupted cognitive functioning in children with insulin-dependent diabetes mellitus (IDDM). School Psychology Review,
36. Ingberg C, Palmér M, Åman J, Larsson S., (1996). Social consequences of insulin-dependent diabetes mellitus are limited: a population-based comparison of young adult patients vs healthy controls. *Diabet Med*, 13, 729–733.
37. James RG, Alberti KGMM, Mayer BD, (2002). Report on the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care, 25: S5–S20.

38. Julius, U., Gross, P., Hanefeld, M., (1993). Work absenteeism in type 2 diabetes mellitus: results of the prospective Diabetes Intervention Study. Diabetes Metab, 19, 202–206.
39. Kautz, T., Heckman, J.H., Diris, R., Weel, B.T., Borghans, L. (2014). Fostering and Measuring Skills: Improving Cognitive and Non-Cognitive Skills to Promote Lifetime Success. NBER Working Paper Series, Working Paper 20749.
40. Kessler, R., Foster, C., Saunders, W., Stang, P.,(1995). Social Consequences of Psychiatric Disorders, I: Educational Attainment. American Journal of Psychiatry, 152, 1026-32.
41. Kraut, A., Walld, R., Tate, R., Mustard, C., (2001). Impact of diabetes on employment and income in Manitoba, Canada. Diabetes Care, 24, 64–68.
42. Kuh, D.J. and Wadsworth, M.E. (1993). "Physical Health Status at 36 Years in the British National Birth Cohort." *Social Science and Medicine* 37(7): 905-916.
43. Lawlor, D., Clark, H., Smith, G.D., Leon, D., (2006). Intrauterine Growth and Intelligence Within Sibling Pairs: Findings from the Aberdeen Children of the 1950s Cohort. Pediatrics, 117 (5), 894-902.
44. Lundborg, P., Nilsson, A., & Rooth, D.-O. (2014). Adolescent health and adult labor market outcomes. *Journal of Health Economics*. *37*, 25–40.
45. Maslow, G. R., Haydon, A., Mcree, A. L., Ford, C. A. & Halpern, C. T. (2011). Growing up with a chronic illness: social success, educational/vocational distress. *Journal Adolesc Health.* 49, 206-12.
46. Matsushima, M., Tajima, N., Agata, T., Yokoyama, J., Ikeda, Y., Isogai, Y., (1993). Social and economic impact on youth-onset diabetes in Japan. Diabetes Care. 16, 824–827.

47. McLeod, J.D., Kaiser, K., (2004). Childhood Emotional and Behavioral Problems and Educational Attainment. American Sociological Review, 69, 636-658.
48. Meo, SA., Alkahlan, MA., Al-mubarak, MA., Al-obayli, MS., Melaibary, BA., Bin Dous2 and, AN., Alhassoun, AI., (2013). Impact of type 1 diabetes mellitus on academic performance. Journal of International Medical Research, 41(3), 855-858.
49. Minor, T., (2011). The Effect of Diabetes on Female Labor Force Decisions: New Evidence From The National Health Interview Survey. Health Economics, 20, 1468–1486.
50. Narayan, KMV., Boyle, JP., Thompson, TJ., Sorensen, SW., Williamson, DF., (2003). The lifetime risk for diabetes mellitus in the United States. JAMA, 290(14), 1884-1890.
51. Northam, E.A., Anderson, P.J., Jacobs, R., Hughes, M., Warne, G.L., & Werther, G.A. (2001). Neuropsychological profiles of children with Type 1 diabetes six years after disease onset. Diabetes Care, 24, 1541–1546.
52. Olivia, J., Lobo, F., Molina, B., Monero, S., (2004). Direct health care costs of diabetic patients in Spain. Diabetes Care, 27, 2616–2621.
53. Oreopoulos, Phillip, Mark Stabile, Randy Walld, and Leslie Roos. (2005). Short, medium, and long term consequences of poor infant health: An analysis using siblings and twins. National Bureau of Economic Research, Cambridge, MA. Working Paper no. 11998.
54. Ramsey, S., Summers, KH., Leong, SA, Birnbaum, HG., Kemner, JE., Greenberg, P., (2002). Productivity and medical costs of diabetes in a large employer population. Diabetes Care, 25(1), 23–29.
55. Rogers, M.A.M., Kim, C., Banerjee, T. et al. Fluctuations in the incidence of type 1 diabetes in the United States from 2001 to 2015: a longitudinal study. BMC Med 15, 199 (2017). https://doi.org/10.1186/s12916-017-0958-6

56. Royer, H. (2005). Separated at girth: estimating the long-run and intergenerational effects of birthweight using twins. Ford School of Public Policy, University of Michigan, Ann Arbor MI.

57. Rusan Chen, Patricia A. Gaudieri, Tammy F. Greer, and Clarissa S. Holmes. Diabetes Care. (2008). Cognitive function in children with type I diabetes: a meta-analysis. 31 (9)(Sept. 2008) p1892.

58. Ryan, C.M. (1988). Neurobehavioral complications of Type 1 diabetes: Examination of possible risk factors. Diabetes Care, 11, 86–93. Sheslow, D., & Adams, W. (1990)

59. Selby, jv., Ray, GT., Zhang, D., Colby, CJ., (1997). Excess costs or medical care for patients with diabetes in a managed care population. Diabetes Care, 20, 1396-1402.

60. Smith, J. P. (2007). The impact of socioeconomic status on health over the life-course. Journal of Human Resources, 42(4), 739–764.

61. Smith, J. P. (2009). The impact of childhood health on adult labor market outcomes. The Review of Economics and Statistics, 91(3), 478–489.

62. Songer T, LaPorte R, Dorman J, Orchard T, Becker D, Drash A. (1989). Employment spectrum of IDDM. *Diabetes Care*, 12, 615–622.

63. Steen, C .K., Landin-Olsson, M., Nystrom, L., Arnqvist, H. J., Bolinder, J., Ostman, J. & Gudbjornsdottir, S. (2010). Long-term detrimental consequences of the onset of type 1 diabetes on annual earnings-evidence from annual registry data in 1990-2005. *Diabetologia*, 53, 1084-92.

CONCLUSION

The findings of my chapters have important policy implications. In Chapter 1, the results show that retirement has a negative impact on health outcomes, and so any policy that prolongs the working period, such as increases in the retirement age, will be accompanied by social benefits that may be greater than the estimated savings in the government's pension burden. Increasing the retirement age may delay the decline in cognitive function and impairment due to aging and may also improve mental health through mitigating factors that increase depression. Moreover, since late retirement is health-preserving, as measured by self-reported health status and two other health outcomes, labor force participation of older people may reduce health care utilization and expenditure, which may add to relieve the government's financial pension burden. In contrast, delaying retirement age could increase drinking participation and intensity among working individuals. Therefore, intervention policies should address the negative impact of delaying retirement age on health-related behavior. The results also shed light on the important role of the retiree's partner's retirement decision on health and health-related behavior of retirees. The costs and benefits arising due to retirement should be internalized when retirement policies are formed. Also, the findings suggest that there is significant heterogeneity across gender and socioeconomic groups. In general, the transition from working to retirement status has a larger negative impact on retired males than retired females. Therefore, a policy that raises the retirement age for men more than females could narrow the gender gap effect of retirement on health.

In Chapter 2, I find that the postnatal period is not the most sensitive for the effect of first-time maternal depression exposure on emotional, physical, cognitive, and social domains. Maternal depression exposure has an effect whenever exposure first occurs, but with a stronger effect during the prenatal period, followed by the preschool period. That is, depression exposure

during the prenatal period is more detrimental than the postnatal period for several profiles of child development at kindergarten age. These findings underscore the need for early detection of maternal depression through several maternal depression screening instruments, ideally by obstetricians during pregnancy and by pediatricians through birth to the preschool period, to reach mothers and children exposed to maternal depression in the most effective manner. Additionally, educational handouts of the effects of maternal depression should target programs that focus on the mother and child together. According to human capital accumulation, intervention during pregnancy may provide a valuable opportunity to improve child early development. Intervention programs should commence prior to the start of school to mitigate early developmental difficulties, which exacerbate if they are not addressed. I also find that among socioeconomically disadvantaged families, the children of teenage mothers at first childbirth, children in families living on income assistance, and children of mothers who have not completed high school are the most at-risk. Therefore, policy interventions that target these at-risk families should be given priority, and they can be identified and reached through existing social services authorities (income assistance system, healthy child program, Healthy child office). I also find that boys are more likely to be vulnerable in EDI domains than girls. Therefore, policy interventions that target vulnerable boys should be prioritized to improve school readiness. These findings are relevant to policymakers in Manitoba, such as Manitoba Health, the Department of Education and Training, the Department of Families, Healthy Child Manitoba, and health care professionals, in particular obstetricians and pediatricians. They are also relevant to the school division, so that they can deliver better policies through their programs.

Finally, in Chapter 3, I find that type I diabetes in childhood has negative impacts on human capital accumulation in early adulthood. Individuals diagnosed with Type I diabetes in childhood

are less likely to be employed, and if they are, they may experience discrimination. For instance, employers may discriminate against Type I diabetics suffering from high-risk hypoglycemia by not allowing them to work on jobs that require a high level of safety and concentration (Matsuuhima et al., 1993; Songer et al., 1989; Kraut et al., 2001). Type I diabetic complications, especially in the case of acute hypoglycemia, may prevent individuals from working a full day or may increase the absence rate (Julius et al., 1993). As Type I diabetes management is costly, a more negative impact of diabetes on educational attainment can be expected among children of parents with low education and income level. Therefore, policy intervention should target the most at-risk families to help them get access to the resources that help them in manage the disease. According to the dynamic complementary effect and self-productivity effect in the human capital model, early intervention through policies or parental investment to mitigate the negative effect of the onset of Type I diabetes are more productive, in terms of human capital accumulation, than later intervention or investment, because accumulated skills increase the success of getting more abilities and skills in the future. Therefore, early detection of Type I diabetes through enforcing screening instruments early in schools will help policymakers to design the programs that mitigate the negative impact of diabetes on individuals' learning and academic performance early in life.

This thesis has presented three chapters that evaluate health effects at different life stages. Two chapters focus on capital accumulation of young people; one chapter focuses on the exit from the labour force at retirement. Health affects outcomes at all stages. The policy implications largely focus on early intervention to minimize the negative effects, at any stage in life.