Essays on Human Capital Disruption

A Dissertation
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Heather Marie Dahlen

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Judy Temple

July 2015

Acknowledgements

There are many individuals who have helped me with this project that I would like to acknowledge. I am grateful for the years of advice, guidance, and encouragement of my advisor, Judy Temple. The members of my committee, Liz Davis, Ezra Golberstein, Katy Kozhimonnil, and Sharon Long, have provided constructive feedback, expertise, and direction on this work throughout this process.

I will be forever indebted to SHADAC for introducing me to wonderful colleagues and friends that have assisted me over the years. Thank you in particular to Donna Spencer and Sharon Long for being inspirational role models, mentors, and occasional therapists. I appreciate the willingness to help and patience demonstrated by Michel Boudreaux, Peter Graven, Jessie Kemmick Pintor, Gilbert Gonzales, and Giovann Alarcon. I am thankful for both the exposure to new ideas and the opportunities provided by SHADAC, as they have shaped my work and piqued my interest in health services research and health economics.

When I decided to pursue a PhD, I knew the process would be humbling. Despite some rather large obstacles, I am grateful for having been challenged. The trials and setbacks encountered along the way helped me discover the kindness of the APEC faculty. In particular, thanks to Liz Davis, Terry Hurley, Steve Polasky, Jay Coggins, and Paul Glewwe for encouraging me, and allowing me, to keep moving forward in the program. I am also forever indebted to Kristine West for a conversation that we had in her kitchen after I had been knocked down by the process and was ready to admit defeat, as it helped me realize that I might not win every battle, but with hard work I can win the war.

The University of Minnesota is where I met many talented friends. Charlotte Tuttle and Sarah Humpage, thanks for sharing non-academic interests and also being great study partners. Katie Genadak, I will never remember how to spell your last name but will always remember how you were willing to drop everything when I needed help. Kristine West, thanks for blazing the “mommy trail” for the rest of us and demonstrating that academics and family are not mutually exclusive.

My parents and sister have always supported me in my academic (and life) adventures. Growing up my mother used to get so nervous during my state piano competitions that she would pace the halls and hide in the stairwell. I’ve always known, and appreciated, that she worries enough for both of us. Mom, you may come out of the stairwell now. By the way my dad lives his life, he has made it well-known that hard work and dedication are the key ingredients to any success in life, big or small. My sister has been a continued source of emotional support and sanity, and her help after the birth of our children made it possible for me to be both a mother and a student. This PhD would not have happened without the love and encouragement of my husband, Jeff. You are the glue that held our family together while I’ve been chasing my dream and I am forever grateful for everything you have sacrificed. Lastly, I want to thank my daughters, Ruby and Samantha, for inspiring me to push through and finish this degree.

Dedication

To my grandpa, Ralph Mach. Those were my thinking shamrocks.

Abstract

This dissertation is comprised of three essays related to disruptions in human capital production. In the first essay, the impact of maternal depression on child cognitive and non-cognitive measures for elementary school-aged children is estimated. After applying a bounding methodology to address the methodological concern of endogeneity, maternal depression negatively affects test scores and reduces a child's ability to learn in the classroom environment, self-control, and interpersonal skills, and increases problem behavior. The second essay examines the effect of earlier school start times on classroom outcomes of fifth grade children. The panel of data follows the same children over time, allowing for a methodology that nets out time-invariant unobserved characteristics that might be influencing results. Findings suggest small movements in start time (1-29 minutes earlier) have no impact on cognitive or non-cognitive outcomes, but large movements (60 minutes earlier or more) lead to lower math scores for girls, lower reading scores for boys, and impaired performance in socioemotional measures for both genders. The last essay measures the effect of "aging out" of the dependent coverage provision of the Affordable Care Act. Using a regression discontinuity design, I find that turning 26 leads to increases in labor force participation and directly purchased private insurance for young men and increases in health insurance plan dissatisfaction for both young men and women.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Human capital is the sum of a person's skills, abilities, and talents. Although individuals determine some of their own human capital production, it is also influenced by the world around them. For children in particular, the actions of external forces such as parents and society play a key role in the production of human capital. What happens to human capital levels of a child or young adult when there are changes in these external forces?

To answer this question, we must first look at the determinants of human capital. The traditional concept of its formation relied on how two inputs, education and training, combined to produce a measurable labor market or educational output. Gary Becker's 1964 model expanded this definition by recognizing that the production of human capital is constrained by time, income, and scarce resources. In addition, his framework incorporated more inputs such as innate ability, acquired skills, and investment decisions. Since parents and society influence inputs and constraints, their behavior has the potential to alter a child's human capital formation. The essays in this dissertation examine how disruptions in human capital affect test scores and non-cognitive outcomes of elementary school-aged children as well as labor market and health insurance choices of young adults.

The potential long-term consequences from reduced childhood human capital formation guided my decision to focus on two potential disruptions to capital investment, maternal depression (Chapter 2), and earlier school start time (Chapter 3). Empirical support for the long-run effects of reduced human capital formation stems from studies that built onto Becker's 1964 model. In particular, Becker (1981) and Cunha and Heckman (2007) found that early human capital is determined predominantly through parental investments and endowments in children. Disruptions in child human capital formation are important, as other studies have shown that human capital stock during childhood predicts adult human capital and earnings (Currie and Thomas 1999; McLeod and Kaiser 2004).

Human capital formation does not end when a child completes school. Therefore, the last essay (Chapter 4) focuses on the impact or reduced capital investment on young adults just beginning to build their careers. In this chapter, the disruption in human capital formation may be the result of a federal policy that loosened the link between employment and health insurance. If individuals used the dependent coverage provision of the Affordable Care Act (ACA) to delay, pause, or reduce labor force participation, long-term disruptions to future earnings may have resulted. Since the majority of earnings growth happens in the first decade of a person's employment (Guvenan et al. 2015), this reduction in capital investment may manifest in lower lifetime human capital accumulation.

A common empirical concern for all three essays is disentangling the true relationship between the measurement of human capital and the explanatory variable. This is due in part because production of human capital might be jointly determined with other choices, and when the choices are not observed, endogeneity becomes a threat to the validity of a study's findings. Therefore, each dissertation essay utilizes a different empirical strategy to address the role of endogeneity. Chapter 2 explores how maternal depression impacts human capital of school-aged children by taking information from observed characteristics to make inferences about how unobserved characteristics might be affecting the results. In doing so, this methodology creates bounds on the maximum effect of maternal depression on various child outcomes. Chapter 3 examines how changes in school start time affect human capital measures of school-aged children and exploits the variation from changes in school start times between two academic periods but within the same school. This fixed effects methodology nets out unobserved time-invariant effects that might be influencing results. Lastly, in the essay that looks at how aging out of the dependent coverage provision of the Affordable Care Act impacts labor market and health insurance choices, endogeneity driven by self-selection into type of insurance plan is addressed via a regression discontinuity design.

Although there is variation in topic, methodological approach, and human capital outcomes, the essays in this dissertation are united in theme--each focuses on a potential disruption in human capital formation and the findings fill existent gaps in the literature. In Chapter 2, I identify the range of causal effects of maternal depression on test scores and non-cognitive outcomes, moving the literature away from associative relationships. Whereas most school start time studies examine high school and college-aged individuals, my findings (Chapter 3) demonstrate that earlier school start times impact younger students, too. Lastly, the topic of Chapter 4 is a very heavily-studied provision of health reform, but results are the first to focus on the labor market and insurance coverage outcomes due to loss of provision eligibility. The results in this series of essays have important implications for policymakers as well as the affected children and young adults.

Chapter 2

The Impact of Maternal Depression on Child Academic and Socioemotional Outcomes

2.1 Introduction

Nearly one in ten American adults currently suffers from depression (Centers for Disease Control and Prevention). Depression is costly to society, with direct costs estimated to exceed 80 billion dollars annually in the U.S. (Greenberg et al. 2015), measured in terms of lost wages and health care needs. There is less clarity surrounding the indirect costs of depression. Among these indirect costs are the spillover effects on family, friends, and coworkers.

Previous research has found a person's mental health status impacts adults around them, manifesting as reductions in mental health status of colleagues (D'Souza et al. 2005) and spouses (Fletcher 2009; Siegel et al. 2004). There are also labor market spillovers for spouses of those with mental health problems, such as foregone employment opportunities and lost wages (Tarricone et al. 2000; Rice and Miller 1996; Access Economics & SANE Australia 2000). Maternal depression adversely affects children's health (Casey et al., 2004; Perry 2008; Clayton et al., 2013; Raposa et al. 2014) and behavior (Frank and Meara 2009). However, less studied are the causal effects of maternal depression on test scores and other classroom outcomes of school-aged children. This study adds to the literature of the costs of depression by identifying the range of causal effects of maternal depression on elementary school-aged academic and socioemotional outcomes. Findings demonstrate a lack of evidence of a causal relationship between maternal depression and child test scores, but moderate impacts between maternal depression and non-cognitive outcomes such as self-control, interpersonal skills, internalized problem behavior, and externalized problem behavior.

Also, the magnitudes of these adverse effects increase with chronicity and severity of maternal depression.

This paper focuses on how maternal depression, defined as having a score of greater than 9 (any) or 15 (severe) on the Center for Epidemiological Studies Depression Scale (CES-D), impacts the academic and socioemotional outcomes of her school-aged child, using data from the nationally-representative Early Childhood Longitudinal Study-Kindergarten cohort (ECLS-K). Non-academic outcomes are studied because depressed mothers spend less time talking and playing with their children (McLearn et al. 2006), which might result in children engaging in more disruptive behavior and being less interested in learning while at school.

The causal effect of maternal depression on a child's academic skills and social functioning is difficult to measure due to endogeneity, stemming from the omission of unobserved factors that affect both maternal depression and child outcomes. For example, unmeasured factors that are not easily available in most data sources, such as family history of depression and neighborhood characteristics (e.g. feeling safe in the home), might negatively impact both a mother's mental state as well as a child's emotional and academic outcomes. Therefore, this paper applies a rigorous robustness check to an important topic in public health, and in doing so, addresses the role of endogeneity by identifying bounds of the maximum effect of maternal depression on various child outcomes. The bounding methodology allows for partial identification of the impact of maternal depression on child outcomes and allows inferences about causal effects from non-causal associations.

Since the outcomes of interest are for school-aged children (kindergarten through eighth grade), a mother is defined as being the parent (by birth, adoption, or marriage) of at least one child aged 54 to 203 months. This paper focuses on mothers because women are twice as likely as men to suffer from (major) depression (Kessler 2003), single-parent households are more likely to be headed by a mother than a father (24 percent versus 4 percent) (U.S. Census Bureau, 2012), and, despite changes in labor force participation, mothers in dual-income households spend double the amount of time each week on child care than fathers (Pew Research Center, 2013). Depressed mothers are identified based on their CES-D scale scores, which allows for varying degrees of severity cut-offs. In addition to severity of depression, the study takes advantage of the panel nature of the data and identifies mothers who either report being depressed in multiple waves of the study or whose depression worsens over time.

By using coefficient estimates and R-squared values from multivariate regression models, identified sets, or bounds, are created to identify the range of causal effects of maternal depression on child outcomes, given a range of plausible assumptions about endogeneity. Results from the multivariate models provide either upper or lower bounds of the effect, and if the bound excludes zero then a statistically significant point estimate is interpreted as being in the range of causal effects. Bounding is an attractive methodology given the inherent endogeneity issues present in observational data and relies on the role of *observable* characteristics to provide information on how much of the impact of maternal depression on child outcomes is being driven by the *unobservable* characteristics. This technique uses the differences in observed traits across children whose mothers have varying degrees of depression (none, any, or severe). These differences demonstrate the size and direction of potentially confounding unobserved traits and shed light on the strength of the results from the multivariate regression models.

2.2 Background

There is a sizable literature in child development and pediatrics documenting a relationship between maternal depression and negative child outcomes. Much of the prior work focuses on infant health and postpartum depression in women occurring in the year following childbirth. A review of these studies is provided by Lovejoy et al. (2000), whose meta-analysis of the early interactions between depressed women and infants found that those women who were depressed during their infants' first 3 months of life were more irritable and hostile, less engaged, displayed less warmth and emotion, and were less likely to play with their infants. Postpartum depression leads to negative reactions in infants such as crying, looking away, directed hand movements, and self-soothing behaviors, (Cohn and Tronick, 1983) and reduced mental and motor development skills at the end of infants first year of life (Field, 1995 and Lyons-Ruth et al., 1986). Early mother-infant interactions also predict poorer infant cognitive outcomes at 18 months of age (Murray et al., 1996).

The negative impacts of maternal depression are not limited to infants of mothers battling postpartum depression. Using the National Longitudinal Survey of Youth (NLSY), Frank and Meara (2009) find depression leads to moderately large effects in child behavioral problems (such as bullying, fearfulness, anxiety, and sudden changes in mood) but not cognitive outcomes once children enter school. Similar to my study, Frank and Meara recognize the potential role of endogeneity as causing the results. Instead of creating bounds of the range of the causal effect of maternal depression on child outcomes, Frank and Meara instead focus on the richness of available observable covariates in their data and then apply astute robustness testing. By estimating models restricted to subpopulations of mothers with more than one child, they are able to examine within-mother differences in child outcomes that might be related to maternal depression. Although my paper and Frank and Meara's use different techniques to assess the role of endogeneity in maternal depression studies, by arriving at similar conclusions bolsters support that maternal depression does negatively affect non-cognitive behavior in school-aged children.

Others have also found increased behavioral problems (Welsh-Allis and Ye, 1988; Weissman et al., 1987; Klein et al., 2005; Shaw, Hyde, and Brennan, 2012), poorer infant and child health (Casey et al., 2004), elevated risks of psychopathology (Beardslee et al., 1983; Downey and Coyne, 1990; Orvaschel, 1983), and higher rates of depression (Fendrich, Warner and Weissman, 1990). Additionally, children with chronic health problems like asthma and diabetes whose mothers are depressed have higher emergency room and hospital utilization rates (Perry, 2008; Clayton et al., 2013).

Three additional themes from the literature guided this study. The first is that severity of depression is important. Brennan et al. (2000) demonstrated a positive relationship between severity of maternal depression and behavioral problems and a negative relationship between maternal depression and vocabulary scores for five-year-old children. Secondly, the longer the mother is depressed, the larger the negative impacts on her child. This result is present both for infants (Campbell, Cohn and Meyers (1995)) and young children (Brennan et al., 2000). Lastly, the literature suggests maternal depression may predict long-term negative outcomes for children. Raposa et al. (2014) find negative effects (increased health-related stress and poor social functioning) of maternal depression on children up to 20 years after depression was first reported. Research by Gilliam et al. (2014) finds maternal depression is linked to aggression in youth, even when it was experienced several years prior.

Prior work demonstrates that the impacts of maternal depression on children may begin during infancy and remain through adolescence (and beyond). Examining the role of maternal depression on school-aged child outcomes is important, as it may stymie the formation of human capital in children and economic success of children is influenced by human capital formation (National Research Council and Institute of Medicine 2000; Heckman 2006; Cunha and Heckman 2007; Heckman 2007). This paper augments the literature by examining the effect of maternal depression on school-age child outcomes using nationally representative longitudinal data, allowing for severity, episodic occurrences (chronicity), and longer-term time trends to be studied. Additionally, it builds on the findings of similar studies (Frank and Meara, 2009) by also recognizing the role of endogeneity and applying alternative statistical techniques to evaluate robustness of findings.

2.3 Conceptual Framework

Historically, economists have studied the determinants of classroom success using educational or child quality production functions and have found human capital accumulation by children depends on a variety of inputs from the home and school (Currie 2003; Hanushek 1996; Behrman and Wolfe 1987; Currie 2001; Currie and Stabile

2003). Traditional economics literature has established that poor adult health impacts labor outcomes both of the individual and of other family members (Currie and Madrian 1999; Parsons, 1977; Berger and Fleisher 1984; Coile, 2004). These altered labor outcomes change the home environment by shifting budget constraints, time commitments, and home productivity. This is important as the home environment plays a key role in a child's current and future skill formation (Becker 1964; Heckman 2006; Cunha and Heckman 2007; Heckman 2007).

A parent's mental health is one component of (their own) overall health and Becker's 1964 model demonstrates that family background is a strong predictor of human capital formation in children. Thus, maternal depression can be viewed as a potential shock to her child's human capital formation. The maternal depression literature suggests that the negative effects of maternal depression on a child may extend well beyond the period in which depression is noted and/or treated. Therefore, it is important to have sound empirical results that establish the magnitude of the impact of maternal depression on child academic and non-cognitive outcomes.

To test this empirical relationship between maternal depression and child outcomes, I use a time allocation model of labor supply framework from Wilcox-Gok and Temple (working paper, 2006) and consider a woman's mental health to be a component of her overall level of health. Leisure is part of consumption (following the work of Heckman, 1974) and both other family income and other family labor supply are assumed to be exogenous. Lastly, all components of the time allocation model of labor are measured at the individual level (that of the mother). Given these assumptions, the mother seeks to maximize her utility, which consists of consumption (*C*), her own health (*M*), and her child's human capital (*D*):

$$maxU = U\ (C,\ M,\ D) \qquad (2.1)$$

The three components of utility (*C, M*, and *D*) are produced when time (t_C, t_M, t_D), market goods and services (g_C, g_M, g_D), efficiency producing the goods ($e_C, e_{M,}$ and e_D), and past exogenous endowments of health and child human capital (M_0 and D_0) are combined. The resultant production constraints on the mother's utility are:

$$C = \varphi(g_C, t_C, e_C) \qquad (2.2\text{ a})$$

$$M = \delta(g_M, t_M, e_{M,} M_0) \qquad (2.2\text{ b})$$

$$D = \theta(g_D, t_D, e_{D,} D_0) \qquad (2.2\text{ c})$$

These production constraints are also impacted by characteristics of the child(ren), family, school characteristics, genetic predispositions, and preferences, and alter the efficiency with which she is able to produce *C, M*, and *D*. The mother's efficiency to produce child health is a function of her own health (*M).* Thus, if she is depressed the stock of *M* decreases and she is less efficient in the production of child health. In order to maintain the same level of utility she must reallocate the other inputs, but those are constrained by time (equation 2.3) and budget (equation 2.4). In these constraints, *w* is the mother's wage rate, *F* is other family income and *h* represents hours worked:

$$h = T - t_C - t_M - t_D \qquad (2.3)$$

$$wh + F = p_C g_C + p_M g_M + p_D g_D \qquad (2.4)$$

The time constraint (2.3) is substituted into the budget constraint (2.4) to form a single constraint:

$$w(T - t_C - t_M - t_D) + F = p_C g_C + p_M g_M + p_D g_D \quad (2.5)$$

Incorporating the production constraints into the utility function, the mother's problem is to maximize her utility, which again consists of consumption (C), her own health (M), and her child's human capital (D), subject to time and budget constraints:

$$\max \mathrm{U} = \mathrm{U}((\varphi(g_C, t_C, e_C),\ \delta(g_M, t_M, e_{M,} M_0), \theta(g_D, t_D, e_{D,} D_0))$$

s.t.

$$w(T - t_C - t_M - t_D) + F = p_C g_C + p_M g_M + p_D g_D \qquad (2.6)$$

Solving the maximization problem yields the following first order conditions:

$$\frac{\partial U}{\partial g_C} = \frac{\partial U}{\partial \varphi}\frac{\partial \varphi}{\partial g}_C - \lambda p_C \leq 0,\ \frac{\partial U}{\partial g_C} g_C{}^* = 0,\ \ g_C{}^* \geq 0 \qquad (2.7)$$

$$\frac{\partial U}{\partial g_M} = \frac{\partial U}{\partial \delta}\frac{\partial \delta}{\partial g}_M - \lambda p_M \leq 0,\ \frac{\partial U}{\partial g_M} g_M{}^* = 0,\ \ g_M{}^* \geq 0 \qquad (2.8)$$

$$\frac{\partial U}{\partial g_D} = \frac{\partial U}{\partial \theta}\frac{\partial \theta}{\partial g}_D - \lambda p_D \leq 0,\ \frac{\partial U}{\partial g_D} g_D{}^* = 0,\ \ g^* \geq 0 \qquad (2.9)$$

$$\frac{\partial U}{\partial t_C} = \frac{\partial U}{\partial \varphi}\frac{\partial \varphi}{\partial t}_C - \lambda w \leq 0,\ \frac{\partial U}{\partial t_C} t_C{}^* = 0,\ \ t_C{}^* \geq 0 \qquad (2.10)$$

$$\frac{\partial U}{\partial t_M} = \frac{\partial U}{\partial \delta}\frac{\partial \delta}{\partial t_M} - \lambda w \leq 0,\ \frac{\partial U}{\partial t_M} t_M{}^* = 0,\ \ t_M{}^* \geq 0 \tag{2.11}$$

$$\frac{\partial U}{\partial t_D} = \frac{\partial U}{\partial \theta}\frac{\partial \theta}{\partial t_D} - \lambda w \leq 0,\ \frac{\partial U}{\partial t_D} t_D{}^* = 0,\ t_D{}^* \geq 0 \tag{2.12}$$

$$\frac{\partial U}{\partial \lambda} = w(T - t_C{}^* - t_M{}^* - t_D{}^*) + F - p_C g_C{}^* - p_M g_M{}^* - p_D g_D{}^* = 0, \frac{\partial U}{\partial \lambda}\lambda^* = 0,\ \lambda^* \geq 0 \tag{2.13}$$

Equations 2.7, 2.8, and 2.9 show that the marginal utility per dollar spent on *C, M,* and *D* are equal, and equations 2.10, 2.11, and 2.12 demonstrate that the marginal utility of time spent producing each good (*C, M*, and *D*) are equal. Lastly, equation 2.13 states that income is fully exhausted. These results give a reduced form expression for the mother's optimal demand for child human capital:

$$D^* = \varphi^*(p_C, p_M, p_D, M_0, D_0, F, w, T, e_C, e_M, e_D) \tag{2.14}$$

If the mother is experiencing depression, her health stock (*M*) will decrease if she does not reallocate production time toward making this good. Therefore, I assume that when the mother's health is diminished by depression, she will increase the amount of time spent on her own health production, which reduces the time spent on the child's capital:

$$\overbrace{\qquad}^{-} \quad \overbrace{\qquad}^{-} \quad \overbrace{\qquad}^{+}$$

$$\left(\partial D/_{\partial M}\right) = \left(\partial D/_{\partial t_D}\right)\left(\partial t_D/_{\partial t_M}\right)\left(\partial t_M/_{\partial M}\right) > 0 \qquad (2.15)$$

The reduction in maternal health might also reduce her productive efficiency of the child's human capital. Ceteris paribus, this also reduces child capital:

$$\left(\partial D/_{\partial M}\right) = \underbrace{\left(\partial D/_{\partial e_D}\right)}_{+}\underbrace{\left(\partial e_D/_{\partial M}\right)}_{+} > 0 \qquad (2.16)$$

Additionally, the mother may treat depression with medication. This increase in market goods (g_M) leads to a reduction in the amount of market goods available for the production of the child's human capital:

$$\left(\partial D/_{\partial M}\right) = \underbrace{\left(\partial D/_{\partial g_m}\right)}_{+}\underbrace{\left(\partial g_D/_{\partial g_M}\right)}_{-}\underbrace{\left(\partial t_M/_{\partial M}\right)}_{-} > 0 \qquad (2.17)$$

There are many other routes through which maternal depression might influence child's human capital formation, but these three partial effects demonstrate the same empirical hypothesis-- holding all else constant, a reduction in the mother's health status, driven by the presence of maternal depression, generates a reduction in a child's human capital formation:

$$\left(\partial D/_{\partial M}\right) > 0 \qquad (2.18)$$

I test this relationship by seeing if an increase in a mother's depression score is associated with a reduction in child cognitive and non-cognitive outcomes, the measures of child human capital. In the models, this change in child human capital formation is captured by the coefficient on maternal depression. The pathway between maternal depression and child human capital formation is likely not as clean and ordered as this conceptual framework suggests. There are many less direct routes through which maternal depression may affect a child's development, as depression symptoms are influenced by a myriad of external (and internal) factors that vary from person to person (like income, employment, or recent loss of a family member), but I argue that the mother's response to depression hinges on her preferences, which in turn affects the production functions for each good.

For example, maternal depression could alter both the quality and quantity of parent-child interactions, which may negatively impact the child's emotional development. Additionally, maternal depression might place strain on the marriage, also potentially detrimental to a child. When faced with depression a mother might become less of a disciplinarian, leading to more behavioral problems of her child. Over time, a chronically depressed mother who remains untreated may retreat from family activities and cause a breakdown in the parent-child relationship, leading to feelings of insecurity by the child and potential increase in internalized behavioral problems. Empirically, though, all of these examples would be modeled as a reduction in time spent on the child's human capital and reach the same testable hypothesis summarized by equation

(2.18). This conceptual framework motivates the empirical analyses that follow, estimating the impact of maternal depression on various measures of child outcomes.

2.4 Data

2.4.1 Sample

The Early Childhood Longitudinal Study-Kindergarten Class of 1998-1999 (ECLS-K) is used in order to examine the role of maternal depression on various child outcomes from kindergarten up to eighth grade. The ECLS-K is a large, nationally representative, longitudinal study sponsored by the National Center for Educational Statistics (NCES) to follow roughly 21,400 kindergartners upon entry and through completion of eighth grade. Data collection began in the fall of kindergarten (1998) and follow-up surveys were administered in the spring of kindergarten (1999), the fall of first grade (1999), the spring of first grade (2000), the spring of third grade (2002), the spring of fifth grade (2004), and the spring of eighth grade (2007). These data come from a collection of parent, teacher, and school administrator interviews as well as child assessments.

As the study is longitudinal, sample attrition is expected. The sample decreases with each round of data collection due to ineligible/nonresponse in the questionnaires as well as movers. Raw sample sizes for this study are 19,914 (kindergarten), 14,395 (third grade), and 9,352 (eighth grade). The estimation sample is limited to mothers of the focal child, whether by birth, adoption, step-parent, or legal guardian, reducing the sample by roughly 15-17% each year (N=16,482 in kindergarten, N=11,974 in third grade, and N=7,900 in eighth grade). Respondents were asked a series of validated questions about depressive symptoms in three survey rounds, when the focal child was in the spring of kindergarten, spring of third grade, and spring of eighth grade. Approximately 85% of mothers answered all parts of this questionnaire. For consistency across models, students missing responses for any of the outcome measures or incomplete maternal depression questionnaires were not included in the estimation (representing less than 1% of the final samples). Final analytic samples included 13,978 kindergartners, 11,831 third graders, and 6,873 eighth graders.

2.4.2 Key Measures

The ECLS-K measure of depression is based on the Center for Epidemiological Studies Depression Scale (CES-D). This scale was developed by Lenore Radloff (1977) and is widely used to measure depression in research studies. Respondents were surveyed about depressive symptoms in three periods: spring of the child's kindergarten year, spring of the child's third grade year, and spring of the child's eighth grade year. As shown in Table 2.1, in addition to the specific question, "how often during the past week have you felt depressed?" 11 other questions were asked to assess the respondents' emotional well-being. For all 12 questions, the respondents selected from the choice set of "never," "some of the time," "a moderate amount of time," or "most of the time."

These emotional well-being indicators compose the 12-item short version of the CES-D and are used to construct two measures of maternal depression. Following the guidelines from the National Center for Education Statistics and other studies (Silverstein et al., 2005; Temple and Wilcox-Gok, 2006; Frank and Meara 2009), I assigned depressed mothers in the sample as having "any" depression if their CES-D score was greater than 9 and as having "severe" depression if the CES-D score was greater than 15 (American Psychological Association; Radloff 1977).

It is important to note that the CES-D scale is based on symptoms of depression as seen in clinical cases and is thus not designed for use in clinical diagnosis. However, nearly 85% of individuals clinically diagnosed with depression by a psychiatrist have high CES-D scores, (Radloff, 1977). Although in 20% of cases an individual with a high CES-D score is not clinically depressed, the test is generally regarded as being a valid and reliable screening mechanism (Radloff, 1977). Depending on the sample (kindergarten, third grade, or eighth grade), approximately 15-18% of mothers were characterized as experiencing any depression and 5-7% of mothers were characterized as experiencing severe depression.

Since the ECLS-K follows the same children over time, it could be determined whether the woman experienced depression in multiple years of the study and created an indicator for episodic depression (my measure of chronicity). A dummy variable was created to be one if the mother's CES-D score was greater than 9 in both kindergarten and third grade or greater than 9 in both third grade and eighth grade. This indicator assessed the impact of moderate depression occurring in multiple periods on third and eighth grade outcomes. Dichotomous indicators equal to one if the mother's CES-D score was greater than 9 in the first period (kindergarten; third grade) and greater than 15 in the second (third grade; eighth grade) allowed for testing of the impacts of chronic and worsening maternal depression on children over time. Table 2.2 summarizes how these various definitions of depression were constructed for analytical purposes.

Child outcomes used in this study were measured in the spring of each survey year (kindergarten, third grade, and eighth grade) and appeared in the child and teacher components of the survey (not the component answered by the parent). Math and reading scores from item response theory (IRT) exams provided measures of cognitive ability. In addition to academic performance outcomes, non-cognitive socioemotional child outcomes were examined for the same years as the test scores. These measures were based on the ECLS-K teacher survey that asked teachers to assess each child's ability in five different areas, ranking them as "never" to "very often," which created continuous scales ranging from 1 to 4. The survey did not collect these non-cognitive socioemotional outcomes in eighth grade. Table 2.3 provides detailed descriptions of each of the five measures, but in general, negative scores indicate worsening performance for approaches to learning, self-control, and interpersonal skills (i.e., the student is less likely to be able to learn in the classroom, show self-control, or interact with others), whereas positive scores indicate poorer outcomes for internalizing and externalizing problems (i.e., more problem behaviors such as hitting and anxiety are observed).

Included in the regression analyses are covariates controlling for child, family, and school characteristics that may impact test scores or non-cognitive classroom outcomes. These measures are consistent with those used in other studies (Brennan et al. 2000; Silverstein et al. 2005; and Wilcox-Gok and Temple 2006) as well as the production constraints identified in the conceptual model. The selected covariates are needed in the estimation process to control for the part of the child outcome (e.g., test score) that can be explained by the particular input (e.g., gender). These measures account for the fact that many factors aside from maternal depression are possible determinants of differences in child outcomes.

Child measures include age, race/ethnicity, birth weight, disability status, English as a second language, and gender. Family characteristics include family type (single-parent or two-parent), if the mother was a teenager at birth, socioeconomic status (mother's education, father's education, mother's occupation, father's occupation, and income), and number of children under age 18 in the family. School characteristics are made up of US Census region of the school, urban/rural location, if teacher turnover is a problem, if student overcrowding in the school is a concern, public/private school, if the school is made up of fewer than 10% minority students, and a school neighborhood quality index that measures crime, drugs, violence, gangs, and tension near the school. Table 2.4 provides an overview of the differences in in these observable characteristics between children with a depressed mother and not depressed mother. Models control for the possibility of more than one respondent within the school by clustering the standard errors on school IDs.

2.4.3 Summary Statistics

The estimates of the differences in mean outcomes between children with and without depressed mothers for each survey period suggest that maternal depression is associated with negative impacts in children (Column (1) in Tables 2.5, 2.6, and 2.7). For all grade levels, maternal depression is associated with lower math and reading test scores. The mean values of the five measures of socio-emotional child outcomes also indicate that children of depressed mothers are less able to benefit from the learning environment, exhibit less self-control, demonstrate lower interpersonal skills, and are more likely to display internal and external problem behavior.

As severity of depression increases, so does the magnitude of the associative relationship with child outcomes. For example, a third grader whose mother exhibited any signs of depression (CES-D score greater than 9) had an unadjusted 10.2 point reduction in reading score (Table 2.6, Panel A, Col. 1), whereas one whose mother was severely depressed (CES-D score greater than 15) had a 13.1 point reduction in reading score (Table 2.6, Panel B, Col. 1).

The summary statistics also reveal that episodes of maternal depression occurring in multiple periods are associated with larger impacts on child outcomes than when the mother is depressed in only one period. Relative to children whose mothers were not depressed in either period, third graders whose mothers reported any level of depression in both kindergarten and third grade scored 12.1 and 13.9 points lower in math and reading, respectively (Table 2.8, Col. 1).When a mother's depression was persistent (more than one period) and potentially worsened over time (from any to severe), this was associated with even larger reductions in child outcomes. Third graders whose mothers experienced depression that became worse over time had a 13.9-15.8 point (0.6 standard deviations) difference in test scores and roughly 0.2-0.3 point (0.5 standard deviations) differences in socioemotional outcomes (Table 2.8, Column (1)). Similar negative associative relationships between maternal depression and test scores occurred for eighth graders (Table 2.9, Column (1)).

Three themes emerged from these associative summary statistics that do not control for any additional child, family, or school characteristics: 1. As the severity of maternal depression increases, the negative impacts on child outcomes increases. 2. Outcomes of children whose mothers report depression in multiple periods are worse than outcomes of children whose mothers are depressed in only one period. 3. Depression that becomes more serious (worsening CES-D score) over time has negative impacts on children. These themes form testable hypotheses outlined below.

2.5 Empirical Models

2.5.1 Linear Cross-Sectional Models

Models were estimated in single and multiple time periods. The single time period models assess the impact of maternal depression in a particular period on the child's outcome in the same period (e.g. the presence of any maternal depression in kindergarten on the kindergartner's math score, or the presence of severe maternal depression in third grade on the third grader's ability to learn in the classroom). The following multivariate linear cross-sectional models were generated for single-period outcomes:

$$Y_{ijt} = \alpha + \beta D_{it} + \psi X_{it} + \gamma_j + \varepsilon_{ijt} \qquad (2.19)$$

where Y_{ijt} is the child outcome of individual i at school j in time t, D_{it} is a dummy representing whether or not the mother of child i in time period t is depressed (CES-D score of greater than 9 to indicate any depression and a CES-D score greater than 15 to indicate severe depression), X_{it} is the set of other explanatory variables listed in Table 2.4, γ_j are the school effects that are constant over time, ε_{ijt} is an error term uncorrelated with D_{it} and X_{it}, and α, β, and ψ are the parameters to estimate.

To estimate the effect of chronic maternal depression (either any, with a CES-D score of 9 in two grades, or worsening, with a CES-D score > 9 in one grade and > 15 in the next), a similar multivariate model is estimated but includes unobserved individual effects that are constant over time (ω_i):

$$Y_{ijt} = \alpha + \beta D_{it} + \psi X_{it} + \omega_i + \gamma_j + \varepsilon_{ijt} \qquad (2.20)$$

2.5.2 Inverse Probability Weighting

The assumption with linear cross-sectional models is that the effect of maternal depression on child outcomes is constant and linear, such that β from equations (2.19) and (2.20) represents the true effect of maternal depression on child outcomes. Unobserved random variables are represented by the error term (ε_{ijt}), and the methodology assumes they are **not** affected by maternal depression but might be correlated with it.

Of concern is that this assumption does not hold and that maternal depression **is** correlated with unobservable characteristics that might also influence test scores and behavioral measures. Thus, estimating equations (2.19) and (2.20) using ordinary least squares (OLS) leads to biased estimates of β. Therefore, the next step taken was to implement inverse probability weighting (IPW) as outlined by Hirano, Imbens, and Ridder (2000). This method allows adjustment for preexisting observed differences among groups (i.e. selection bias, or the fact that assignment among children to depressed or non-depressed mothers is not random). It creates a reweighted data set that better resembles a randomized experiment. Individuals are assigned smaller (larger) weights if their observed treatment status is overrepresented (underrepresented), given their covariates.

First, logit models are estimated to predict the probability of being depressed, p, controlling for child, family, and school characteristics. Multivariate regression models are then estimated, reweighting each unit by 1/p. This method is advantageous over the more common propensity score matching estimation, which requires quite a bit of overlap of observed characteristics between the treatment and control groups, as fewer distributional assumptions about the underlying data are needed.

Inverse probability weighting reduces the bias and improves the efficiency of OLS estimates (Wooldridge 2007; Posner et al. 2012; Williamson, Forbes, and White 2014) but does not fully address the concern of potential endogeneity resulting from omitted variables. This is because data are reweighted using observed, not unobserved, characteristics. The results from these models augment the findings from the linear cross-sectional models and also shed light on some of the predictors of maternal depression. For example, logit models estimating the presence of any depression for mothers of kindergartners demonstrate that being Asian (non-Hispanic), having a female child, eating dinner as a family together more frequently, and increased socioeconomic status are all associated with reductions in maternal depression. Additionally, having a child who is disabled, having a child attend a public school, and increased school participation in the free lunch program are associated with increases in any maternal depression. Increased socioeconomic status, family dinners, and child's weight at birth are negatively associated with severe maternal depression, with disability increasing severe maternal depression (Appendix A).

2.5.3 Bounding

To address the concern of endogeneity, one possible technique is the use of instrumental variables, which relies on finding a measure that is correlated with the variable of interest (maternal depression) but not correlated with the outcome (child test scores and non-cognitive outcomes). However, as an alternative, bounding is advantageous since it does not rely on identifying variables like instruments that might be weak in explanatory power and require assumptions that might be difficult to validate. Instead, bounding allows partial identification of the impact of maternal depression on child outcomes and constructs consistent bounds on its true value. If the bounded set excludes zero, this suggests robustness in the range of what would have been seen if the effect of maternal depression had been randomized.

The bounding method outlined by Oster (2015) builds off prior studies that make assumptions about the relationship between observed and unobserved measures in a model in order to determine bounds, or ranges, of the estimated coefficients (Murphy and Topel, 1990; Altonji, Eder, and Taber, 2005; Altonji et al, 2011). However, Oster's methodology establishes bounds on estimates using both movements in the coefficients as well as movements in the R-squared that result when more observed covariates are introduced in a model. By using information about the R-squared, Oster's bounding methodology nets out the added noise that may arise when confounding measures are included in a model.

Adopting the methodology of Altonji, Elder, and Taber (2005), Oster begins with a simple model: $Y = \beta X + W1 + W2$, where $W1$ is observed, $W2$ is unobserved, and β is the coefficient of interest. Applying this to my data, these are equations (2.19) and (2.20) but assuming ε_{ijt} consists entirely of unobserved measures, X_{it} is the vector of observed covariates, and β is the coefficient of interest. Altonji, Elder, and Taber state that information about the relationship between the treatment and the unobserved measures can be learned via the relationship between the treatment and the observed measures, which is called the proportional selection assumption. Recall that *D* is the outcome (e.g., math score) that represents the child's human capital. Using equation (2.19) or equation (2.20), this proportional selection assumption is as follows:

$$\frac{Cov(D_{it}, \varepsilon_{ijt})}{Var(\varepsilon_{ijt})} = \delta \frac{Cov(D_{it}, X_{it})}{Var(X_{it})} \qquad (2.21)$$

In order for equation (2.21) to be useful, values of the degree of proportionality between the observed and unobserved covariates δ, must be assumed. This method captures both the relative importance of the unobserved measures that are related to the observed measures as well as variations in outcomes unique to individuals. Therefore, Oster refines the Altonji, Elder, and Taber model by assuming that the unobserved random variables in the error term (ε_{ijt}) are, in fact, correlated with the treatment (depression). The model incorporates a measure of the unobserved variables after they have been predicted, with respect to the observed characteristics. Following this methodology, I define $\varepsilon_{ijt}{}'$ as the best linear predictor of ε_{ijt} given the observed characteristics (X_{it}):

$$\varepsilon_{ijt} = \varepsilon_{ijt}{}' + v_{ijt} \qquad (2.22)$$

Substituting (2.22) into (2.18):

$$Y_{ijt} = \alpha + \beta D_{it} + \delta X_{it} + \omega_i + \gamma_j + \varepsilon_{ijt}{}' + v_{ijt} \qquad (2.23)$$

Equation (2.23) is similar to equation (2.18) but no longer requires the assumption that the correlation between maternal depression and the residual is zero. Using equation

(2.23), Oster's method defines $\tilde{\delta}$ as the proportionality value between the observed characteristics (X_{it}) and the unobserved variables after an intermediate model has been estimated (providing $\varepsilon_{ijt}{}'$). In doing so, $\tilde{\delta}$ shows precisely how much of the effect of maternal depression on child outcomes is explained by observables versus the unobservables, but now the unobservables contain *only* the omitted variables related to D_{it} and proxied by X_{it}:

$$\frac{Cov(D_{it}, \varepsilon_{ijt}{}')}{Var(\varepsilon_{ijt}{}')} = \tilde{\delta}\,\frac{Cov(D_{it}, X_{it})}{Var(X_{it})} \qquad (2.24)$$

Equation (2.24) provides a relationship between the observed and unobserved characteristics. It provides bounds on the degree of selection and therefore ranges between zero and one ($\tilde{\delta} \in [0,1]$). If the coefficient of interest, $\tilde{\delta}$, is less than one, it indicates that the observed measures have *more* impact on the outcomes than the unobserved measures.

The other piece of information needed to establish the identified set (or bounds) for maternal depression's effect is the maximum R-squared value, R_{max}. This value can be determined using two methodologies. The first uses the R-squared values from prior studies as a guide, and the second takes the R-squared achieved from the controlled models, $\widetilde{R}$, multiplies them by1.3, and if this value is greater than one then R_{max}=1 is used. The cutoff value 1.3 is derived by Oster's sample of 65 papers that use randomized controlled trials. She determined that $\tilde{R}$= 1.3 allowed 90% of the randomized results to survive. As the previous literature is largely associative, the latter method was chosen in selection of the maximum R-squared value ($R_{max} = min\{1.3\widetilde{R}, 1\}$) .

These two pieces of information, $\tilde{\delta}$ and R_{max}, allow for construction of an identified set of the effect of maternal depression. If the effect is negative, the set is:

$$[\tilde{\beta}, \beta^{*\prime}(min\{1.3\tilde{R}, 1\},1)] \qquad (2.25)$$

If the effect is positive, as in the cases where the teacher's ratings of behavioral problems move from low scores (where "never"=1) to high scores ("very often"=4), the set is:

$$[\beta^{*\prime}(min\{1.3\tilde{R}, 1\},1), \tilde{\beta}] \qquad (2.26)$$

Equations (2.25) and (2.26) bound the impact of the treatment effect (maternal depression) using the beta estimated from the model with full controls ($\tilde{\beta}$) and the beta estimated under the assumption that $\tilde{\delta}$=1 and $R_{max} = min\{1.3\tilde{R}, 1\}$, or $\beta^{*\prime}$. Should the bounds exclude zero, then the effect of the treatment on the outcomes is also not zero. Equations (2.25) and (2.26) are used as robustness checks for the estimates of maternal depression on child outcomes produced using multivariate regression models.

In addition, Oster's method allows for hypothetical testing of how much more important omitted variables would have to be than the included control measures for the outcomes. In other words, by setting $\beta = 0$ (indicating maternal depression has no impact on the outcome), a proportional selection coefficient, δ^0, can be obtained. This value states how much larger the omitted variables would have to be, relative to the control variables, in order for the treatment effect to be zero.

2.6 Results

2.6.1 Contemporaneous impact of maternal depression

Single-period multivariate regression results are presented in Column 2 of Table 2.5 (kindergarten), Table 2.6 (third grade), and Table 2.7 (eighth grade). The supplementary IPW results are shown in Appendix B. These multivariate models control for child, family, and school characteristics (described earlier). I begin with the results shown in Panel A of each table, which report the coefficients from the models when the mother reports any level of depression (dichotomous measure equaling one when the CES-D score is greater than 9).

Contemporaneous maternal depression leads to lower test scores across all three time periods (Col. 2 of Panel A in Tables 2.5, 2.6, and 2.7). There are significant reductions in test scores and socioemotional outcomes in all grade levels. Maternal depression is also associated with lower socioemotional scores. Recall these outcomes are based on teacher surveys and reductions in ability to learn, self-control, and interpersonal skills demonstrate movement away from scores of “very often” and toward “never.” When the coefficients for the two problem behavior measures (internalizing and externalizing) are positive, this suggests movement away from scores of “never” and toward “very often.” As shown in Col.2 in Panel A of both Tables 2.5 and 2.6, the magnitude of these reductions is roughly 0.05 to 0.15 points.

The inverse probability weighting (IPW) models are presented in Appendix B. Recall that this method reweights the sample to address the issue of non-random selection into treatment (depressed or not). In general, the same pattern emerges in the IPW results as the linear cross sectional models, although for most models the magnitudes from the IPW models are slightly larger. Significance across models is similar except in eighth grade, for which the IPW models found the presence of any maternal depression did not have an impact on math scores.

Overall, when these changes in scores are translated into changes in standard deviations, presence of maternal depression reduces kindergarten math and reading scores by 0.2 and 0.1 standard deviations, respectively (Table 2.5, Panel A, Col. 2). Similarly, third reading scores declined by 0.2 standard deviations (Table 2.6, Panel A, Col. 2). These reductions are similar in magnitude to the "smaller class size" effects of 0.2 to 0.3 standard deviation increases in test scores (Unlu, 2005).

After converting the socioemotional measures into standard deviations, the presence of any maternal depression lead to roughly 0.1 to 0.2 standard deviation changes in outcomes for both kindergartners and third graders. Again, these results are in the neighborhood of the findings from the smaller class size effect (20 or fewer students) on test scores (Unlu, 2005).

The contemporaneous regression models demonstrate that as severity of maternal depression increases, the magnitudes of the impact on children also increase (Panel B, Col. 2, in Tables 2.5, 2.6, and 2.7). Recall mothers with a CES-D scale of greater than 15 are classified as severely depressed. Consider the case of kindergarten interpersonal skills—a child whose mother reported any depression scored 0.08 points (-0.15 standard deviations) lower than a child whose mother was not depressed, whereas a child whose mother reported severe depression scored (0.10 points (-0.22 standard deviations) lower (Table 2.5, Panel B, Col. 2).

As shown Tables 2.5, 2.6, and 2.7, in both the linear cross-sectional (Col. 2) and the IPW models (Appendix B), when the measure was statistically significant in both the "any" depression (Panel A) and "severe" depression models (Panel B), the magnitude of the relationship between maternal depression and child test scores was greater in the severe models. A kindergartner whose mother had any depressed scored roughly 1.2 points lower in math (-0.15 standard deviations), whereas a kindergartner whose mother was severely depressed score 2 points lower (Table 2.5, Panels A and B, Col. 2).

Kindergartners with severely depressed mothers also fared much worse than children whose mothers were not depressed in terms of socioemotional outcomes. These magnitudes increased from roughly 0.05 standard deviations to 0.15 standard deviations across nearly all measures in kindergarten (Table 2.6, Panels A and B, Col. 2). There were no statistically significant effects for any measure for third graders with severely depressed mothers. Additionally, maternal depression did not predict test scores for eighth graders (Table 2.7). Results from the IPW models were similar (Appendix B).

2.6.2 Multiple episodes of maternal depression

Linear cross-sectional models were estimated to examine the role of chronicity of depression. Tables 2.8 and 2.9 report these coefficients. Again, for the models in which the coefficient of interest is "any" maternal depression (Panel A), this was measured as the mother having a CES-D score greater than 9 in both kindergarten and third grade or both third grade and eighth grade. Maternal depression that potentially worsened[1] over time (Panel B) was based on the mother having a CES-D score greater than 9 in the first period (kindergarten; third grade) and greater than 15 in the second (third grade; eighth grade). In all models, the same set of control measures were used as in the contemporaneous models, but since these models incorporate changes over time, a

[1] If the mother's CES-D score was greater than 15 in period one then her depression may have remained severe in both periods.

lagged outcome variable is added in order to account for prior performance (e.g., the third grade math models included the kindergarten math score on the right-hand side).

If a mother reported any depression when her child was in both kindergarten and third grade, her child scored 5.5 points lower in reading (0.2 standard deviations) than his/her counterpart whose mother did not report depression in those two time periods (Table 2.8, Panel A, Col. 2). These children also had less-developed interpersonal skills (-0.01). For children whose mothers were depressed in both third and eighth grade, there were no significant effects on test scores (Table 2.9, Panel A, Col. 2).

Larger impacts occurred when maternal depression was both recurring and potentially increasing in severity, as shown in Panel (B) of Tables 2.8. When the mother's depression was present in the child's kindergarten year and worsened to severe depression in the child's third grade year, the child's reading score fell over 6 points and there were increases in internal problem behavior such as sadness, anxiety, loneliness, and low self-esteem (+0.14). When maternal depression worsened between third and eighth grade there were no significant impacts on test scores (Table 2.9, Panel (B)).

2.6.3 Bounding

As mentioned, one potential problem with estimates from the above-described models is that results might be driven by unmeasured variables. Thus, bounds were estimated to place limits on the magnitude of the selection on unobservables (Col. 3 of Tables 2.5, 2.6, 2.7, 2.8, and 2.9. Although this methodology has been used to correct for selection bias in other studies (Dee et al., 2006/2007), and a similar methodology has been applied to mental disorder studies (Banerjee, Chatterji, and Lahiri, 2013), it has not previously been applied to maternal depression estimates.

Beginning with the contemporaneous impacts on kindergartners (Table 2.5), the impact of any maternal depression (Panel A) remained statistically significant after addressing the potential role of endogeneity (Col. 3) in all seven models that showed significance via linear cross-sectional analyses (approaches to learning, interpersonal skills, and internalizing problem behavior). For these outcomes, the bounds of the true effect of any maternal depression do *not* include zero (Col. 4), which indicates that the coefficients generated by the linear cross-sectional models are robust. Models for severe depression (Panel B) are similar, with all significant findings representing one bound of the causal estimates.

The additional robustness results (Col. 5) indicate how much larger the role of unobservables would have to be in each model to make the effect of maternal depression on child outcomes zero. These results suggest that the role of unobserved characteristics range between 1 (reading score) and 4 (internalizing problem behavior) in the "any" maternal depression models (Panel A) and range from 1.5 (self-control) to 2.4 (math score) in the "severe" models (Panel B).

Turning to the single-period results for third graders (Table 2.6), after addressing the role of endogeneity the presence of any maternal depression (Panel A, Col. 3) is associated with worsening child outcomes in three models (self-control, interpersonal skills, and externalizing problem behavior). The bounds for the significant -2.6 point reading estimate include zero (Col. 4), and thus it cannot be concluded that this result is due to the presence of maternal depression. For the three models whose bounds exclude zero, I also find that the role of the unobserved measures would have to be roughly twice as important as the role of observed measures in order to make the statistically significant findings fall to zero (Col. 5). There were no statistically significant results for third graders whose mothers were severely depressed (Table 2.6, Panel B, Col. 2). Thus, even in the models for which the calculated causal bounds excluded zero (interpersonal and internalizing), the coefficients from the models had relatively weak explanatory power.

Bounds were also calculated to determine if the results from the chronicity models would remain after addressing endogeneity. As shown in Panel A of Table 2.8, estimates from both models with significant multivariate regression results (reading score and interpersonal skills) passed the bounding threshold. (Columns 3 and 4). For both models, the role of unobserved measures would need to be between 1.5 and 2 times greater than the role of observed measures in explaining the relationship between maternal depression and the outcome (Col. 5). Reduced reading scores and increased rates of internal problem behavior remained significant when the mother reported any depression in kindergarten and severe depression in third grade (Panel B). For the eighth grade outcomes (Table 2.9), there were no significant findings.

By calculating identified sets using Oster's bounding methodology and adjusting for unmeasured confounders in this manner, maternal depression does not predict child outcomes. This is important as omitted variables are of particular concern to the validity of results in many survey-based observational studies. Findings suggest that even after controlling for observed characteristics, for most models that show significant effects in simply multivariate models, the significant effects remain. This paper points to the necessity of reevaluating the causal strength of estimates when the role of unobserved characteristics is not addressed in statistical models. The methodology used in this paper to correct for endogeneity has applications that reach beyond maternal depression and can be utilized in many areas of public health for which an instrumental variable approach is simply not feasible.

2.6.4 Limitations

The analysis has important limitations. First, depression is self-reported by mothers. However, self-reported depressive symptoms are typical measures in screening tools and these data do not contain clinical notes on which I could potentially validate the CES-D scores. Second, although the results are more generalizable because the data are national, the study itself is still observational and not a randomized trial, so internal validity, or degree to which the results are being driven by maternal depression and not other explanations, is a concern. However, by taking advantage of the panel nature in the models for chronicity and by producing robust coefficient bounds, the concern of omitted variable bias is mitigated. Third, as the survey did not capture maternal depression in fifth grade, there is a large gap between third and eighth grade, during which time changes in other factors, such as transitioning to middle-school or junior high, might be driving the changes in outcomes. I believe this is why the results for eighth graders are weak in statistical strength. Lastly, the depression questionnaire has a 1-week look back period, so the models of chronicity may not indicate that the mother was depressed over the entire period. Despite these limitations, this study produces important, policy-relevant results that are not simply associative in nature.

2.7 Discussion

In the existing literature on the consequences of depression within a family, depression of a parent has been shown to be associated with a number of adverse outcomes for family members. Researchers focusing specifically on the impact of maternal depression on infants and children report evidence from dozens of studies in the consequences for child well-being. However, the adverse outcomes associated with maternal depression represent causal impacts only if all relevant variables are included in the analysis. This reliance on the assumption of selection on observables is a concern due to the possibility that unincluded factors affecting both maternal depression and child outcomes (like family history) may exist, and when not accounted for within a modeling strategy may generate spurious results.

This study employs two methodologies to generate estimates of the impact of maternal depression on the test scores and non-cognitive skills of school-aged children, then corrects for endogeneity by establishing bounds on the results. Oster's bounding methodology separates the causal results from the non-causal associative results, and does so by incorporating both the relationship between observed and unobserved variables as well as changes in R-squared values (Oster 2015).

The findings from this study find evidence of a causal relationship between maternal depression and math scores for kindergartners, maternal depression and reading scores for kindergartners and third graders, and maternal depression and reduced socioemotional measures for both kindergartners and third graders. There were no effects of maternal depression on eighth grade test scores, suggesting that this issue is of particular importance for children in their first few years of elementary school.

Additionally, after addressing the role of endogeneity via bounding, this paper finds that severity and chronicity of maternal depression both lead to negative impacts on non-cognitive measures for children-- as the severity of maternal depression increases (higher CES-D values), so do the magnitudes of the effects for kindergartners. For children whose mothers reported chronic depression that worsened over time, the magnitude of causal effect on interpersonal skills increased.

These results motivate a role for health and education policy, as policy efforts require information on causal relationships. The significant effects of maternal depression on non-cognitive child outcomes found in this paper are similar in size to those of the effects of smaller class sizes on test scores. Therefore, diagnosis and treatment, as well as preventive screenings, of depression in mothers could be used to improve or mitigate disruptions of human capital formation of children. This is especially important given the relatively low cost of screening and treatment for mothers compared to potential human capital savings in children.

Table 2.1: Emotional well-being measures

Variable Name	Description (all measures reference the previous week)
Bothered	Felt that you were bothered by things that don't usually bother you?
Appetite	Felt that you did not feel like eating, that your appetite was poor?
Blue	Felt that you could not shake off the blues even with help from your family and friends?
Focus	Felt that you had trouble keeping your mind on what you were doing?
Depress	Felt depressed?
Effort	Felt that everything you did was an effort?
Fearful	Felt fearful?
Restless	Felt that your sleep was restless?
Less talk	Felt that you talked less than usual?
Lonely	Felt lonely?
Felt sad	Felt sad?
Not go	Did not go somewhere you should have?

Source: ECLS-K Codebook

Table 2.2: Variable coding of maternal depression

Depression measure created in study	Center for Epidemiological Studies Depression Scale (CES-D) score
Any	> 9
Severe	> 15
Chronic	> 9 in kindergarten and > 9 in third grade > 9 in third grade and > 9 in eighth grade
Chronic and worsening	> 9 in kindergarten and > 15 in third grade > 9 in third grade and > 15 in eighth grade

Table 2.3: Child socioemotional measures

Variable Name	Description
Learn	Six-item scale that rate the child's attentiveness, task persistence, eagerness to learn, learning independence, flexibility, and organization. This measures the ease with which children benefit from the learning environment.
Control	Four-item scale that includes the child's ability to control behavior by respecting the property rights of others, controlling temper, accepting peer ideas for group activities, and responding appropriately to pressure from peers.
Interpersonal Skills	Five-item scale measuring a child's skill in forming and maintaining friendships, getting along with people who are different, comforting or helping other children, expressing feelings, ideas and opinions in positive ways, and showing sensitivity to the feelings of others.
Externalizing Problem Behavior	Five-item scale measuring the frequency with which a child argues, fights, gets angry, acts impulsively, and disturbs ongoing activities.
Internalizing Problem Behavior	Four-item scale measuring loneliness, sadness, low self-esteem, and anxiety.

Source: ECLS-K Codebook

Table 2.4: Differences in observable covariates between children of depressed and non-depressed mothers

	Kindergarten				Third Grade				Eighth Grade			
	Any		Severe		Any		Severe		Any		Severe	
Child characteristics												
Race/ethnicity												
White, non-Hispanic	-0.07	***	-0.09	**	-0.10	***	-0.14	***	-0.02		-0.04	
	(0.018)		(0.033)		(0.026)		(0.042)		(0.021)		(0.039)	
Black, non-Hispanic	0.06	***	0.09	***	0.04	*	0.00		0.02		0.04	
	(0.013)		(0.024)		(0.020)		(0.017)		(0.011)		(0.051)	
Hispanic	0.02		0.01		0.05	**	0.14	***	0.39		-0.13	
	(0.014)		(0.027)		(0.025)		(0.039)		(0.422)		(0.112)	
Female	-0.04	*	-0.03		0.00		0.02		0.01		0.02	
	(0.017)		(0.031)		(0.026)		(0.043)		(0.012)		(0.024)	
Age	0.19		0.68	*	-0.08		-0.25	*	0.10	***	0.02	
	(0.161)		(0.312)		(0.083)		(0.121)		(0.067)		(0.107)	
Weight at birth	-0.14	**	-0.24	**	-0.04		-0.11		-0.16		-0.23	
	(0.046)		(0.084)		(0.074)		(0.116)		(0.102)		(0.176)	
Disabled	0.04	**	0.07	**	0.06	*	0.08	*	0.101	**	0.21	**
	(0.013)		(0.023)		(0.025)		(0.043)		(0.037)		(0.067)	
English is a second language	0.01		0.01		0.05	**	0.11	***	0.01		0.02	
	(0.010)		(0.020)		(0.017)		(0.032)		(0.022)		(0.037)	
Family Characteristics												
Two biological parents	-0.05	***	-0.09	***	-0.09	***	-9.19	*	-0.18	***	-0.22	***
	(0.008)		(0.019)		(0.022)		(0.029)		(0.036)		(0.062)	
Mother was a teen at child's birth	0.11	***	0.17	***	0.08	***	0.13	**	0.11**		0.148	**
	(0.015)		(0.027)		(0.023)		(0.040)		(0.036)		(0.065)	
Number of children in the home	0.08	*	0.16	*	0.17	**	0.12		0.15		-0.02	
	(0.042)		(0.078)		(0.056)		(0.080)		(0.099)		(0.180)	
Socioeconomic status	-0.35	***	-0.46	***	-0.35	***	-0.51	***	-0.38	***	-0.44	***
	(0.027)		(0.042)		(0.040)		(0.054)		(0.061)		(0.097)	
School Characteristics												

Urban	0.002		-0.03		0.07	**	0.10	*	0.02		-0.05	
	(0.018)		(0.279)		(0.026)		(0.043)		(0.037)		(0.053)	
Region												
Midwest	-0.01		-0.03		-0.02		-0.01		-0.02		-0.10	**
	(0.017)		(0.025)		(0.025)		(0.040)		(0.033)		(0.042)	
South	0.04	*	0.08	*	0.03		0.03		0.08	*	0.07	
	(0.020)		(0.032)		(0.028)		(0.042)		(0.040)		(0.067)	
West	0.00		-0.02		0.01		0.03		-0.01		-0.01	
	(0.017)		(0.028)		(0.023)		(0.038)		(0.032)		(0.052)	
Index of neighborhood problems	0.45	***	0.66	**	0.82	*	1.09		0.59	*	0.02	
	(125)		(0.200)		(0.365)		(0.618)		(0.279)		(0.377)	
Teacher turnover is a problem	0.02		0.02		-0.01		0.03		0.14		0.041	
	(0.040)		(0.058)		(0.049)		(0.091)		(0.072)		(.119)	
School is overcrowded	-0.01		0.01		0.13		0.21	*	-0.09		-0.16	
	(0.047)		(0.081)		(0.068)		(0.115)		(0.098)		(0.143)	
Less than 10% minority in school	-0.05	**	-0.06	*	-0.08	**	-0.06		-0.03		0.01	
	(0.019)		(0.031)		(0.026)		(0.043)		(0.035)		(0.059)	
Public school	0.07	***	0.08	***	0.08	***	0.13	***	0.08	***	0.03	
	(0.011)		(0.017)		(0.014)		(0.015)		(0.023)		(0.045)	
Sample Size	13,978		13,106		11,831		11,400		6,873		6,255	

Note: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers.

Table 2.5: Maternal depression, test scores, and socioemotional outcomes (kindergarten)

Panel A: Any depression (CES-D > 9)

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3\tilde{R}, 1\},1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-3.775	***	-1.169	***	[-1.169, -0.600]	Yes	1.456
	(0.261)	[0.0152]	(0.287)	[0.2671]			
Reading	-4.002	***	-0.960	**	[-0.960, -0.034]	Yes	1.034
	(0.295)	[0.0129]	(0.347)	[0.1842]			
Learning	-0.173	***	-0.076	***	[-0.076, -0.042]	Yes	2.008
	(0.017)	[0.0094]	(0.020)	[0.1435]			
Control	-0.134	***	-0.062	***	[-0.062, -0.035]	Yes	2.094
	(0.016)	[0.0068]	(0.018)	[0.0949]			
Interpersonal	-0.151	***	-0.075	***	[-0.075, -0.050]	Yes	2.428
	(0.016)	[0.0082]	(0.019)				
Externalizing	0.123	***	0.0542	**	[0.029, 0.054]	Yes	1.986
	(0.017)	[0.0054]	(0.196)	[0.1083]			
Internalizing	0.081	***	0.057	***	[0.049, 0.058]	Yes	3.710
	(0.014)	[0.0037]	(0.017)	[0.0390]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 13,978. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.5: Maternal depression, test scores, and socioemotional outcomes (kindergarten)

Panel B: Severe depression (CES-D > 15)

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), [$\dot{R}$]		Controlled Effect $\tilde{\beta}$, (S.E.), [$\widetilde{R}$]		Identified Set [$\tilde{\beta}, \beta^{*\prime}$(min{1.3 $\tilde{R}$, 1},1)]	Exclude Zero?	δ^0 for $\beta = 0$
Math	-5.328	***	-1.970	***	[-1.970, -1.297]	Yes	2.406
	(0.406)	[0.0127]	(0.458)	[0.2690]			
Reading	-5.089	***	-1.022		[-1.022, 0.177]	No	0.862
	(0.509)	[0.0087]	(0.643)	[0.1830]			
Learning	-0.258	***	-0.122	***	[-0.122, -0.073]	Yes	1.997
	(0.030)	[0.0089]	(0.036)	[0.1400]			
Control	-0.202	***	-0.073	*	[-0.073, -0.029]	Yes	1.519
	(0.029)	[0.0065]	(0.033)	[0.0928]			
Interpersonal	-0.233	***	-0.099	**	[-0.099, -0.054]	Yes	1.787
	(0.027)	[0.0083]	(0.032)	[0.0875]			
Externalizing	0.172	***	0.055		[0.007, 0.055]	Yes	1.134
	(0.029)	[0.0045]	(0.032)	[0.1043]			
Internalizing	0.126	***	0.070	*	[0.040, 0.070]	Yes	1.676
	(0.241)	[0.0037]	(0.030)	[0.0287]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 13,106. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.6: Maternal depression, test scores, and socioemotional outcomes (third grade)

Panel A: Any depression (CES-D > 9)

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3\tilde{R}, 1\}, 1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-8.187	***	-1.429		[-1.429, 1.051]	No	0.589
	(0.915)	[0.0138]	(0.905)	[0.2713]			
Reading	-10.182	***	-2.622	*	[-2.622, 0.175]	No	0.942
	(1.049)	[0.0165]	(1.096)	[0.2711]			
Learning	-0.154	***	-0.020		[-0.020, 0.033]	No	0.390
	(0.028)	[0.0061]	(0.028)	[0.1588]			
Control	-0.130	***	-0.060	*	[-0.060, -0.030]	Yes	1.910
	(0.025)	[0.0054]	(0.027)	[0.0980]			
Interpersonal	-0.134	***	-0.062	*	[-0.062, -0.033]	Yes	1.920
	(0.027)	[0.0050]	(0.028)	[0.1080]			
Externalizing	0.126	***	0.055	*	[0.025, 0.055]	Yes	1.699
	(0.026)	[0.0052]	(0.028)	[0.1122]			
Internalizing	0.111	***	0.047		[0.022, 0.047]	Yes	1.700
	(0.024)	[0.0051]	(0.026)	[0.0702]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 11,831. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.6: Maternal depression, test scores, and socioemotional outcomes (third grade)

Panel B: Severe depression (CES-D > 15)

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3\tilde{R}, 1\},1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-10.286	***	-1.791		[-1.791, 1.307]	No	0.595
	(1.281)	[0.0108]	(1.355)	[0.2621]			
Reading	-13.091	***	-3.126		[-3.126, 0.557]	No	0.862
	(1.444)	[0.0137]	(1.629)	[0.2603]			
Learning	-0.191	***	-0.033		[-0.033, 0.087]	No	0.571
	(0.039)	[0.0044]	(0.040)	[0.1523]			
Control	-0.124	**	-0.031		[-0.031, 0.005]	No	0.861
	(0.038)	[0.0023]	(0.041)	[0.0923]			
Interpersonal	-0.164	***	-0.066		[-0.066, -0.043]	Yes	2.393
	(0.038)	[0.0035]	(0.042)	[0.0993]			
Externalizing	0.103	*	0.022		[-0.016, 0.022]	No	0.586
	(0.041)	[0.0016]	(0.045)	[0.105]			
Internalizing	0.130	***	0.060		[0.034, 0.060]	Yes	1.928
	(0.039)	[0.0032]	(0.042)	[0.068]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 11,400. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.7: Maternal depression and test scores (eighth grade)

Panel A: Any depression (CES-D > 9)

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3 \tilde{R}, 1\}, 1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-7.108	***	-0.790		[-0.790, 0.340]	No	2.105
	(1.404)	[0.012]	(1.123)	[0.307]			
Reading	-10.319	***	-1.264		[-2.531, -0.717]	Yes	1.889
	(1.627)	[0.016]	(1.562)	[0.307]			

Panel B: Severe depression (CES-D > 15)

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3 \tilde{R}, 1\}, 1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-7.993	***	0.412		[-0.412, 1.447]	No	1.560
	(2.471)	[0.006]	(1.597)	[0.297]			
Reading	-11.477	***	-2.542				
	(2.611)	[0.008]	(2.109)	[0.322]	[-2.542, 0.264]	No	2.468

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N=6,873 (Panel A), N=6,255 (Panel B). Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.8: Chronic maternal depression, test scores, and socioemotional outcomes (third grade)

Panel A: Any depression in both kindergarten and third grade

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3\tilde{R}, 1\}, 1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-12.196	***	-2.072		[-2.072, 1.407]	No	0.607
	(1.383)	[0.0170]	(1.150)	[0.5627]			
Reading	-13.868	***	-5.496	***	[-5.496, -2.181]	Yes	1.534
	(1.809)	[0.017]	(1.578)	[0.4086]			
Learning	-0.268	***	-0.049		[-0.049, 0.031]	No	0.627
	(0.045)	[0.0105]	(0.043)	[0.3078]			
Control	-0.217	***	(-0.062)		[-0.062, -0.015]	Yes	1.271
	(0.046)	[0.0087]	(0.045)	[0.0204]			
Interpersonal	-0.231	***	-0.010	*	[-0.010, -0.054]	Yes	1.894
	(0.050)	[0.0085]	(0.047)	[0.2042]			
Externalizing	0.244	***	0.093		[0.049, 0.093]	Yes	1.929
	(0.052)	[0.0111]	(0.048)	[0.2982]			
Internalizing	0.182	***	0.064		[0.018, 0.064]	Yes	0.172
	(0.042)	[0.0078]	(0.043)	[0.1322]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 6,078. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.8: Chronic maternal depression, test scores, and socioemotional outcomes (third grade)

Panel B: Any depression in kindergarten and severe depression in third grade

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), [$\dot{R}$]		Controlled Effect $\tilde{\beta}$, (S.E.), [$\widetilde{R}$]		Identified Set [$\tilde{\beta}, \beta^{*\prime}$(min{1.3 $\tilde{R}$, 1},1)]	Exclude Zero?	δ^0 for $\beta = 0$
Math	-13.873	***	-2.506		[-2.506, 1.868]	No	0.589
	(2.004)	[0.0094]	(1.705)	[0.5663]			
Reading	-15.756	***	-6.079	**	[-6.079, -2.184]	Yes	1.446
	(2.674)	[0.0094]	(2.301)	[0.4043]			
Learning	-0.308	***	-0.026		[-0.026, 0.088]	No	0.235
	(0.065)	[0.0052]	(0.057)	[0.3153]			
Control	-0.231	***	-0.036		[-0.036, 0.046]	No	0.045
	(0.066)	[0.0037]	(0.067)	[0.2046]			
Interpersonal	-0.283	***	-0.112		[-0.112, -0.042]	Yes	1.463
	(0.066)	[0.0049]	(.067)	[0.1976]			
Externalizing	0.232	**	0.052		[-0.014, 0.052]	No	0.803
	(0.072)	[0.0039]	(0.067)	[0.2874]			
Internalizing	0.266	***	0.143	*	[0.108, 0.143]	Yes	2.349
	(0.062)	[0.0062]	(0.063)	[0.1369]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 5,225. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.9: Chronic maternal depression and test scores (eighth grade)

Panel A: Any depression in third grade and any depression in eighth grade

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3\tilde{R}, 1\},1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-10.610	***	-0.490		[-0.490, 2.941]	No	0.144
	(3.046)	[0.0121]	(1.408)	[0.6421]			
Reading	-11.213	***	-0.346		[-0.346, 3.861]	No	0.083
	(3.213)	[0.0087]	(1.676)	[0.5823]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 4,359. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Table 2.9: Chronic maternal depression, test scores, and socioemotional outcomes (eighth grade)

Panel B: Any depression in third grade and severe depression in eighth grade

	(1)		(2)		(3)	(4)	(5)
Dependent Variable	Baseline Effect $\dot{\beta}$, (S.E.), $[\dot{R}]$		Controlled Effect $\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		Identified Set $[\tilde{\beta}, \beta^{*\prime}(\min\{1.3\tilde{R}, 1\},1)]$	Exclude Zero?	δ^0 for $\beta = 0$
Math	-9.506	*	1.578		[1.578, 5.402]	Yes	-0.389
	(4.436)	[0.0027]	(1.782)	[0.6492]			
Reading	-9.272		0.932		[0.932, 4.778]	Yes	-0.236
	(5.025)	[0.0017]	(2.787)	[0.5725]			

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N= 4,110. Results of the uncontrolled models (Column 1) are from OLS regressions. Results of the controlled model are from OLS regressions that include child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school). Results in columns 3and 5 are computed using Oster's (2015) Stata code, psacalc.

Chapter 3

Morning Bell: The Effect of Changing School Start Time on Child Test Scores, Behavior, and Learning Outcomes

3.1 Introduction

When an elementary school begins to ring the starting bell at an earlier time, can children compensate simply by going to bed earlier? The primary argument in favor of "no" to this question is biological-- there is a delay in the timing of sleep onset and awakening (Thorleifsdottir et al. 2002; Carskadon and Acebo 2005; Wolfson and Carskadon 1998; Carskadon et al. 1998; Crowley, Acebo, and Carskadon 2007; Gau and Soong 1995) and an inverse relationship between age and hours spent sleeping each night (Frederiksen et al. 2007; Drake et al. 2003). One aspect of puberty is a change in the body's production of melatonin, the sleep-inducing hormone produced by the brain (Randall 2012). Whereas adult brains begin to release melatonin at 8 p.m. and reach peak production at 4 a.m., adolescents (ages 10-19) do not begin producing it until 11 p.m. and reach peak levels at 7 a.m. (Carskadon 2011). Thus, a child entering puberty cannot simply go to bed and get out of bed earlier to accommodate an earlier school start time, as their bodies are naturally fighting against them to stay awake later and rise later.

As with many questions in education, the relationship between learning and start of school day would benefit from more rigorous research. Much of the focus in prior research has been on the effect of changing school start time on the academic performance of older children (Wahlstrom 2002; Wahlstrom et al. 2014; Carrell, Davis, Maghakian, and West 2011; Hinrichs 2011; Edwards 2012) or has been limited to changes within a particular school district (Wahlstrom 2002; Wahlstrom et al. 2014). This study uses nationally representative panel data that follows the same students over time and contains information on start time for several hundred schools and several thousand children. The survey's inclusion of a rich set of individual-level, family-level, and school-level observed characteristics lends itself to models that are able to account for individual-specific, time-invariant, unobserved heterogeneity.

Results from this study contribute to this literature in three key ways: First, this paper demonstrates that changing school start time has negative implications for younger students, too. I find that pubescent-aged (10-12 years old) elementary school children attending schools that shifted their start time 60 or more minutes earlier had lower math scores (girls) and lower reading scores (boys). Second, earlier start time affects non-academic performance of children as well, demonstrated by increased socioemotional problems (both genders) in students whose schools began 60 or more minutes earlier. Third, relatively smaller movements in school start time did not impact elementary school-aged children. School schedules moving either 1 to 29 minutes earlier or any amount later had no effect on children's test scores or socioemotional outcomes.

3.2 Background

The relationship between school start time and a child's academic and classroom-related performance is not direct. Changes in start time also impact sleep patterns, which then affect learning processes of children. Therefore, in order to understand how a change in school start time might impact how a child performs in the classroom, it is important to understand how sleep preferences are influenced by biology, how sleep impacts child learning and behavior, and how school scheduling impacts sleep.

The sleep-wake cycle, or circadian rhythm, regulates production of melatonin, the hormone that induces sleep. Circadian rhythms begin to change during puberty, causing children to go to bed later at night and rise later in the morning (Cardinali 2008; Crowley, Acebo, and Carskadon 2007; Wolfson and Carskadon 1998). The National Sleep Foundation recommends children aged 6-13 receive between nine and eleven hours of sleep each night. When these recommended sleep durations are not met by children, both academics and socioemotional outcomes are negatively affected. Sleep surveys have found older children are more likely than younger children to report daytime sleepiness, unplanned naps, and inadequate sleep (Carskadon et al. 1993; Gau and Soong 1995). Reduced sleep is also associated with a higher body mass index in children (Magee, Caputi, and Iverson 2013; El-Sheikh et al. 2014). Additionally, increased sleep for middle school-aged students all the way through college-aged students positively impacts test scores (Edwards 2012; Wolfson and Carskadon 2003; Carrell, Davis, Maghakian, and West 2011).

One cause of inadequate sleep is a child's school schedule. While circadian rhythms shift to make adolescents more alert at night and less alert in the morning, school schedules are often at odds with this natural rhythm. School start time becomes earlier as the level of education increases, usually to accommodate transportation costs within a district (Fuenschuh 2009) or to free up time for after-school activities (Wolfson and Carskadon 2005)

Literature suggests that when schools begin earlier, children are not able to compensate by going to bed earlier. As a result, they are sleepier during the school day, as shown by studies linking earlier start times and sleep-deprived high school students (Randall 2012; Hansen et al. 2005; Dexter et al. 2003). Sleep-deprivation matters because it is correlated with lower grades (Curcio, Ferrara, and DeGennaro 2006), lower standardized test scores (Wahlstrom 2002), and lower college entrance exam scores

(Carskadon, Viera, and Acebo 1993). The literature also suggests that schools that do the opposite in terms of scheduling also have the opposite effects, with associations between schools that start later and more hours slept (Drake et al. 2003), reduced daytime fatigue (Epstein, Chillag, and Levie 1998), fewer depressive feelings (El-Sheikh et al. 2014), and improved mood and behavior (Owens, Belon, and Moss 2010).

More recent research has focused on moving beyond association between school start time and academic performance with stronger study methodologies. Although not causal in nature, Wahlstrom et al (2014) uses a pre-post design to show that later start times had positive effects on student achievement at the high-school level. Pre-post models do not account for changes in student achievement in schools that did not change start time, so the gains may have been determined by peer and teacher effects, or changes occurring in the schools in addition to the changes in start time. For example, course evaluation, course selection by students, and possibility that high-achieving students were selecting into these school districts all may have caused these results.

Other studies have been able to control for these peer and teacher effects in their study design and produced causal estimates of the role of start time on academic outcomes. Carrell, Davis, Maghakian, and West (2011) took advantage of two policy changes in freshman student schedules at the US Air Force Academy, coupled with the fact that students were randomly placed in courses. This allowed them to evaluate the causal effect of start time on academic achievement at the university level. Their results showed that delaying start time by 50 minutes leads to a one standard deviation improvement in academic achievement. However, using a quasi-experimental design, Henrichs (2011) finds no impact of later start times on college entrance exam scores for high school-aged children (Hinrichs 2011).

The association between puberty and a later sleep-wake cycle is well-known, so much of the prior research on changing school start time focuses on how the resultant reduction in total sleep hours for older students (middle-school aged, high school-aged, and college-aged individuals) affects learning. However, in the U.S., middle school typically begins in grade 6 (age 11), meaning older-aged elementary school children (grades 4 and 5) are being overlooked or grouped with their younger pre-pubescent counterparts whose sleep-wake cycles are more in line with those of adults. The average onset of puberty in the U.S. occurs in grades 3 and 4, between the ages of 9 and 10 for both girls (Biro, Greenspan, and Galvez 2013) and boys (Herman-Giddens, Steffes, and Harris 2012).

Whereas prior research suggests that earlier school start times have no effect on pre-pubescent students (Edwards 2012), this study recognizes that the sleep-wake patterns of a first-grade child are potentially quite different from that of a fifth-grade child. Therefore, shifting school start time might impact older elementary school-aged children in a more similar fashion to those of middle and high school-aged children than younger elementary school-aged children. By examining how shifting school start time in a child's fourth or fifth grade year impacts academic and behavioral outcomes for older elementary school-aged children (fifth graders), results from this study indicate that elementary school administrators should not apply a one-size-fits-all approach to start time decisions.

3.3 Conceptual Framework

In many households with elementary school-aged children, the amount of sleep a child receives at night is a function of several inputs, including bedtime (e.g., preferences over the time at which this begins and strictness of parental bedtime rule enforcement), number of and timing of extracurricular activities (e.g., if the child has soccer practice until 9 p.m. and if that is only one night a week or several), number of and timing of social activities (e.g., spending time at friends' homes and using technology for communication with peers), school work, desired sleep duration, wake time, and biology. When a school schedule shifts earlier, in order to produce the same number of hours of sleep in a night, the child could shift bedtime by an equal amount of minutes that wake time moved to accommodate the earlier start time. However, biology has a larger influence on bedtime than other factors (Crawley, Acebo, and Caskadon 2007), suggesting there is not a one-to-one tradeoff between bed time and wake time.

Therefore, early start times potentially reduce the total amount of sleep received by older children. This is important because sleep is a determinant of health (Moore et al., 2002), and health is a determinant of a child's human capital (Case, Fertig, Paxson 2005). In this study, the child's level of human capital is measured by test scores and socioemotional outcomes in this study. Conceptually, an earlier (later) start time coupled with a later sleep-wake cycle would lead to children getting less (more) sleep and negatively (positively) affect human capital formation.

In addition to affecting the health component of a child's human capital production function, sleep has the potential to influence several other inputs. Thus a change in school start time, which impacts sleep, also influences these other factors of a child’s human capital production. For example, in order to maximize performance on a reading exam, the child should not be hungry. As the argument has been made that earlier start times make it more difficult for (pubescent-aged) children to get up in the morning, this might mean that the child also has less time to eat a good breakfast (or any breakfast) before heading off to school. Another input affected might be the amount of quality time the child has to spend with a parent (or parents) in the morning, as the child is once again sleepier and/or reducing the amount of total time spent preparing for school. Similarly, amount of time and/or quality of time spent on homework each night might diminish in response to having to go to bed at an earlier time. Driven by the biological processes of melatonin production, a child whose bedtime moves 60 minutes earlier might still fall asleep around the same time as they did before the change, leading to lower quality sleep (i.e., fewer REM cycles) and negatively affecting the reading exam score.

There are a number of reasons schools begin earlier, such as bus schedules, extra-curricular activities, and parental preference to align with work schedules. Since these data do not discern why schools made changes to their start times, endogeneity is a concern. In other words, human capital measurements (test scores and soioemotional outcomes) might be correlated with the unmeasured characteristics from the underlying equation predicting start time, and start time might be correlated with unmeasured characteristics from the equation predicting human capital. By estimating a reduced form equation, the system has been solved for this endogeneity and the coefficients from the models represent the treatment effect among the treated (those whose school schedule changed).

3.4 Data

3.4.1 Sample

The Early Childhood Longitudinal Study-Kindergarten Class of 1998-1999 (ECLS-K) is a large, nationally representative, longitudinal study of students who entered kindergarten in the fall of 1998. The ECLS-K was sponsored by the National Center for Educational Statistics (NCES) to follow roughly 22,000 kindergartners from 1,000 schools upon entry and through completion of 8th grade. Data collection began in the fall of kindergarten (1998) and follow-up surveys were administered in spring of kindergarten (1999), fall of first grade (1999), spring of first grade (2000), spring of third grade (2002), spring of fifth grade (2004), and fall of eighth grade (2007). Data are sourced from a collection of parent, teacher, and school administrator interviews as well as child assessments (Tourangeau et al. 2009).

Within each ECLS-K school, between 12 and 24 kindergartners are sampled (not necessarily all in one classroom), leading to a primary sample of 21,409. The ECLS-K was designed to follow the same children over time, tracking them as they move through the U.S. educational system. Start time was only measured in third and fifth grade, reducing the sample to just over 9,000 students with information for both years. Lastly, students that transferred schools were not included in the analyses, as these students may have been changing schools for personal reasons (e.g., bullying) and that would have also affected the measures of human capital.

Since these data do not provide a reason for why a student transfers schools, one concern regarding the exclusion of transfer students is that they may have simply moved to the next "phase" of schooling—from elementary school to middle school. However, in the U.S., most middle schools begin in sixth grade, and based on estimates produced by alternative school years in the ECLS-K, I believe that this is also the case with my sample. From first to third grade, 18.1 percent of students transferred schools. Since a change in school format does not typically occur during these two academic years (e.g., from elementary school to middle school), this estimate is a good baseline indicator of the share of students that transfer schools for personal reasons during elementary school. Next, I calculated the transfer rate of ECLS-K students between fifth and eighth grade and found it to be 79.2 percent. This period of time **does** correspond to one when school type typically changes (from elementary school to middle school or junior high). Between the two academic years in this study, third to fifth, the transfer rate was 18.6 percent, nearly mirroring the first to third grade rate. Therefore, it is likely that students who transferred during the two years in my study (third and fifth grade) were not doing so because of a natural progression of type of school but for personal reasons that may have biased results either positively or negatively (e.g., a child's start time moved earlier because they transferred schools on account of being accepted at a school for the gifted or because they were being bullied), justifying exclusion from the sample. The final analytic sample included 7,420 students with full data available in both third and fifth grade.

Analyses are weighted using the panel weights, stratum, and primary sampling units designated for use with the third through fifth grade sample. The ECLS-K followed a multistage probability sample, such that in kindergarten, the primary sampling units (PSUs) were geographic areas of counties (or groups of counties). Within each PSU, both public and private schools were included in the sample frame. Selected public schools had a minimum of 24 students whereas private schools needed 12. Schools with fewer than these minimum numbers were clustered together, and were then selected with probability proportional to size. The final-stage units included children within schools, leading to standard errors are clustered at the school level for analytical purposes.

3.4.2 Key Measures

The primary variable of interest is the time at which the academic day began in schools. Start time information was collected via the administrative component of the survey. The times were then recoded in the ECLS-K into 15-minute increments. As shown in Table 3.1, the share of the sample that began school before 8:15 a.m. grew from 37.6 percent in third grade to 40.7 percent in fifth grade. Between the two waves of data collection, 765 schools maintained the same start time between the two waves, 105 began later, and 205 began earlier (Table 3.2). At the student level, this translates to roughly 15 percent of the sample attending schools that moved their start time earlier between 3rd and 5th grade, 11 percent attending schools that moved their start time later between 3rd and 5th grade, and the remainder attending schools that did not change their start time.

This variation in start times provides the identification to test the effect of changing school start times on child outcomes. The ECLS-K recoded all administrative start time data into 15-minute increments, with the earliest being before 7:45 a.m. and the latest 9:00 a.m. or later (Table 3.1). By calculating the change in start time between third in fifth grade, the resultant changes in start time were up to 90 minutes (or more) earlier and up to 60 minutes (or more) later. To insure adequate sample sizes in analyses, categories were collapsed to 1-29 minutes earlier, 30-59 minutes earlier, and 60 or more minutes earlier. Additionally, although fewer students were impacted by later start times and most in this group (92%) attended schools moving 1-29 minutes later, a fourth categorical outcome for start time captures the effect of any later start time.

All child outcomes are collected in the spring of the academic year. Test scores are measured by math and reading item response theory (IRT) exams. Math scores have a potential range of 0-174 and reading scores range from 0-212. These scale scores estimate children's performance on the entire set of assessment questions by using the pattern of right, wrong, and omitted responses to the items administered in an assessment. IRT scores have several advantages over raw number-right scoring, including compensation for the possibility of low-ability children guessing difficult items correctly, reduction of score distortions caused by omission, and are most appropriate for longitudinal measurement of gains in achievement over time.

The teacher questionnaire contains three self-administered portions and is weighted to compensate for non-response and differential probability of selection. This questionnaire asks teachers to assess children's ability in five non-cognitive areas, with several questions designed to evaluate performance in each category. The reported outcomes have a continuous range of 1.0-4.0, with 1 indicating never, 2 indicating sometimes, 3 indicating often, and 4 indicating very often. Included in these non-cognitive measures are internalizing problem behavior (measurement of loneliness, sadness, low self-esteem, and anxiety), externalizing problem behavior (indicates the frequency with which a child fights, gets angry, argues, acts impulsively, and disturbs ongoing activities), ability to learn (attentiveness, task persistence, eagerness to learn, flexibility, and organization), self-control (child's ability to control behavior, control temper, accept peer ideas for group activities, and appropriateness of response to pressure from peers), and interpersonal skills (ability to form and maintain friendships, get along with others who are different, comfort and help other children, express feelings, ideas, and opinions in a positive manner, and show sensitivity to others' feelings). Higher scores are more desirable for ability to learn, self-control, and interpersonal skills whereas lower scores are more desirable for internalizing and externalizing problem behavior.

Plotting weighted mean outcomes by school start time suggests that later start times are associated with higher test scores (Figures 3.1, 3.2, and 3.3). Across reading and math assessments, regardless of gender (Columns (B) and (C)), scores in both third and fifth grades are lowest for students attending schools that began before 8:15 a.m. For the behavioral measures, the trends are somewhat less straight-forward. In four of the six graphs, the worst score (highest numerically) for internal problem behavior occurred prior to 8:30 a.m., whereas this was the case only half of the time for external problem behavior. In the remaining three non-cognitive measures the relationship between start time and performance is clearer. Assessments of ability to learn, self-control, and interpersonal skills, show lowest average scores occurring prior to 8:30 a.m. seventeen of eighteen times, and in ten cases, before 8:00 a.m.

The information provided in Tables 3.1 and 3.2, and Figures 3.1-3.3 suggest earlier school start time is correlated with lower test scores and reduced performance in non-cognitive measures. Yet nearly one in six children attended a school that moved its start time earlier in their fourth or fifth grade year of schooling. As this coincides with a point in time during which children's sleep-wake cycles are naturally beginning to shift to favor later bedtimes and later wake times, these administrative changes to the start of the school day are potentially disruptive to human capital formation of children. Through empirical testing, I am able to discern first if there are impacts on academic and non-cognitive outcomes, and second if these effects are statistically different from zero.

3.4.3 Time-varying covariates

Models control for changes in time-varying observable characteristics that may have impacted cognitive and non-cognitive outcomes during the study period. At the family level, these include changes in income, father's employment status, mother's employment status, family type (single- or two-parent), if the child was asked to do household chores during the week, number of nights the family eats dinner together per week, and number of children under age 18 in the house.

The ECLS-K also includes rich information regarding the schools, allowing for control of measures that might be influencing child outcomes between third and fifth grade. For example, whether or not either teacher turnover or over-crowdedness were indicated as being a problem in the school, number of in-school gym classes per week, if less than 10% of the school was comprised of minority students, share of students receiving free lunch, and participation in the U.S.D.A. school breakfast program. The survey also collects characteristics of the neighborhood in which the school resides, allowing for construction of a neighborhood problem index to measure changes in gang activity, drug problems, crime, and tension. Time-invariant characteristics (e.g., birth weight, birth order, and race/ethnicity) are removed during the estimation process and thus not included as control measures in the analyses.

3.5 Empirical Model

The effect of a change in school start time is identified by examining how a child's change in test score or non-cognitive outcome differs with incremental changes in school start time. By using children as their own control measures, a fixed effects methodology employs variation from changes in school start times between the two time periods but within the same school. The difference in outcomes between the two time periods is regressed on the changes in school start time. By doing so, any time-invariant constant effects resulting from the school or home environment drop out of the model, including unobserved characteristics, and identification hinges on the within-school variation in start times. As long as the factors predicting the adoption of changes in start time are school-specific and unvarying over the fairly short time period considered here, the model will generate unbiased estimates of the effect of changes in school start time on student outcomes.

For all multivariate linear regression models the treatment is the change in school start time. A child may have attended one of the 765 schools that did not change their start time over the period, one of the 205 schools that began earlier, or one of the 105 schools that moved to a later start time between the 2001-2002 and 2003-2004 schools years. Models define the control group as those students attending schools that did not change their start time over the study period. The treatment group consists of all children attending a school whose start time changed between the 2001-2002 and 2003-2004 school years. The final categorical start time variable reflects 4 potential changes: 1-29 minutes earlier, 30-59 minutes earlier, 60 or more minutes earlier, and any amount of time later. All coefficients are interpreted as the change in the outcome for the treatment group (e.g., those whose schools started 30-59 minutes earlier) relative to the change in the outcome for the control group (those whose schools did not change times).

Models take advantage of the panel nature of the data and use *changes* in these outcomes between third and fifth grade. If the change is negative for test scores, this indicates the child's performance in that area became worse between the two periods. For internal and external problem behavior, a negative change indicates the outcome was better for the child between the two waves[ii] as it demonstrates movement away from the teacher reporting the child was very frequently exhibiting the poor behavior (e.g., hitting).

The general model of the relationship between school start time and child outcomes is:

$$Outcome_{ijt} = S_{it}\,\beta + \delta X_{it} + \varepsilon_{ijt} \qquad (3.1)$$

where $Outcome_{ijt}$ is the test or socioemotional score of individual *i* at school *j* in time *t*, S_{it} is a measure of school start time for student *i* in time period *t*, X_{it} is a set of other explanatory variables, and ε_{ijt} is an error term.

The standard assumption in panel data analysis is to write the error as:

$$\varepsilon_{ijt} = \theta_i + \gamma_j + \varphi_{ijt} \quad (3.2)$$

[ii] A simplified example would be if the child often acted up in third grade and received an external behavior score of 3.5. If in fifth grade he was less likely to misbehave in class and received a score of 1.5, this 2 point decrease in his external behavior score would be indicative of improved behavior.

where θ_i and γ_j represent individual- and school-level unobserved time-invariant effects. The concern is that there is selection on these unobservable characteristics that might be biasing the effect of the treatment (start times) on outcomes (scores). However, it is possible to consistently estimate the effect, β, by first-differencing the data. This allows replacement of the assumption that $E(\varepsilon_{ijt}|S_{it}) = E(\theta_i + \gamma_j + \varphi_{ijt}|S_{it}) = 0$ by the weaker assumption $E(\varphi_{ijt}|S_{it}) = 0$. A first-differences methodology corrects for the bias by allowing time-invariant unobservable individual (θ_i) and school measures (γ_j) be "differenced out" in the estimation.

The following equations illustrate how the differencing process leads to unbiased estimates of β. Equation 3.3 incorporates the new error term assumption:

$$Outcome_{ijt} = \beta S_{it} + \delta X_{it} + \theta_i + \gamma_j + \varphi_{ijt} \tag{3.3}$$

The coefficient of interest is β, the effect of school start time on scores. The concern is that start times might be correlated with individual-level and school-specific unobservable characteristics, and these will drive the variation in scores. For example, if students with low unobservable ability are clustered in one type of school, and test score losses are related to ability, then the low scores of an individual could be attributed to the school type as opposed to the nonrandom distribution of students. Therefore estimating equation (3.3) using pooled OLS would lead to biased estimates of β. Instead, this paper estimates β using techniques that take advantage of the panel nature of the ECLS-K data. Although the study periods are two years apart, it is assumed that relevant unobserved individual- and school-level characteristics correlated with both child outcomes and scores are constant over the time period. By taking first differences equation (3.3) becomes:

$$Outcome_{ijt} - Outcome_{ijt-1} = \beta(S_{it} - S_{it-1}) + \delta(X_{it} - X_{it-1}) + (\theta_i - \theta_i) + (\gamma_j - \gamma_{j-1}) + (\varphi_{ijt} - \varphi_{ijt-1}) \qquad (3.4)$$

or

$$\Delta Outcome_{ijt} = \beta \Delta S_{it} + \delta \Delta X_{it} + \Delta \varphi_{ijt} \qquad (3.5)$$

The goal is to identify the effect of school start time, β, by studying how a child's score differs with changes to the school start time over the two time periods. Estimating Model 3.5 using this first differences approach reduces the problem of unmeasured characteristics by using children as their own control and takes advantage of the variation that comes from changes in school start time over time but within the same school. Time-invariant characteristics (θ_i and γ_j) drop out of the model and identification hinges on the within-school variation in start times. The model will generate unbiased estimates of the effect of changes in school start time on student outcomes as long as factors predicting the adoption of changes in start time are school-specific and unvarying over the fairly short time period considered here.

Between the two periods, schools might have moved their start times earlier, later, or remained the same. Therefore, Equation 3.5 is more precisely defined as:

$$\Delta Score_{ijt} = \beta_1 \Delta earlier\_1\_to\_29_{it} + \beta_2 \Delta earlier\_30\_to\_59_{it} + \beta_3 \Delta earlier\_60\_or\_more_{it} + \beta_4 later_{it} + \delta \Delta X_{it} + \Delta \varphi_{ijt} \qquad (3.6)$$

The coefficients from β_1, β_2, and β_3 represent the change in outcomes between students attending schools that moved earlier (by 1-29, 30-59, and 60 or more minutes, respectively), relative to the change in outcomes between students attending schools that did not change start times. β_4 corresponds to the change in outcomes between students attending schools that began later, relative to those attending schools that did not change start time, with standard errors clustered at the school level.

3.6 Results

Estimates of the other covariates included in the models were mostly statistically insignificant (Table 3.3), suggesting both little variation in the time-varying covariates and that these impacts of school and family characteristics were not driving the results. The latter strengthens the argument that the fixed-effects modeling does remove the constant impacts of the time-varying characteristics outlined above.

Regression-adjusted results (Table 3.4) report the coefficient estimates of earlier start time (β_1, β_2, and β_3) and later start time (β_4). In general, findings suggest that moving to an earlier start time leads to reductions in both cognitive and non-cognitive outcomes. More specifically, girls attending schools that shifted start time by 60 or more minutes earlier at the beginning of their fourth or fifth grade years of schooling saw changes in math scores of -5.67 points, relative to girls attending schools that did not change times (Table 3.4). This point-estimate represents a 0.5 standard deviation reduction in math score. Additionally, girls whose schools began 1-29 minutes earlier also had 0.1 point reductions in teacher-assessed interpersonal skills, representing a decrease of 0.2 standard deviations.

Compared to boys attending schools that did not change start time, boys attending schools beginning 60 or more minutes earlier scored 4.43 points (0.3 standard deviations) lower in reading (Table 3.4). These boys also fared worse in the teacher's assessments of their classroom behavior. Boys whose schools began 60 or more minutes earlier had 0.17 point increases in externalizing problem behavior scores, indicating a movement away from the teacher reporting that he/she seldom saw the boy acting out or being aggressive and toward a report of very often. Compared to boys attending schools that did not change start time, this is a 0.3 standard deviation increase in negative classroom behavior. These boys also experienced losses in teacher-assessed self-control, evidenced by a 0.44 point (0.5 standard deviation) reduction in scores, relative to the control group.

Narrowing the subpopulation to only children whose schools began earlier and the new time was before 8:15 a.m. and children whose schools began later and the new time was 8:15 or later, the same pattern emerges (Appendix C). Students attending schools beginning 60 minutes or more earlier, with the start time in fifth grade being 8:15 or before, had reductions in math scores of roughly five points (girls), readings scores of 4 points (boys), increases in external problem behavior (boys), and reduced self-control (boys). Additionally, when the population is defined this way, significant results occur with smaller changes in start time, with reductions in interpersonal skills at start times between 1-29 minutes earlier (overall and girls) and 30-59 minutes earlier (overall and boys).

3.6.1 Robustness testing: replication of the study in a period prior to the time change

Because these data are a panel that follow the same students over time, by estimating Model (6) using the same children but with the outcomes measured as changes between the children's first and third grade years of schooling, it can strengthen the argument that results are due to changing school start time and not unmeasured variables. If the same children negatively impacted by earlier start times also demonstrated worse performance in an earlier period of time, this suggests that they might simply be lower-performing individuals. As demonstrated in Table 3.5, by replicating the model on the same children but in a period of time prior to their school star time changing, nearly all of the outcomes show statistically insignificant differences, relative to the control groups. However, girls whose schools moved later between third and fifth grade initially showed an improvement in internalizing problem behavior (Table 3.4), but an improvement of roughly the same size was also observed between grades first and third. Therefore, this study cannot conclusively suggest that later start times lead to improved behavior for girls.

3.6.2 Limitations

The ECLS-K provides a rich source of panel data with numerous advantages for this sort of study. However, these data are not without their limitations. First, although results from this study are more generalizable because data are at the national level, it is still an observational study, not a randomized trial. Therefore, internal validity remains a concern. However, by examining how a change in start time between third and fifth grade affected the differences in child outcomes for the same child, all unobserved fixed factors correlated with *both* start time and child outcomes are controlled. For example, it may have been the case that administrators in some schools did not wish to begin earlier but were mandated by the district to do so in order to accommodate one fleet of buses. These same administrators might also have elected to increase classroom sizes. If this were the case, then schools with earlier start times may have had reduced child outcomes but the start time would not have played a causal role.

Additionally, despite having large sample sizes for the entire sample, subgroup comparisons (girls and boys by various treatment levels) are limited in terms of statistical power. Larger sample sizes of students impacted by relatively big start time changes would help. Results from placebo tests (Table 3.5) strengthen the findings that small subpopulations are not swaying results—in other words, the group of students attending schools that started 60 minutes earlier are not simply "different" from students attending schools that did not change start time and would not have had changes in outcomes regardless of changes in start time.

A limitation in the model is the inability to control for unobserved characteristics correlated with either start time *or* child outcomes that may have changed between third and fifth grade. For example, the public use ECLS-K data file does not measure either the end time of the school day or the length of the school year, so the motivations for why the school moved its start time, and the potential direction of the bias, are not clear. A school might have been beginning the school day a few minutes earlier in order to lengthen the number of minutes in the day, or they may have begun the day earlier in order to end the day earlier and make more time for extra-curricular activities. Based on data from the U.S. Department of Education, the average length of an elementary school day increased six minutes between the 2001-2002 and 2003-2004 academic years. It is unknown if the schools that moved earlier were trying to add time, and if so, if they only began earlier of if they also released students a bit later to add time. However, longer days don't necessarily translate to student gains. The research surrounding length of school day and student learning indicate is not the additional time that makes a difference but rather how the extra time is used (Aronson, Zimmerman, and Carlos 1999). Therefore, although the overall length of school days lengthened slightly over the study period, and the ECLS-K does not provide information for why schools in the study were changing their start time, this does not pose a large threat to the validity of findings. Results from the placebo models (Table 3.5) strengthen the argument that any potential changes in unobserved measures correlated either with start time or with child outcomes are not the primary determinant of the results.

3.7 Discussion

School start time is an issue that has received increasing amounts of attention over the past several years, prompting changes from local district level all the way to national level. Most of these changes have been directed at high school start times. In 1997, the Minneapolis School District was one of the first in the nation to shift start times for high school students later, from 7:15 a.m. to 8:40 a.m. More recently, national policy initiatives like House Concurrent Resolution 176, called for all high schools in the U.S. to begin at 9:00 a.m. or later.

For many schools, the decision to move start times is a function of non-academic inputs—bus schedules, extra-curricular schedules, cost, and parental feedback. If finances within a school district require a shared bus fleet, results from this study demonstrate that elementary schools could potentially move start time a bit earlier in order to accommodate a later high school start time, without negative repercussions in terms of academic or socioemotional outcomes. This is demonstrated by a lack significant findings in child outcomes of children attending schools that moved their start time 1 to 29 minutes earlier.

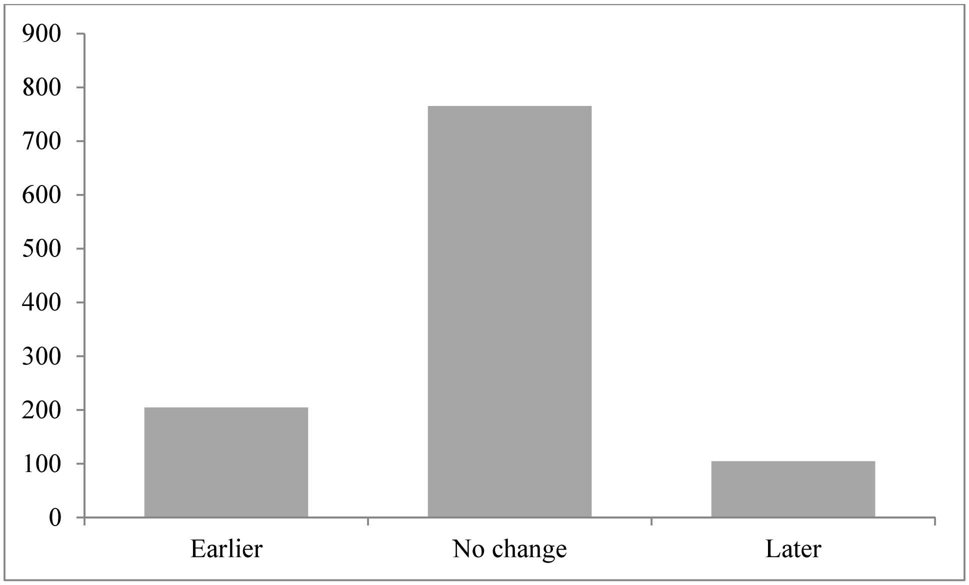
In addition, this research also demonstrates that large changes in start times impact elementary school test scores and non-cognitive outcomes. Schools that have pushed up the start of the day to much earlier times (60 or more minutes) should recognize the potential benefits of reversing (or partially reversing) these policies. If by starting the school day at a later time would result in girls recovering the point differential lost in math and interpersonal skills and boys regaining the points lost in reading, self-control, and externalized problem behavior, these results are similar in magnitude (in terms of standard deviations) to the gains made by having smaller class sizes (Unlu 2005). However, if the school is able to begin slightly later and still share a bus fleet with the high school(s) in the district, the same gains as smaller class sizes might be realized but at a lower per pupil cost.

Table 3.1: Weighted distribution of students by school start time

	Third Grade (%)	Fifth Grade (%)
Before 7:45	2.1	3.4
7:45 to before 8:00	13.2	14.6
8:00 to before 8:15	22.3	22.7
8:15 to before 8:30	17.6	17.1
8:30 to before 8:45	21.3	19.6
8:45 to before 9:00	11.3	10.7
9:00 or after	12.2	11.9

N= 7,420

Table 3.2: Weighted distribution of schools by start time



N= 1,075 schools

Figure 3.1: Weighted outcomes by start time (full sample)

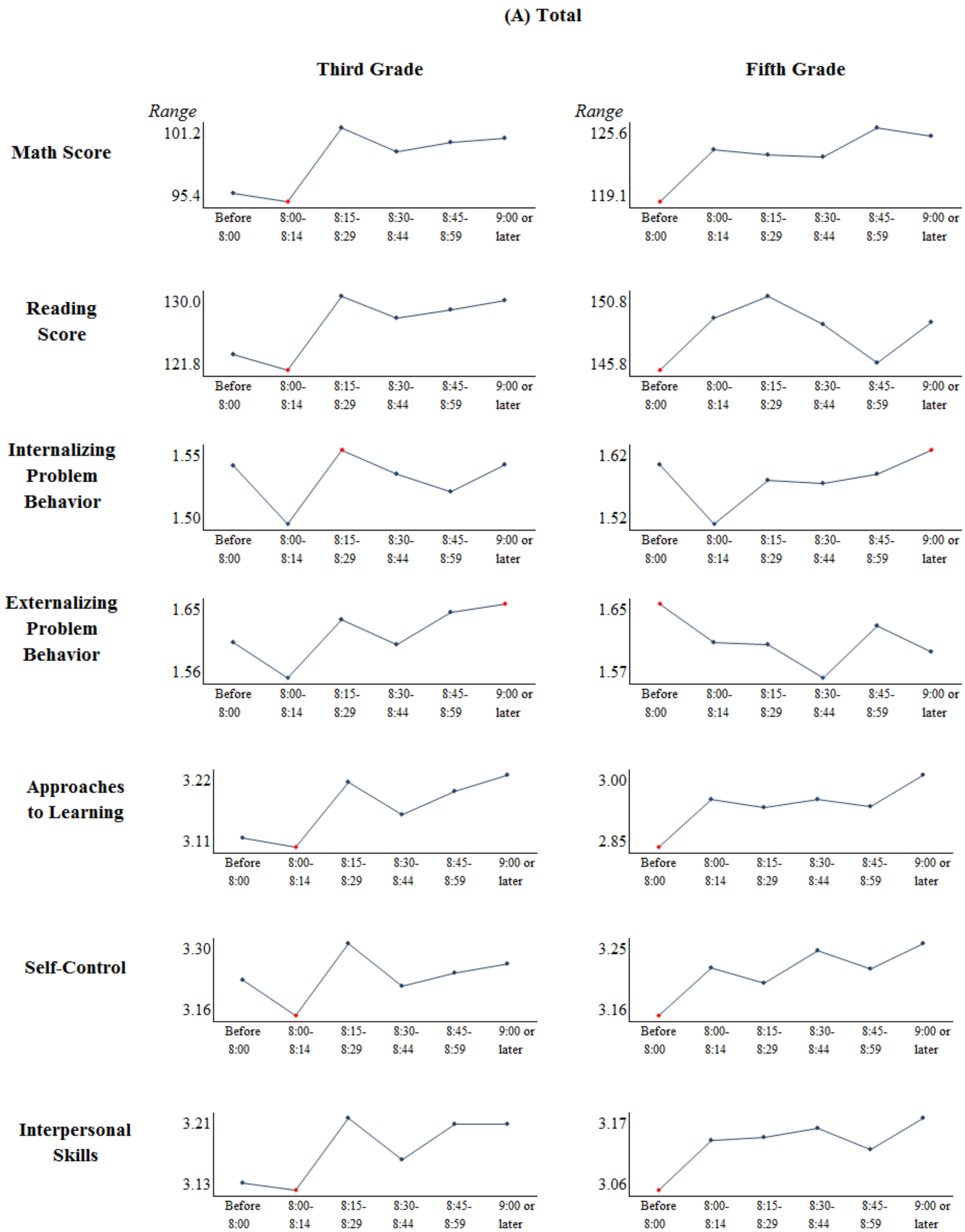


Figure 3.2: Weighted outcomes by start time (boys)

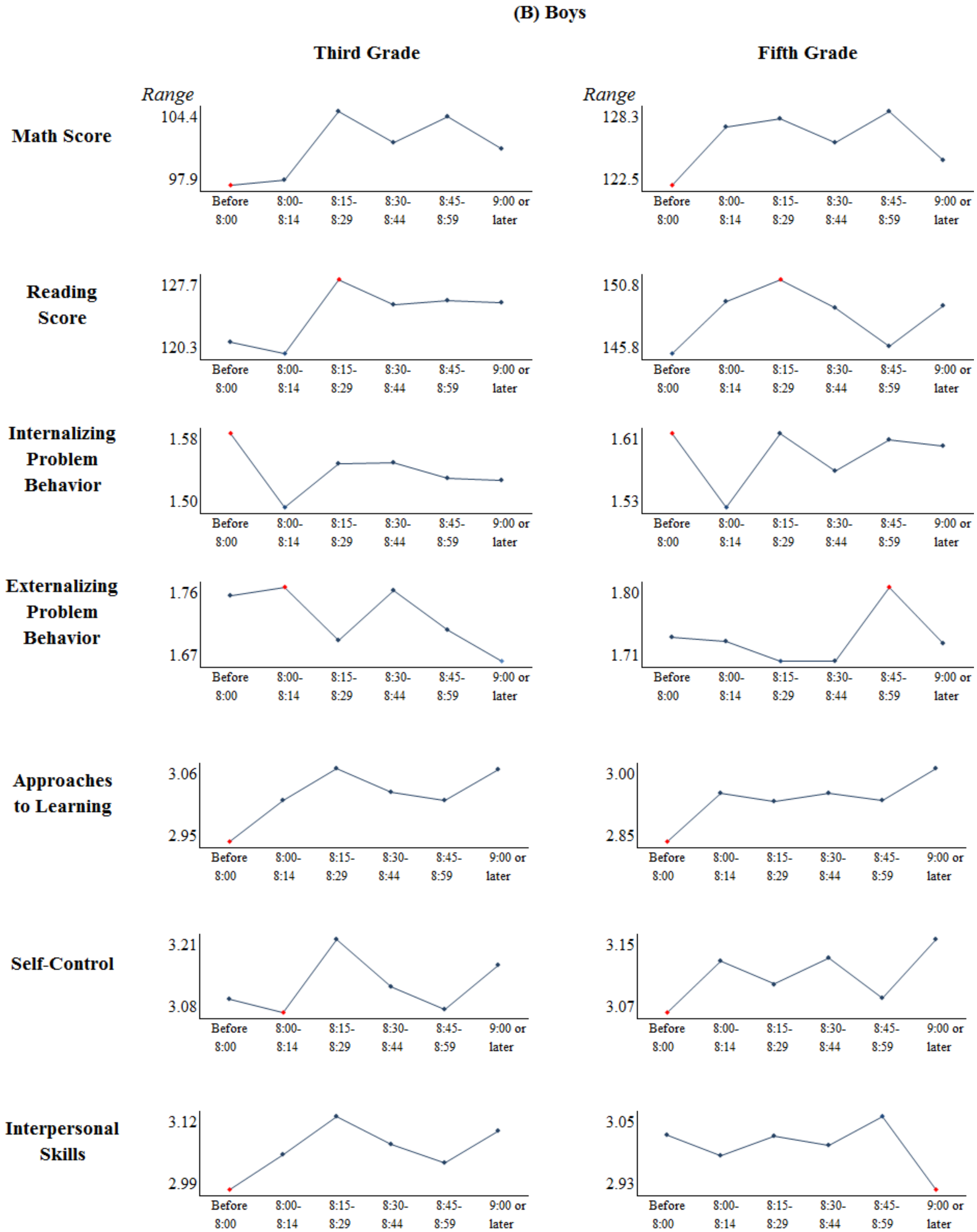


Figure 3.3: Weighted outcomes by start time (girls)

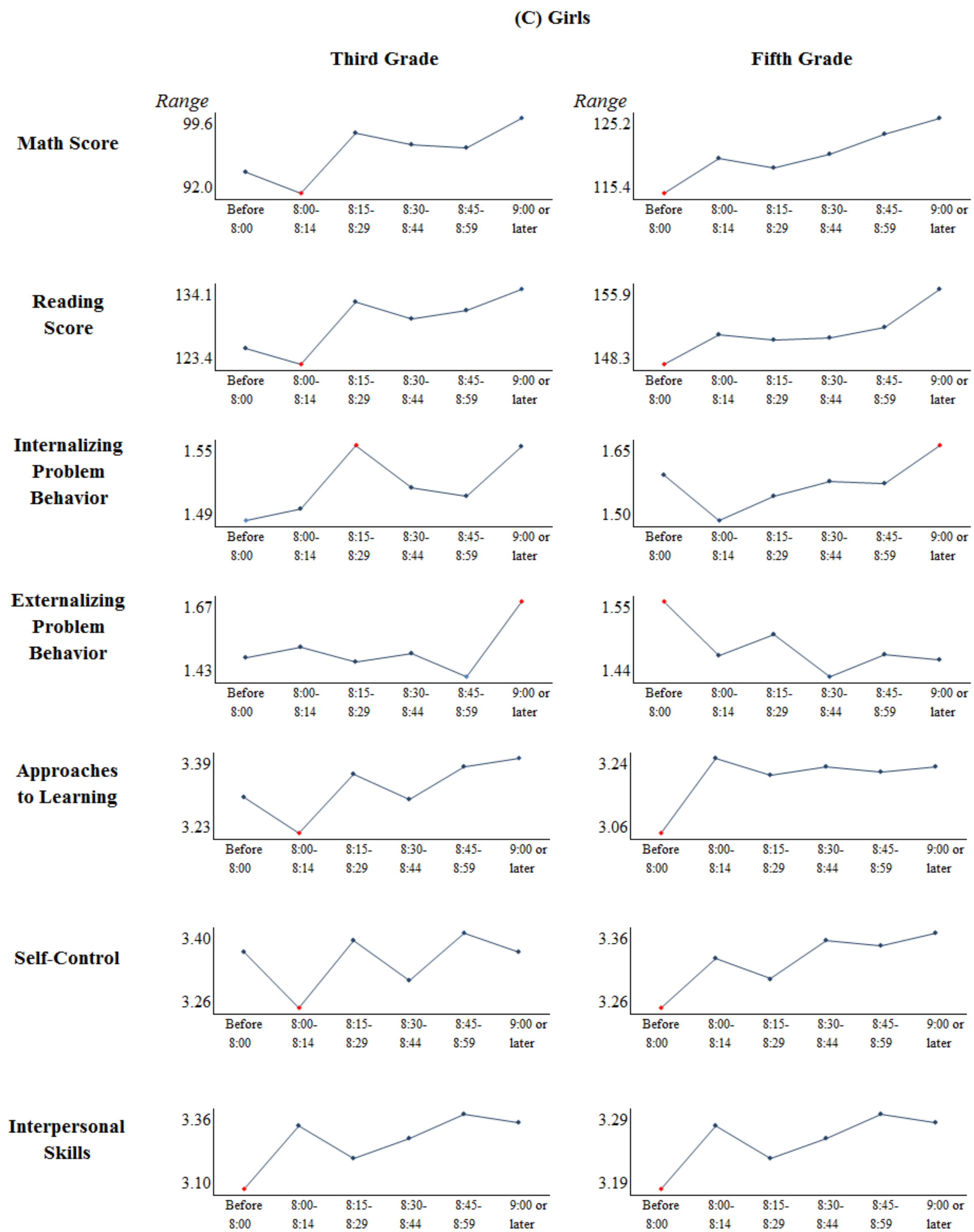


Table 3.3: Changes in observable characteristics by start time

Changes between 3rd and 5th grade in the following:	Same Time[a]	Earlier Time[b]	Later Time[c]	Difference between earlier and same start time Est.		(95% CI)	Difference between later and same start time Est.		(95% CI)
Income									
$0-$25,000	-0.01	-0.02	0.02	-0.01		(-0.04, 0.02)	0.03		(-0.04, 0.10)
$25,001-$50,000	-0.03	-0.02	-0.03	0.01		(-0.05, 0.06)	0.00		(-0.07, 0.07)
$50,001-$75,000	0.00	0.03	0.00	0.03		(-0.02, 0.08)	-0.01		(-0.05, 0.03)
$75,001 and up	0.03	0.01	0.01	-0.02		(-0.07, 0.03)	-0.03		(-0.06, 0.01)
Family characteristics									
Father is employed	0.02	-0.01	0.03	-0.03		(-0.09, 0.04)	0.01		(-0.04, 0.07)
Mother is employed	0.05	0.08	0.05	0.02		(-0.05, 0.09)	0.00		(-0.08, 0.07)
Single-parent family	0.03	0.03	0.03	0.00		(-0.07, 0.06)	0.00		(-0.06, 0.05)
Required to do some chores	-0.05	-0.05	-0.17	0.01		(-0.08, 0.09)	-0.11	***	(-0.09, 0.04)
# of nights /wk family has dinner together	-0.18	-0.24	-0.15	-0.06		(-0.32, 0.21)	0.03		(-0.22, 0.27)
# of children under 18 in the house	0.01	-0.01	0.06	-0.02		(-0.10, 0.06)	0.05		(-0.12, 0.22)
School characteristics									
Minutes of gym classes in school/week	-0.22	-0.47	-0.16	-0.26		(-0.54, 0.02)	0.06		(-0.24, 0.35)
Neighborhood problem index	-0.07	-1.79	1.02	-1.72	**	(-3.21, -0.22)	1.09		(-1.22, 3.39)
Teacher turnover is a problem	-0.01	-0.03	0.01	-0.01		(-0.07, 0.04)	0.03		(-0.10, 0.16)
Problems with crowded school	-0.06	-0.01	0.02	0.05		(-0.05, 0.14)	0.08		(-0.02, 0.18)
Fewer than 10% of the school is minority	-0.02	0.01	-0.01	0.03		(0.00, 0.06)	0.01		(-0.04, 0.06)
Percent of students eligible for free lunch	2.30	0.55	4.33	-1.75		(-4.17, 0.68)	2.03		(-1.56, 5.62)
School participates in free breakfast program	0.02	-0.02	-0.02	-0.04	**	(-0.07, -0.01)	-0.05	**	(-0.08, -0.01)

Notes: [a]N= 5,522; [b]N= 1,101; [c]N=797. *p<.05;**p<.01;***p<.001

Table 3.4: Regression-adjusted impact of changing school start time on child outcomes

	1-29 minutes earlier	30-59 minutes earlier	60+ minutes earlier	Any later
Math score				
All	0.89	-2.22	-3.33	-1.73
	(1.05)	(1.48)	(1.05)	(0.94)
Boys	0.75	-1.09	0.26	-1.61
	(1.64)	(1.55)	(2.39)	(0.97)
Girls	1.02	-3.63	-5.67 **	-1.68
	(1.06)	(1.93)	(2.13)	(1.59)
Reading score				
All	1.17	-1.49	-1.95	-0.70
	(0.94)	(1.55)	(1.38)	(0.79)
Boys	1.56	-3.60	-4.43 ***	-0.97
	(1.38)	(2.18)	(1.17)	(1.07)
Girls	0.78	1.09	0.67	-0.18
	(1.19)	(1.99)	(2.59)	(1.11)
Internalizing problem behavior				
All	0.002	0.02	-0.31	-0.07
	(0.05)	(0.08)	(0.25)	(0.04)
Boys	0.02	0.01	-0.47	-0.01
	(0.05)	(0.11)	(0.37)	(0.05)
Girls	-0.02	0.07	-0.12	-0.13 *
	(0.09)	(0.12)	(0.11)	(0.06)
Externalizing problem behavior				
All	-0.01	0.02	0.08 *	0.06
	(0.05)	(0.08)	(0.04)	(0.04)
Boys	0.003	0.004	0.17 **	0.06
	(0.10)	(0.06)	(0.07)	(0.05)
Girls	-0.02	0.08	-0.04	0.06
	(0.07)	(0.07)	(0.03)	(0.06)
Approaches to learning				
All	-0.07	-0.14	0.31	0.01
	(0.05)	(0.08)	(0.28)	(0.05)
Boys	-0.07	-0.22 *	0.51	0.26
	(.06)	(0.11)	(0.44)	(0.06)
Girls	-0.07	-0.05	0.06	-0.01
	(0.07)	(0.06)	(0.10)	(0.04)
Self-control				

All	-0.06		0.00		-0.17		-0.02	
	(0.05)		(0.06)		(0.22)		(0.05)	
Boys	-.01		-0.10		-0.44	*	-0.01	
	(0.05)		(0.06)		(0.20)		(0.06)	
Girls	-0.12		0.13		0.14		-0.04	
	(0.08)		(0.09)		(0.15)		(0.05)	
Interpersonal skills								
All	-0.12		-0.17	**	-0.07		-0.05	
	(0.06)		(0.06)		(0.19)		(0.05)	
Boys	-0.08		-0.23	**	0.17			
	(0.07)		(0.07)		(0.28)			
Girls	-0.16	*	-0.07		-0.07		-0.13	
	(0.07)		(0.08)		(0.08)		(0.05)	

Notes: N=7,420 (all); N=3,776 (boys); N=3,644 (girls). *(**)(***) Significantly different from the estimate for children whose schools did not change start time at the (.05) (.01) (.001) level, two-tailed test.

Table 3.5: Regression-adjusted placebo testing of the impact of changing school start time on child outcomes

	1-29 minutes earlier	30-59 minutes earlier	60+ minutes earlier	Any later	
Math score					
All	-0.55	-1.03	1.71	-0.81	
	(0.92)	(2.83)	(2.13)	(1.12)	
Boys	-0.53	-1.86	1.45	-0.53	
	(0.98)	(3.26)	(1.86)	(1.09)	
Girls	-0.95	-0.56	1.82	-0.83	
	(1.34)	(2.70)	(3.27)	(1.59)	
Reading score					
All	1.57	-2.49	-1.11	-3.05	
	(1.23)	(2.94)	(2.17)	(1.56)	
Boys	1.21	-3.10	-3.38	-2.52	
	(1.61)	(3.46)	(3.32)	(1.61)	
Girls	1.94	-2.12	1.37	-3.47	
	(1.37)	(2.72)	(1.14)	(1.97)	
Internalizing problem behavior					
All	0.01	-0.01	-0.04	-0.05	
	(0.04)	(0.75)	(0.10)	(0.04)	
Boys	0.01	-0.05	-0.08	0.00	
	(0.04)	(0.08)	(0.21)	0.04	
Girls	0.00	0.03	0.00	-0.11	*
	(0.07)	(0.14)	(0.11)	(0.05)	
Externalizing problem behavior					
All	-0.02	0.05	0.03	0.05	
	(0.05)	(0.05)	(0.05)	(0.04)	
Boys	-0.02	0.04	0.09	0.04	
	(0.06)	(0.06)	(0.05)	(0.04)	
Girls	0.00	0.07	0.00	0.06	
	(0.06)	(0.06)	(0.10)	(0.06)	
Approaches to learning					
All	0.02	-0.07	-0.03	-0.06	
	(0.05)	(0.06)	(0.09)	(0.06)	
Boys	0.05	-0.07	-0.05	-0.08	
	(0.07)	(0.06)	(0.15)	(0.06)	
Girls	-0.01	-0.07	-0.02	-0.02	
	(0.05)	(0.07)	(0.06)	(0.09)	

Self-control					
All	-0.06	0.04		0.01	-0.01
	(0.05)	(0.05)		(0.11)	(0.05)
Boys	0.00	-0.03		-0.17	0.02
	(0.05)	(0.06)		(0.13)	(0.05)
Girls	-0.13	0.13		0.17	-0.03
	(0.07)	(0.09)		(0.14)	(0.05)
Interpersonal skills					
All	-0.11	-0.11		-0.05	-0.05
	(0.06)	(0.06)		(0.09)	(0.04)
Boys	-0.09	-0.16	*	-0.01	-0.05
	(0.07)	(0.07)		(0.16)	(0.06)
Girls	-0.13	-0.06		-0.07	-0.06
	(0.07)	(0.09)		(0.05)	(0.05)

N=7,420 (all); N=3,776 (boys); N=3,644 (girls). *(**)(***) Significantly different from the estimate for children whose schools did not change start time at the (.05) (.01) (.001) level, two-tailed test.

Chapter 4

Labor Market and Health Insurance Impacts Due to "Aging Out" of the Dependent Coverage Provision of the Affordable Care Act

4.1 Introduction

In September of 2010 one of the first provisions of the Patient Protection and Affordable Care Act (ACA) went into effect, allowing young adults up to age 26 to remain on a parent's health insurance plan as a dependent, provided that they did not have an offer for health coverage through their own employer. The goal was straightforward— to expand health insurance to a group of individuals that had historically high rates of uninsurance. Since the predominant source of health insurance in the United States for working-age adults is through an employer (Kaiser Family Foundation 2012), this provision relaxed the tie between employment and insurance for young adults, allowing more flexibility in job choice and a potential reduction in job lock, or inability to leave a job for fear of losing health insurance benefits (Madrian 1994). For qualifying individuals seeking health insurance, the provision altered insurance choice set, leading to changes in labor market and health insurance outcomes. As eligibility for this program expires on an individual's 26th birthday (Healthcare.gov), these changes are most likely to affect an individual whose age is close to age 26. This paper estimates the impact of turning 26, or "aging out," on labor and health insurance market outcomes for young adults in the United States.

The provision has expanded health coverage to millions of young adults, covering more than 3 million individuals from implementation through December of 2011 (ASPE Issue Brief, 2012). By the end of 2012 an estimated 8 million young adults between the ages of 19 and 25 were able to remain on their parents' plans (The Commonwealth Fund Health Insurance Tracking Survey of Young Adults, 2013). In terms of changes in labor market outcomes, studies have found no evidence of the provision changing the likelihood of a young adult being employed, but small reductions in the probability of working full-time and the number of hours worked per week (Antwi, Moriya, and Simon 2013).

Rather than comparing changes in coverage and employment for the entire targeted group (young adults aged 19-25) to changes in coverage for older adults (adults aged 26-30), this paper focuses on what happens to individuals at or above the eligibility threshold, or those whom have aged out of the provision. The threshold that occurs at age 26 as a result of this provision leads to variation in labor and health insurance outcomes of the marginally ineligible young adult, which I estimate using a regression discontinuity (RD) design. This aim of this study is to determine if aging out of the dependent coverage provision leads to changes in labor market and health insurance outcomes for young adults.

A key finding is that young adults desire to remain insured, as measured by no statistical change in the uninsurance rate but a significant increase in directly purchased non-group health insurance plans. Males increase their labor force participation at the threshold, with this effect increasing in magnitude when the sample is limited to unmarried individuals. Additionally, young adults that have aged out of the dependent coverage provision are much more likely to report that their health insurance plan is worse than it was one year ago. Results from this study not only provide insight into possible behavior of young adults when faced with the individual health insurance mandate of the ACA but also pertinent information for health insurance marketplace outreach coordinators.

4.2 Background

4.2.1 Health insurance and labor supply

Many studies have analyzed how health insurance affects labor supply. Recall that job lock is defined as remaining in a job for fear of losing health insurance benefits (Madrian 1994). Thus, numerous researchers have analyzed labor market behavior among individuals after their health insurance choice set extends beyond employment. In theory, access to health insurance outside of an individual's own-employment may affect retirement timing decisions, employment choices of second earners, and labor force choices (Gruber and Madrian 2002). As a result, the literature focuses on three groups: 1. Older adults close to age 65, as they are able to access publicly-provided health insurance, or Medicare; 2. Married females, who may have health coverage through a spouse; and 3. Low-income unmarried mothers, as they may have access to Medicaid coverage (Gruber and Madrian 2002).

Studies focused on retirees find eligibility for outside health insurance choices after retirement, either through continuation of an employer plan or through Medicare increases the probability of earlier retirement (Karoly and Ragowski 1994; Gruber and Madrian 1995; Rust and Phelan 1997; Blau and Gilleskie 2006). Having access to spousal insurance reduces full-time work and labor force participation of married women (Olsen 2000; Honig and Dushi 2005). There is less consensus surrounding the impact of access to Medicaid on low-income single mothers, with most studies finding no changes in labor force participation in response to increases in Medicaid eligibility levels (Meyer and Rosenbam 2000; Ham and Shore-Sheppard 2003; Tomohara and Lee 2006). However, Yelowitz (1995) observes small increases in labor force participation rates of low-income single mothers following Medicaid expansions, and Moffitt and Wolfe (1992) demonstrates that an increase in the value of Medicaid leads to reductions in labor force participation by single mothers.

To place the dependent coverage provision's potential impacts on labor market outcomes in context with existing literature of the effect of labor supply decisions made in connection with health insurance status, it must be noted that the provision also extended the health insurance choice set beyond employment. Prior to the provision, a young adult desiring to have health coverage had three primary options in his choice set: first, through an employer offering health insurance; second, through Medicaid or Medicaid-like public program; and third, through directly-purchased non-group coverage. The provision gave individuals a fourth option—through a parent's employer-sponsored health insurance plan.

Missing from the dependent coverage provision and unlike ESI and non-group coverage, key requirements to receiving health insurance coverage, employment and (own) income, are not present. However, the dependent coverage provision is somewhat similar to Medicaid coverage in that in order to receive health insurance, (own) employment is not required, but dissimilar in its lack of an income limit. The provision has the potential to reduce income barriers to health insurance coverage and increase the value of parental health insurance. Therefore I would expect results from this study to be similar in nature to Moffitt and Wolfe's 1992 labor market findings among low-income single mothers—just as the expansion of Medicaid eligibility increased the value of Medicaid for them and reduced labor force participation, I expect the decreased value of parental ESI that results from the aging out of the dependent coverage provision to increase labor force participation and subsequent employment.

4.2.2 Dependent coverage provision

Prior work has relied on the "natural experiment" that resulted following the passage of the dependent coverage provision by using a difference-in-differences framework to analyze changes in outcomes for eligible young adults. These models estimate the change in an outcome (e.g., insurance rate) for the targeted group (individuals aged 19-25 or 19-26, depending on the study) and subtract from that the change in the same outcome for a control group (e.g., individuals aged 27-30). This methodology accounts for underlying trends not related to the provision, such as changes in the economy, and the difference-in-differences estimator is interpreted as the policy effect (Wooldridge 2002) of the dependent coverage provision.

Using different data, studies show gains in coverage of roughly 2 to 6 percentage points. Cantor et al. (2012) and Sommers and Kronick (2012) show this effect using data from the Current Population Survey; Sommers et al. (2013) use the National Health Interview Survey (NHIS) and the Annual Social and Economic Supplement of the Census Bureau's CPS; Antwi et al. (2013) utilize panel data from the Survey of Income and Program Participation (SIPP; O'Hara and Brault (2013) use the American Community Survey (2013); and Mulcahy et al. (2013) utilize hospital claims data. In general, coverage gains are larger for males, non-students, and unmarried individuals (Sommers et al. 2013).

The provision clearly improved health insurance coverage for young adults, but it also has had some impacts on labor markets. Although many affected individuals moved from uninsured to insured, by removing the link between employment and insurance, the law has the potential to reduce job lock. Antwi et al. (2013) find no significant evidence of the mandate changing the likelihood of a young adult being employed, but do see a 2 percentage point reduction in the probability of working full-time. Additionally, they find small reductions in the number of hours worked per week.

One concern with these studies is that the methodological approach, difference-in-differences, requires satisfaction of an assumption that the treatment and control groups would have followed parallel trends in the absence of the policy (Slusky 2015). Without this condition being satisfied, differences in trends between the treatment and control groups might not be a true reflection of the effect of the policy. After plotting the age and outcome trends for the control and treatment groups and noting age bands for which they are not parallel, Slusky replicates the findings of several of these papers (Sommers and Kronick 2012; Cantor et al. 2012; Antwi et al. 2013). He then estimates placebo models using the same treatment and control groups but in years prior to the provision being in place and finds significant effects for the treatment groups in these earlier years, suggesting that age and labor force characteristics might be explaining some of the "gains" from the dependent coverage provision. One method offered by Slusky to increase the precision of estimates surrounding a sharp age cut-off is to reduce the age bandwidth, motivating my study's use of individuals aged just below and just and above the eligibility threshold.

4.3 Conceptual Framework

Aging out of the dependent coverage provision removes a health insurance option for individuals, leading to a change of all available consumption-leisure combinations in the opportunity set. Using the neoclassical model of labor-leisure choice, I will show that removal of access to health insurance through a parent can be viewed as a downward shift (reduction) of the young adult's budget constraint. At a lower budget constraint, new optimum levels of labor market outcomes will result.

Young adults draw utility from consumption, (C), leisure (L), and health insurance (HI), with β capturing individual preferences of the value of health insurance:

$$U(C, L, HI) = f(C, L) + \beta HI \tag{4.1}$$

Individuals are constrained by time and income, including nonlabor income (V), and make choices regarding how many hours (h) to work each pay period and whether or not to purchase health insurance. These choices are impacted by the individual's wage rate (w), which is assumed to be equal to the individual's level of productivity and the same for both part-time and full-time work ($w_{pt} = w_{ft}$). The cost of health insurance (HI_{cost}) also impacts the possible consumption-leisure bundles available to young adults. Thus a young adult is faced with the following budget constraint:

$$C + wL = hw + V - HI_{cost} \tag{4.2}$$

To capture changes in outcomes for young adults whose behavior would change in response to aging out of the dependent coverage provision, I restrict the conceptual model to those whose value of health insurance, β, is high enough that they would seek health insurance coverage once parental insurance is no longer in the choice set. This restriction is consistent with RD estimation in that individuals on either side of the cutoff should be similar in order to address the concern of self-selection into coverage under a parent (further explained in section 4.5). In other words, I assume individuals with low valuations of β would not have changed their behavior in response to losing eligibility for the provision as they would not have been insured while eligible. Thus, for each individual the following utility maximization problem results:

$$max_{C,L}\, f(C,L) \text{ subject to } C + wL = hw + V - HI_{cost} \tag{4.3}$$

One key result from solving this maximization problem is that the marginal rate of substitution will equal the wage rate $(\partial U_L/\partial UC = w)$. ∂U_L is the change in utility from an additional hour of leisure and ∂U_C is the change in utility from spending an additional dollar on consumption goods. Since wages are positive, it follows that $\partial U_L/\partial UC > 0$. Incorporating this result into the budget constraint leads to a positive relationship between hours worked and the cost of insurance ($h = L + [C - V + HI_{cost}] / (\partial U_L/\partial UC)$ Ceteris paribus, an increase in the cost of health care will reduce the demand for leisure and increase hours worked.

I assume that losing eligibility for parental ESI increases the cost of health insurance for young adults for several reasons. The first is that the marginal cost of adding a dependent is less than the cost of purchasing a separate plan (Antwi et al. 2013; United States Department of Labor 2015). In other words, if the young adult were to obtain ESI through his own employer, his monthly premium would likely exceed that of the additional cost to add him to a parent's plan. Thus even if the parent were having the child reimburse for the full amount of their child's portion of the health insurance premium, the cost would still be less than if the young adult were paying for single-coverage health insurance through his own employer. Trends in other markets that group the entire family under one bill suggest parents do not necessarily have their adult children reimburse them for their portion of the family bill. For example, in the cellular phone market, 29% of parents paid the bill for their adult (aged 18-35) child's portion even when they did not reside in the same house (Harris Interactive Poll, 2013). Lastly, if the young adult transitioned to directly purchased non-group health insurance that directly purchased non-group coverage, due to the favorable tax treatment of ESI that would be relatively more costly (U.S. Department of Labor, 2015).

Graphical representations of expected labor supply impacts of aging out of the dependent coverage provision are shown in Figures 4.1 and 4.2. These models denote hours of leisure on the x-axis and (non-labor) consumption in dollars on the y-axis. In general, an increase in the cost of health insurance will reduce an individual's budget constraint. In addition to this downward shift, there is also a kink in the budget constraint at the point where hours worked equals full-time employment. This kink occurs for two reasons: 1. ESI is more commonly offered to full-time workers versus part-time or temporary workers (in 2013, 47% of firms with more than 200 employees offered health insurance compared to only 25% of small firms, and 3% of temporary workers received an offer of health insurance coverage) (Kaiser Family Foundation, 2013 Employer Health Benefits Survey); 2. ESI has favorable tax treatment when compared to the directly purchased private insurance available for part-time and temporary workers (i.e., has more of a negative impact on budget).

The solid black line is the budget constraint for young adults eligible for parental insurance. The dotted black line is the budget constraint after losing eligibility, showing both the overall reduction as well as the kink occurring at full-time employment. Leisure-consumption preferences are denoted by utility curves U_0, U_1, and U_2. Tangency of the utility curves with the budget constraints represent various optimal leisure-consumption combinations (i.e., where utility is maximized).

Based on the assumptions made, this model predicts ineligibility for parental ESI will increase labor force participation and employment. Figure 4.1 models the potential behavior of an individual working part-time before turning 26. He is maximizing utility at point A (C_0, L_0). When ESI through a parent is no longer an option, to remain insured he could remain a part-time worker (L_0) and purchase non-group private insurance. This would give him C_1 units of consumption and leave him at point B. Alternatively he could reduce his leisure hours to L_1 and work full-time in order to receive ESI through an employer (point C). Since utility associated with point C is higher than the utility at point B, this suggests full-time employment will increase as a result of aging out of the dependent coverage provision. (Note that public coverage is also a health insurance choice, allowing the young adult to potentially remain at point B, but because this time period was prior to the Medicaid eligibility expansion to low-income childless adults I believe this effect will be weak.)

Figure 4.2 represents an individual with stronger preferences toward leisure than the individual in Figure 4.1. Due to non-labor income, even though he spends zero hours working (L_0), he still consumes C_0 units of consumption goods (point A). After aging out of the dependent coverage provision, he may continue to enjoy L_0 units of leisure, but this decreases the amount of consumption goods to C_1 (point B). If he desires to remain at C_0, he will increase his labor hours (and decrease leisure hours) to L_1 (point C). Thus, I expect to see an increase in both employment and labor force participation after aging out of the young adult provision.

Since I assume that only individuals with a high enough valuation of health insurance β will be impacted by dependent coverage ineligibility, I do not expect to see changes in the share of uninsured individuals. Figures 4.1 and 4.2 both suggest possible substitution of the source of private insurance, with increases in directly purchased non-group coverage or increases in (own) ESI. As these data do not decipher whether ESI belonged to the respondent or the parent, the reductions in health insurance coverage through a parent's employer might equal the gains in (own) ESI by the individual, leading to no change in this outcome at age 26. In addition, with the average wait time of 2 months between being hired and being offered health insurance (Kaiser Family Foundation, 2014 Employer Health Benefits Survey) I do not expect to see changes in health insurance offer rates at age 26. Lastly, as health insurance plan satisfaction is related to health insurance plan type and I am treating this change in health insurance coverage as an increase in health insurance costs, I expect plan satisfaction rates to decrease. That is, after aging out of the provision, these individuals are experiencing a reduction in their budget constraints, so I expect that by shifting to their own ESI plan, public insurance, or directly purchased non-group coverage, this will lead to increased rates of individuals reporting that their plans were worse than they were one year prior.

4.4 Data

4.4.1 Sample

The National Health Interview Survey (NHIS) provides detailed information on health, health insurance, and employment for a representative sample of the overall civilian, non-institutionalized population of the United States. Data were drawn from a harmonized version of the NHIS, the Integrated Health Interview Series (IHIS), provided by the Minnesota Population Center. The sample is restricted to the years 2011-2013, a time period after implementation of the provision but before the ACA individual mandate and expansion of the dependent coverage provision to individuals offered health insurance through their own employer as well. Within the NHIS, labor and health insurance outcomes are asked of all individuals, with the exception of whether health coverage type was better/worse/the same as the previous year, which is limited to a randomly selected sample adult within the household.

Since the provision allowed married individuals to remain on a parent's insurance until age 26, the sample included both married and unmarried young adults. However, it should be noted that married men have higher labor market participation rates than their unmarried counterparts, whereas married women are less likely to be in the labor market than unmarried women (U.S. Census Bureau 2014). Additionally, marriage may impact health insurance coverage status, as literature has found increases in health coverage and employer-sponsored insurance offer rates for women (Bernstein et al. 2007). Models account for these differences in outcomes based on marital status, but because prior studies have found the dependent coverage provision had stronger effects for unmarried individuals (O'Hara and Brault 2013; Sommers et al. 2013), models limited to single young adults are estimated as part of the sensitivity analyses of the overall study. The provision also had differential effects based on gender (O'Hara and Brault 2013; Sommers et al. 2013), so estimates are produced for the full sample as well as by gender.

4.4.2 Key measures

An attractive feature in the NHIS is that it contains month and year of interview and month and year of birth. This is used to create a more precise definition of age at the time of the interview as well as how much older or younger the respondent was, relative to the eligibility cut-off. Following studies using a similar methodology (Carpenter and Dobkin 2009; Yoruk and York 2011), the selected age bandwidth includes respondents up to 2 years younger or older than the eligibility threshold occurring on the individual's 26th birthday. Because the NHIS age and interview information is in months, not days, simply subtracting interview date from birth date results in a window of plus or minus one month in which an individual may have been miscoded. In other words, a person stating they are age 26 in the survey may have had a calculated age of 25.9, or a person stating that they are age 25 in the survey may have had a calculated age of 26.1. Therefore, I recoded these individuals very close to age 26 but with the incorrect sign into the nearest age bin; thus a person I calculated as being 25.9 but because his/her age on the survey was 26 was then recoded into the first age bin after age 26. Models were also estimated eliminating individuals whose age was within one month of 26 (Appendix E).

The series of questions regarding employment had a 2-week reference period, whereas the health insurance questions referred to the status at the time of survey. Since the empirical strategy compares young adults who are slightly younger than the young adult provision age cut-off to those who are slightly older, these short reference periods for outcome measures are ideal. The final analytic sample includes 13,235 individuals. Analyses are weighted using the survey estimation procedure (svy) in Stata 12 (StataCorp 2012).

Outcomes focus on changes in employment, employment-related health coverage, coverage type, and plan satisfaction. The three labor market measures referenced the prior two weeks and include employed: if the individual worked for pay; in the labor force: if the individual was working for pay or looking for work; and employed full-time: the individual worked 32 or more hours per week. The set of employer-related health coverage measures included if the individual had employer-sponsored insurance (ESI), and of those with a job, whether or not the employer offered health insurance (ESI offer). Health coverage type referred to coverage at the time of the survey and is captured by one of three measures, including public insurance: the individual had coverage through Medicaid, Medicare, or another public assistance or state sponsored plan; private coverage: insurance was provided in part or in full by an employer or union or purchased directly; uninsured: no health insurance coverage. These data also allowed for separation of private insurance not provided by an employer, or directly purchased non-group coverage. Lastly, an indicator of health plan quality (added to the NHIS in 2011 for a subpopulation of randomly selected sample adults only) is analyzed. A summary of these measures with full definitions is found in Table 4.1.

4.5 Empirical Model

A regression discontinuity (RD) design is used to estimate how aging out of the dependent coverage provision impacts labor and health insurance outcomes among young adults. The advantage of using RD design is that it adjusts for potential selection. In other words, young adults who value health insurance might be more likely to sign up for coverage while eligible for the dependent coverage provision as well as to remain insured after aging out, leading to endogeneity. The intuition is that the characteristics of an individual aged 25.9 will be similar to those of an individual aged 26.1, but the older individual will not receive coverage through a parent and thus changes in outcomes between the two will represent the policy effect.

This methodology takes advantage of the sudden change in health coverage options that might result after an individual turns 26 and becomes ineligible for health insurance through a parent. To ensure correct profiling of age, I first followed the methodology of previous RD literature, visual data inspection (Carpenter and Dobkin 2009; Yoruk and York 2011) to determine the highest-order of terms for age. This led to age being expressed as quadratic polynomials. Next, I followed RD model fit guidelines as outlined in Trochim (2006) and reestimated the models using age-cubed and age-cubed interacted with treatment to see if it was appropriate to not include these terms. An example of this test is found in Appendix D. In the model including age-cubed and age-cubed interacted with treatment, both terms are insignificant, suggesting they may not be needed. After dropping these terms from the model and reestimating, there was a gain in efficiency as noted by the decrease in magnitude of the standard error of the treatment effect. The model with the age profile fully interacted with the treatment is:

$$Y_i = B'X_i + \theta_1 Treat_i + \theta_2(age_i) + \theta_3(age_i{}^2) + \theta_4(Treat_i * age_i) + \theta_5(Treat_i * age_i{}^2) + \varphi_t + \gamma\varepsilon_{it} \qquad (4.4)$$

Y_i is a labor market or health insurance outcome for individual i. Vector X_i contains observable characteristics for individual i, including dummy variables to control for marital status, highest educational attainment, region of residence, health status (measured by self-reported health status as being poor, fair, good, very good, or excellent), presence of a chronic health condition, citizenship, gender (for the models including both males and females), race/ethnicity, and poverty status. This indicator of poverty status reflects the U.S. Department of Health and Human Services poverty guidelines and was created by the Integrated Health Interview Survey staff using family size and imputed family income. The study opted to use income at or below 138% of the federal poverty guidelines (FPG) as the public data file does not contain continuous family income and it was important to control for the possibility that poverty influences both labor market and health insurance outcomes. Thus, family income at or below 138% FPG was the best available proxy for poverty. The selection of 138% FPG by IHIS staff reflects the guidelines commonly used in administrative purposes for public eligibility (in particular, the Medicaid expansion, which allows single childless adults with incomes at or below 133% of FPG, plus a 5% income disregard, to qualify for Medicaid if they reside in a state that has agreed to expand the program).

Since the study design is over three years, year fixed effects (φ_t) are included. The treatment measure is captured by $Treat_i$, which is zero for all individuals younger than 26 and one for those 26 and older. For all models, age_i is the time in months before or after the individual's 26th birthday. Logit models estimate θ_1 and control for the complex design of the sample survey using the survey estimation procedure (svy) in Stata 12. The results (Table 4.3) show the average marginal effect, or percentage point increase (decrease) in an outcome in response to turning 26.

In regression discontinuity (RD) design, it is important to ensure that the discontinuities observed for the outcomes at the threshold are not also occurring with other covariates in the model (the right-hand side measures). Table 4.2 shows the results from testing the smoothness of the observable characteristics for unmarried young adults around the Young Adult Provision eligibility cutoff (age 26). This test demonstrates a lack of significance at age 26 for gender, highest educational status, citizenship, poverty status, residential region, and health status, suggesting that the jumps in the insurance and labor market outcomes occurring at age 26. While the share of whites, Hispanics, and other non-Hispanic minorities did not change at the threshold, the share of non-Hispanic blacks increased slightly. In RD design, some variables may differ for the two groups based on random chance (Lee, Moretti, and Butler 2004). As there were no significant differences in any other race/ethnicity measure (or any other observable covariate), I do not believe this slight increase in the share of non-Hispanic blacks is a threat to the validity of the results.

Also key to RD design is that the respondent does not have any control over the measure that has the known cutoff, or the forcing variable. Since age is the forcing variable, this condition is satisfied. Additionally, nonrandom sorting of young adults to one side of the threshold should not occur. The NHIS survey was conducted monthly to individuals of all ages, so while this occurrence would have been unlikely, Figure 4.3 shows the distribution of young adults around the eligibility threshold. It suggests nonrandom sorting around age 26 did not occur.

4.6 Results

Summary statistics suggest differences in outcomes at the eligibility threshold (Table 4.1). Overall, being 26 or older is associated with increased employment, full-time employment, offers of employer-sponsored health insurance, uninsurance, directly purchased health insurance, and perceptions of health insurance plan being worse than one year prior. Aging out of the provision is also associated with reductions in private insurance, driven by reductions in employer-sponsored private insurance.

Results report the coefficient on the treatment (θ_1) from estimating Equation (4.1) for each outcome. Regressions include the quadratic polynomial of age fully interacted with a dichotomous indicator of treatment, individual-level characteristics, and year fixed effects, with standard errors clustered at the individual level. Figures 4.4, 4.5, and 4.6 show the pre/post outcomes for each of the ten outcomes, plotting the quadratic fitted lines from the estimated parametric models (without controls) over the mean values of the share of individuals in each age bins.

Aging out of the provision has differential impacts by gender, but for the full sample there are changes in health insurance coverage and coverage-related outcomes. At the threshold, there is no change in the share of individuals reporting being uninsured, but there is a 4.4 percentage point increase in the purchase of nongroup health insurance coverage.

This uptick in directly-purchased private health insurance suggests an interest in remaining insured. Panel (I) of Figure 4.4 also shows this discrete change, and demonstrates that after the increase at age 26, the probability of purchasing non-group coverage initially rises and then begins to decline.

There is a 15.1 percentage point increase in the probability of reporting that coverage was worse than one year prior. This finding is supported by the fact that directly purchased (i.e., non-ESI) health insurance typically provides less generous benefits than group (i.e., ESI) coverage (Reschovsky and Hadley 2007) and this study shows an increase in directly purchased coverage once individuals aged out of the dependent coverage provision. The jump in dissatisfaction with health insurance plan is shown graphically in Panel (J) of Figure 4.4. After the increase at age 26, the probability of reporting coverage is worse than it was one year prior declines.

When models are estimated for men and women separately, the main theme that emerges is that both young men and women find health insurance coverage quality is worse at age 26 than it was the prior year. However, by relaxing the employment-insurance connection, results suggest there are subsequent changes in employment choices for young men and indicate that any reduction in job lock during the eligibility phase erodes on the 26th birthday.

For young men, turning 26 led to 7.5 percentage point increase in labor force participation (Table 4.3). Panels (A) and (B) of Figure 4.5 show the unadjusted pre and post age 26 jumps in these labor market outcomes and suggest that even 6 months, one year, 18 months, and 2 years past the eligibility threshold, employment and labor force participation rates are higher than they were during the eligibility period. These results imply that unmarried young men might have been using the ability to stay on a parent's insurance as a reason to (temporarily) not participate in the labor market, or a loosening of job lock during the eligibility phase. The delayed entry hypothesis is supported by a recent study by the Urban Institute (Nichols and Linder 2013) that found declining labor force participation rates observed during the Great Recession were driven by declining labor force entry rates, not increased labor force exit rates. Regardless of whether or not young men had delayed entry or temporarily exited the labor force, my finding of increased labor force participation was coupled with a 6.2 percentage point increase in the purchase of non-group health insurance coverage by young men. As shown in Figure 4.5 (Panel I), this trend continues after age 26. Young men are also likely to report large increases in health insurance plan quality dissatisfaction (+12.2 percentage points).

Aging out of the provision did not lead to any labor market changes for young women. However, when asked to compare current health coverage to that of one year prior, turning 26 is associated with a nearly 18 percentage point increase in the share of women responding that their coverage was worse in that year versus the prior year (Table 4.3). Again, this jump is supported by the graphical results (Panel (J) of Figure 4.6). As there are no significant changes in this measure for men, these results are possibly due to the fact that women are higher utilizers of health care (Bertakis et al. 2000) and may have had a health care experience under both types of insurance from which to judge coverage (e.g., annual pap and pelvic exam).

4.6.1 Robustness testing: alternative samples

A concern with RD design is selecting the appropriate sample of individuals. It is worth investigating if the main results from the paper dissipate when the sample narrows or widens. Estimates using a narrower band of individuals are smaller in sample size and thus less biased (Table 4.4, Panel (A)) and those using a wider band are including individuals further removed from the threshold and have thus had more time to adjust to losing eligibility (Table 4.4, Panel (B)). Despite less precise estimates, the finding of health insurance coverage being worse (overall and for females) and increase in non-group coverage (males) remains (Panel (A)). Inclusion of individuals further from the threshold leads to results having the same general pattern is evident as in the main findings—increases in shares reporting worse coverage (overall and across genders), increase in labor force participation rates (men), and an increase in non-group private coverage (overall).

Models were estimated for unmarried individuals as well (Table 4.4, Panel (C)). As mentioned, other studies have found differential effects of the dependent coverage provision based on marital status (O'Hara and Brault 2013; Sommers et al. 2013). While these models show increases in employment (for men) and offers of ESI (for women), they also continue to show a jump in the labor force participation rate of men and reports of health insurance plan being worse (overall and for women), reinforcing the main findings.

4.6.2 Robustness testing: model fit

In addition to testing the effect of alternative samples on the results, Table 4.5 presents the results from testing RD model fit under three alternative scenarios. Panel (A) shows the results from estimating Equation (4.1) using a sample of individuals aged 24 up to 28 but in 2004-2006, a period of time several years prior to implementation the dependent coverage provision and in a similar downward trending unemployment (prior to the Great Recession). There are no significant jumps at the threshold (age 26), supporting the notion that the discontinuities found (Table 4.5) are being driven by aging out of the young adult provision and not simply from turning 26.

The remaining two panels in Table 4.5 present results from models that use the same years of data (2011-2013) and data source (NHIS) as the primary models, but use artificial eligibility thresholds and samples involving only young adults to the right or left of age 26. The lack of significant results from these model specification tests (Table 4.5, Panels (B) and (C)) reinforces confidence in both appropriateness and accuracy of the RD model used to address this research topic.

Models excluding individuals who were one month older or younger than 26 were also estimated, with the results paralleling the main findings—males increased labor force participation, females reported health insurance plans were worse, and there was an increase in directly purchased private coverage (Appendix E). Different from the primary findings were that when individuals very close to the cut-off are excluded from the sample, there is an increase in employment (overall) and decrease in ESI (overall).

4.6.3 Limitations

The study design and data are not without limitations. The NHIS public use data file does not contain state identifiers, so the study could not control for the fact that more than half of the states had already extended the age that young adults can remain on a parent's health insurance plan when the ACA provision went into effect (National Council of State Legislatures 2014). However, Section 514 of the Employee Retirement Income Security Act of 1974 (ERISA) preempts state laws for self-insured plans (29 U.S.C. § 1144, Section 514), and during the study period nearly 60 percent of private-sector employees with ESI had self-insured plans. (Employee Benefit Research Institute 2012) Additionally, most of these state-sponsored plans had stringent eligibility requirements in order to qualify for state coverage (e.g., unmarried, financially dependent on the parent, living in the same states as the parent, full-time student, under age 25). For these reasons, it is plausible that the results found using national data are being driven by the federal law. Also, region of residence is used to control for geographic area, and other research that did include state identifiers found no significant differences in outcomes between young adults residing in states with prior dependent coverage laws and those without (O'Hara and Brault 2012; Antwi et al. 2012).

Related, by not having a state identifier I am not able to control for whether or not the young adult resided in a state that was an "early Medicaid expansion" state. These are states that expanded Medicaid eligibility to low-income childless adults (income at or below 138 percent of the federal poverty guidelines) prior to the optional provision beginning in 2014 (Centers for Medicare and Medicaid, 2015). However, I am able to control at the individual level if the respondent had income at or below this threshold (roughly 20% in my sample) and also during the first two years of my study only six states had expanded their Medicaid programs to cover these individuals (Kaiser State Health Facts, 2015). Future work using restricted NHIS would analyze changes based on expansion and non-expansion states, and I expect in the expansion states there might have been changes in public coverage rates at the dependent coverage threshold.

Many colleges and universities mandate the purchase of health insurance (American College Health Association 2014) and thus these individuals are more likely to be insured than non-students. For example, in the 2009-2010 school year the overall rate of uninsurance for graduate and undergraduate students was 7.4% (American College Health Association 2010) compared to rates of roughly 30% uninsured among the total population targeted by the dependent coverage provision. (Centers for Medicare and Medicaid Services 2014) However, the NHIS does not include an indicator for student status, and even though the ACA dependent coverage provision extended coverage to all students up to age 26, many of these individuals were already insured prior to implementation. Highest educational attainment is used to control for differences in education among young adults.

As mentioned, these data do not allow me to discern the policyholder of employer-sponsored health insurance. Had this been possible, I may have been able to tease out the percentage gains/losses in ESI as a result of losing parental coverage/gaining own ESI coverage. Lastly, loss of eligibility for parental insurance may have been mitigated through the continuation of health insurance coverage via the Consolidated Omnibus Budget Reconciliation Act (COBRA). Young adults turning 26 are able to remain on a parent's employer-provided plan for up to 36 months. However, the cost of coverage through COBRA is equal to the total cost of the health premium for a single plan plus a 2 percent administrative fee, whereas most employers pay between 50 and 90 percent of the cost of health insurance, making COBRA a relatively expensive alternative. For example, in 2013 the average monthly premium for an individual health plan was $490, compared to $272 for an individual plan purchased on the health insurance marketplace (Kaiser Employer Health Benefits, 2013). The increase in directly-purchased private insurance (+4.4 percentage points overall and +6.2 percentage points for males) suggests that it may have been more affordable for individuals (and families) to have the young adult purchase his or her own plan on the private market instead of through COBRA.

4.7 Discussion

Using the NHIS, which contains birth date and interview date by month, loss of eligibility for parental insurance is precisely identified and used to determine the immediate effects of aging out of the young adult provision of the ACA. While existing literature demonstrates that the provision has many positive effects for the target population while eligible, to the best of my knowledge, this is the first analysis of how loss of eligibility alters individuals' labor market and health coverage choices.

This paper finds that ineligibility for the young adult provision did not lead to increases in uninsurance rates, suggesting an interest in remaining insured, even in an era prior to the ACA's individual health insurance mandate being in place. From a policy perspective, the differences in outcomes based on gender are particularly important. Results show males appear to have been either more willing or more able (or both) to delay entry into the labor force or temporarily exit the labor force while eligible for the provision. Trends in living arrangements during the study period support the hypothesis that men may have had more financial support to allow a delay or retraction from the labor force—in 2012, more than one in three young adults aged 18-31 resided at home with their parent(s), with men being more likely than women to do so. Another explanation for the jump in employment and labor force participation rates among young men is that young women's skills may have been better matched with their employment choices when the provision went into effect, resulting in fewer of them leaving the labor force when becoming eligible. Job skill matching may have been a carry-over from the Great Recession (December 2007-June 2009), which was harder on men in terms of job loss during the economic contraction and job growth in the subsequent recovery. (Pew Research Center, 2015).

The exogenous change that reduced the possibility of job lock for eligible young adults may have had positive or negative welfare implications worth exploring in future research. Assuming that labor force exit was voluntary (e.g., not the result of discourage workers that have given up looking for employment), then young men could have used eligibility to temporarily exit the labor force and improve their labor market match. Upon returning to the labor force after this temporary leave, such men could then be in a career better suited to their job skill set, representing a welfare increasing outcome of the dependent coverage provision.

However, the provision may have been welfare decreasing if the individual simply exited the work force because they did not need employment in order to be insured, and after their temporary exit returned to a job that required the same (or worse, fewer). In this case, the provision could be viewed as disrupting human capital formation, with longer-term effects potentially being observed over time (e.g., reduced lifetime earnings). While these data do not allow one to decipher if individuals were using the provision to return to school, the period of time was associated with increased growth in graduate school applications but declining enrollment rates (-2.3% in the 2011-2012 academic year and -0.2% in the 2012-2013 academic year) (The Council of Graduate Schools), suggesting individuals may not have been substituting education for employment. Additionally, if young adults were returning to school, the increase in labor force participation at age 26 would suggest that many graduated on or around their 26th birthday, which seems implausible unless most in the sample were interviewed and born in the months of May and June.

Although no statistically significant jumps in broad coverage type occurred at the threshold (public, private, or unnsured), changes within plan type (e.g., from a parent's private insurance to their own directly purchased non-group private insurance) may have contributed to insurance coverage quality being perceived as worse than one year prior. For some, the increase in plan dissatisfaction might be a reflection of the first time that the young adult navigated the health care system on their own, and might not necessarily be a true indicator of plan quality.

This study focuses on a period in time prior to the individual health insurance mandate of the ACA going in effect. Although there were no changes in uninsurance rate at the threshold, coefficients were negative. Therefore, my results indicate that moving forward there might be an increased interest among those aging out of the dependent coverage provision in remaining insured at age 26. Many young adults will turn to state and federal health insurance marketplaces for information about health coverage. As more than half of young adults (aged 18-29) regularly use two or more social media sites (Pew Research Center 2015), marketplace education and outreach coordinators could use these sites to advertise to individuals getting ready to celebrate a 26th birthday. This finding is especially important for young men, as this study demonstrates they are more rapidly reentering the labor market and not necessarily selecting employment based on the potential offer of health insurance.

Table 4.1: Definitions of outcome variables and summary statistics for young adults

Variable: Definition	All Age 24-25	All Age 26-28		Male Age 24-25	Male Age 26-28		Female Age 24-25	Female Age 26-28	
Employed: working for pay in the last two weeks.	73.1 (0.61)	76.0 (0.63)	***	76.0 (0.89)	76.3 (0.80)		76.3 (0.83)	81.7 (0.94)	***
In the labor force: working for pay or looking for work in the last 2 weeks.	83.3 (0.60)	84.4 (0.74)		84.4 (0.67)	87.2 (0.57)	***	87.2 (0.73)	91.1 (0.84)	***
Employed full-time: working 32 or more hours for pay in the last two weeks.	75.1 (0.70)	79.4 (0.64)	***	79.4 (0.95)	78.5 (0.79)		78.5 (0.95)	84.4 (0.79)	***
Employer Sponsored Insurance (ESI): of the privately insured, share covered through employer.	87.9 (0.60)	84.0 (0.74)	***	84.0 (0.80)	87.6 (0.96)	***	87.6 (0.85)	83.8 (1.10)	***
ESI Offer: of those employed, share working for an employer offering health coverage	60.7 (0.77)	65.8 (0.77)	***	65.8 (1.06)	61.3 (1.06)	***	61.3 (1.12)	64.8 (1.02)	***
Uninsured: did not have health insurance coverage at the time of survey.	27.1 (0.67)	31.7 (0.71)	***	31.7 (0.96)	30.9 (1.01)		30.9 (0.80)	37.5 (0.84)	***
Public coverage: Medicaid, Medicare, or other public assistance/state sponsored plan.	12.8 (0.44)	12.5 (0.51)		12.5 (0.50)	7.6 (0.51)	***	7.6 (0.68)	7.3 (0.77)	
Private coverage: insurance provided in part or in full by an individual's employer or union, or purchased directly by a person.	60.2 (0.80)	55.7 (0.76)	***	55.7 (1.02)	61.5 (1.04)	***	61.5 (1.05)	55.2 (0.94)	***
Direct purchase: private health coverage purchased directly, rather than through an employer or union.	5.3 (0.42)	9.1 (0.59)	***	9.1 (0.64)	5.7 (0.79)	***	5.7 (0.51)	9.1 (0.86)	***
Worse insurance[a]: compared to one year ago, health insurance coverage is worse.	11.4 (0.66)	16.9 (0.81)	***	16.9 (0.90)	10.8 (1.17)	***	10.8 (0.94)	15.1 (1.05)	***

Notes: N= 13,235 (all); N= 7,192 (males); N=8,043 (females)
[a]N= 6,765 (all); N= 2,996 (males); N= 3,769 (females)
Significant difference between groups: *p<0.05,**p<0.01,***p<0.001. Standard errors are in parenthesis.

Table 4.2: Testing the smoothness of observable characteristics for young adults around the Young Adult Provision eligibility cutoff

Outcome	All	Males	Females
Highest educational level completed			
Less than high school	0.7	1.9	-0.3
	(1.42)	(2.28)	(2.02)
High school graduate (or equivalent)	-0.1	-3.6	3.5
	(2.28)	(3.62)	(3.09)
Some college or college degree	1.6	7.1	-3.5
	(1.64)	(3.88)	(3.50)
Graduate degree or beyond	-2.4	-5.0	0.0
	(2.96)	(3.60)	(3.98)
US citizen	-1.2	-2.2	-0.3
	(1.50)	(2.19)	(2.28)
Income </= 138 FPG	1.9	2.2	1.6
	(2.31)	(3.61)	(3.20)
Region of residence			
South	-0.4	-2.6	1.7
	(2.86)	(3.88)	(3.84)
Northeast	-0.5	-0.5	-0.5
	(2.03)	(3.64)	(2.83)
Midwest	0.0	0.0	0.0
	(2.78)	(3.83)	(3.59)
West	0.9	3.1	-1.1
	(2.50)	(3.50)	(3.51)
Race/ethnicity			
White, non-Hispanic	-2.7	-3.3	-2.1
	(2.91)	(4.21)	(3.86)
Black, non-Hispanic	4.4*	3.6	5.1*
	(1.85)	(3.63)	(2.29)
Hispanic	-1.6	-1.9	-1.3
	(1.92)	(3.00)	(2.64)
Other, non-Hispanic	-0.1	1.5	-1.8
	(1.12)	(1.60)	(1.67)
Health			
In fair or poor health	-2.3	-3.2	-1.7
	(1.19)	(1.84)	(1.57)
Has a chronic health condition	-0.7	-1.1	-0.3
	(2.44)	(3.56)	(3.41)

Notes: Estimates report the coefficient for T, a binary treatment variable equal to one if the respondent is at least 26 years old. All regressions include age, age-squared, and their interactions with the treatment variable. FPG stands for federal poverty guidelines. *p<0.05, **p<0.01, ***p<0.001. Standard errors are in parenthesis.

Table 4.3: The effect of turning 26 on labor market and health coverage outcomes for young adults

	All		Male		Female	
Labor market outcome and labor-related coverage measures						
Employed	4.4		4.7		3.4	
	(2.24)		(3.17)		(3.05)	
In the labor force	3.3		7.5	**	-0.3	
	(1.78)		(2.54)		(2.80)	
Employed full-time	2.3		-0.2		4.6	
	(2.80)		(3.76)		(3.94)	
Employer-sponsored insurance (ESI)	-2.2		-3.8		-0.6	
	(2.64)		(3.58)		(3.78)	
ESI offered	4.8		3.7		6.1	
	(2.89)		(4.00)		(4.1)	
Insurance coverage and insurance-related measures						
Uninsured	2.1		1.6		2.7	
	(2.17)		(3.54)		(2.70)	
Public	0.5		-1.0		1.8	
	(1.76)		(2.12)		(2.78)	
Private	-2.6		-0.9		-3.9	
	(2.27)		(3.49)		(3.21)	
Directly purchased private insurance	4.4	*	6.2	*	2.9	
	(2.04)		(2.63)		(3.23)	
Insurance coverage is worse (than prior year)	15.1	***	12.2	*	17.6	***
	(3.62)		(5.78)		(4.90)	

Notes: Estimates report the coefficient for T, a binary treatment variable equal to one if the respondent is at least 26 years old. In addition to the set of control variables, all regressions include age, age-squared, and their interactions with the treatment variable. *p<0.05, **p<0.01, ***p<0.001. Standard errors are in parenthesis.

Table 4.4: Robustness testing: alternative samples modeling the effect of turning 26 on labor market and health coverage outcomes for young adults

	Panel (A) Models using a narrower age band (ages 25 up to 27)			Panel (B Models using a wider age band (ages 23 up to 29)			Panel (C) Models with unmarried Individuals (ages 24 up to 28)		
	All	Male	Female	All	Male	Female	All	Male	Female
Labor market and labor-related coverage measures									
Employed	4.0	6.6	0.3	4.5 *	4.3	4.0	5.4	7.9 *	3.3
	(3.23)	(4.64)	(4.38)	(1.89)	(2.47)	(2.60)	(2.93)	(3.93)	(4.11)
In the labor force	2.6	5.9	-0.7	3.2 *	5.8 **	0.6	3.0	9.8 **	-2.87
	(2.67)	(4.03)	(3.89)	(1.58)	(2.01)	(2.42)	(2.37)	(3.29)	(3.56)
Employed full-time	0.7	1.5	0.8	1.7	1.3	2.4	-0.1	-1.9	1.9
	(3.75)	(5.28)	(5.26)	(2.21)	(2.91)	(3.23)	(3.52)	(4.59)	(5.2)
Employer-sponsored insurance (ESI)	0.3	0.7	0.5	-2.9	-3.7	-1.9	-6.8 *	-6.51	-6.2
	(4.14)	(5.17)	(5.53)	(2.01)	(2.82)	(2.83)	(3.20)	(4.34)	(5.23)
ESI offered	7.8	11.6 *	2.8	3.7	2.5	5.1	7.9 *	5.3	11.4 *
	(4.08)	(5.78)	(5.71)	(2.41)	(2.45)	(3.52)	(3.69)	(5.02)	(5.20)
Insurance coverage and insurance-related measures									
Uninsured	0.1	-4.1	3.9	3.3	2.8	4.0	4.1	4.2	3.5
	(3.23)	(5.31)	(3.78)	(1.84)	(2.80)	(2.33)	(2.71)	(4.45)	(3.69)
Public	0.3	-1.5	1.5	-0.4	-1.9	1.0	0.5	-0.7	2.18
	(2.64)	(2.94)	(3.84)	(1.30)	(1.66)	(2.19)	(0.74)	(2.37)	(3.63)
Private	-0.2	5.1	-4.8	-3.0	-1.3	-4.7	-4.2	-3.7	-4.7
	(0.15)	(5.24)	(4.39)	(1.91)	(2.75)	(2.65)	(2.88)	(4.42)	(4.05)
Directly purchased	6.3	12.4 **	3.0	3.7 *	4.1	3.3	5.1	5.2	5.2
	(3.52)	(4.96)	(4.63)	(1.55)	(2.31)	(2.17)	(1.21)	(2.96)	(5.09)
Insurance coverage is worse (than prior year)	17.2 **	17.5	17.0 *	10.7 ***	9.5 *	11.8 ***	14.1 ***	10.4	17.8 ***
	(6.14)	(9.13)	(7.89)	(2.62)	(4.18)	(3.50)	(3.32)	(5.41)	(4.31)

Notes: Estimates report the coefficient for T, a binary treatment variable equal to one if the respondent's age is equal to or greater than threshold age. In addition to the set of control variables, all regressions include age, age-squared, and their interactions with the treatment variable.
*p<0.05,**p<0.01,***p<0.001. Standard errors in parenthesis.

Table 4.5: Robustness testing: model specification testing

	Panel (A) Ages 24-28, years 2004-2006 Threshold: age 26			Panel (B) Ages 26.5-30.5, years 2011-2013 Threshold: age 27.5			Panel (C) Ages 21.5-25.5,years 2011-2013 Threshold: age 23.5		
	All	Male	Female	All	Male	Female	All	Male	Female
Labor market outcomes and labor-related coverage measures									
Employed	0.7 (2.37)	0.8 (2.68)	0.6 (3.71)	-2.1 (1.89)	-3.1 (2.80)	-0.9 (2.73)	2.8 (2.47)	2.0 (3.25)	2.8 (3.50)
In the labor force	0.0 (2.14)	2.8 (2.79)	-2.9 (3.53)	-2.9 (1.61)	-3.2 (2.10)	-2.4 (2.62)	2.0 (2.02)	3.0 (2.77)	0.3 (3.20)
Employed full-time	-3.4 (2.53)	-2.5 (2.94)	-3.9 (4.40)	-2.4 (2.18)	-0.1 (2.90)	-5.5 (3.72)	-6.2 (3.39)	-7.8 (4.59)	-5.0 (4.66)
Employer-sponsored insurance (ESI)	-0.2 (2.49)	-0.3 (3.10)	0.3 (3.85)	-1.6 (2.25)	0.2 (3.11)	-3.3 (3.42)	-4.6 (2.52)	-3.6 (3.18)	-6.0 (3.34)
ESI offered	-5.1 (2.89)	-2.0 (4.00)	-8.8 (4.51)	2.0 (2.46)	2.5 (3.49)	1.3 (3.61)	0.5 (3.48)	-0.9 (4.50)	1.6 (4.75)
Insurance coverage and insurance-related measures									
Uninsured	-4.0 (2.59)	-4.4 (3.69)	-3.4 (3.31)	-1.5 (2.11)	0.10 (3.19)	-2.9 (2.74)	0.4 (2.24)	-4.6 (3.31)	4.6 (3.19)
Public	2.9 (1.67)	1.3 (1.80)	4.6 (2.77)	0.3 (1.38)	1.0 (1.72)	-0.6 (2.16)	0.0 (1.71)	-0.8 (2.11)	0.9 (2.46)
Private	0.8 (2.62)	3.2 (3.18)	-1.3 (3.65)	1.0 (2.15)	-1.2 (3.37)	3.3 (2.73)	-0.2 (2.28)	5.2 (3.38)	-5.4 (3.19)
Directly purchased	1.3 (2.08)	2.5 (2.49)	0.1 (3.38)	0.3 (1.79)	-0.8 (2.41)	1.3 (2.80)	2.4 (1.31)	2.27 (3.18)	1.5 (2.57)
Insurance coverage is worse (than prior year)	*n/a*	*n/a*	*n/a*	0.6 (1.22)	4.4 (3.88)	-3.3 (3.64)	2.9 (2.42)	-5.6 (3.81)	-3.2 (3.43)

Notes: Estimates report the coefficient for T, a binary treatment variable equal to one if the respondent's age is equal to or greater than threshold age. In addition to the set of control variables, all regressions include age, age-squared, and their interactions with the treatment variable. *$p<0.05$,**$p<0.01$,***$p<0.001$. (SEs)

Figure 4.1: Leisure-consumption model for an individual working part-time prior to aging out of the dependent coverage provision

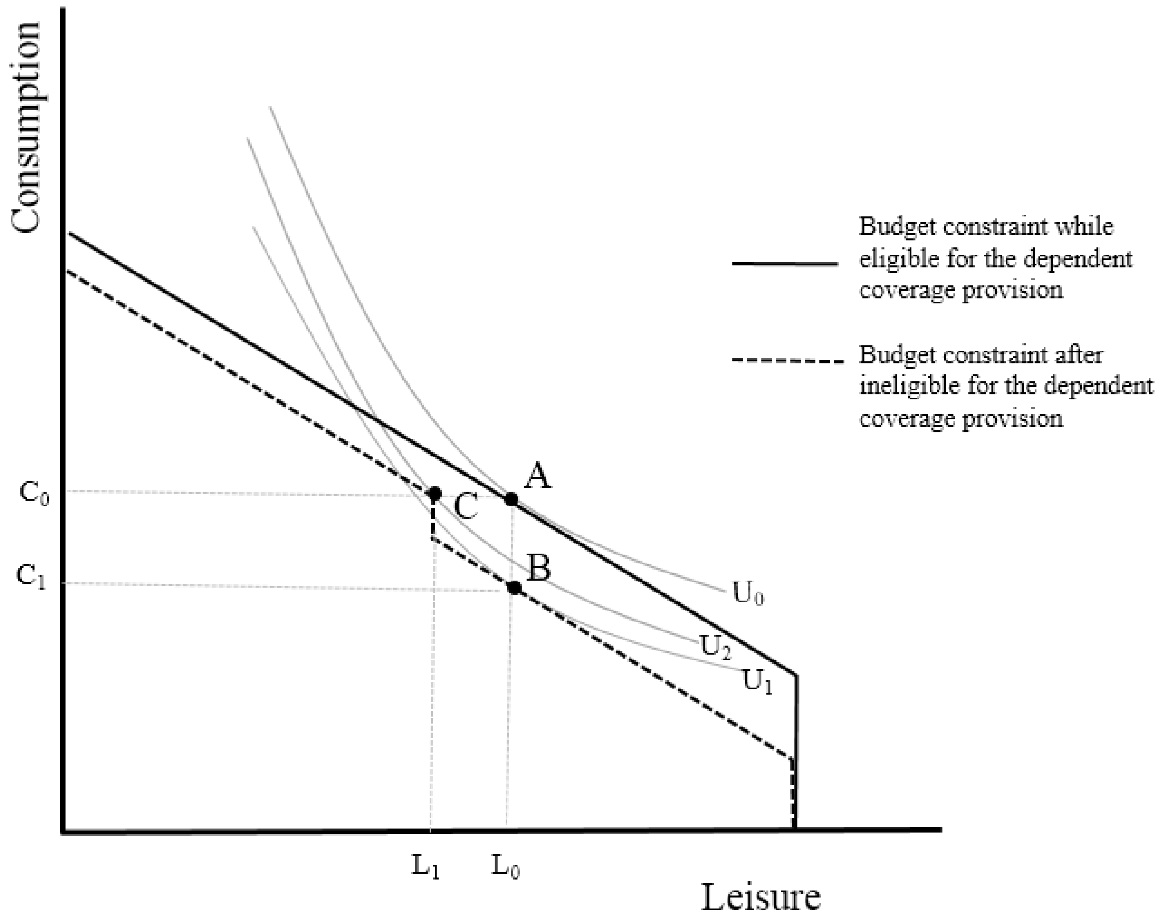


Figure 4.2: Leisure-consumption model for an individual not in the labor force prior to aging out of the dependent coverage provision

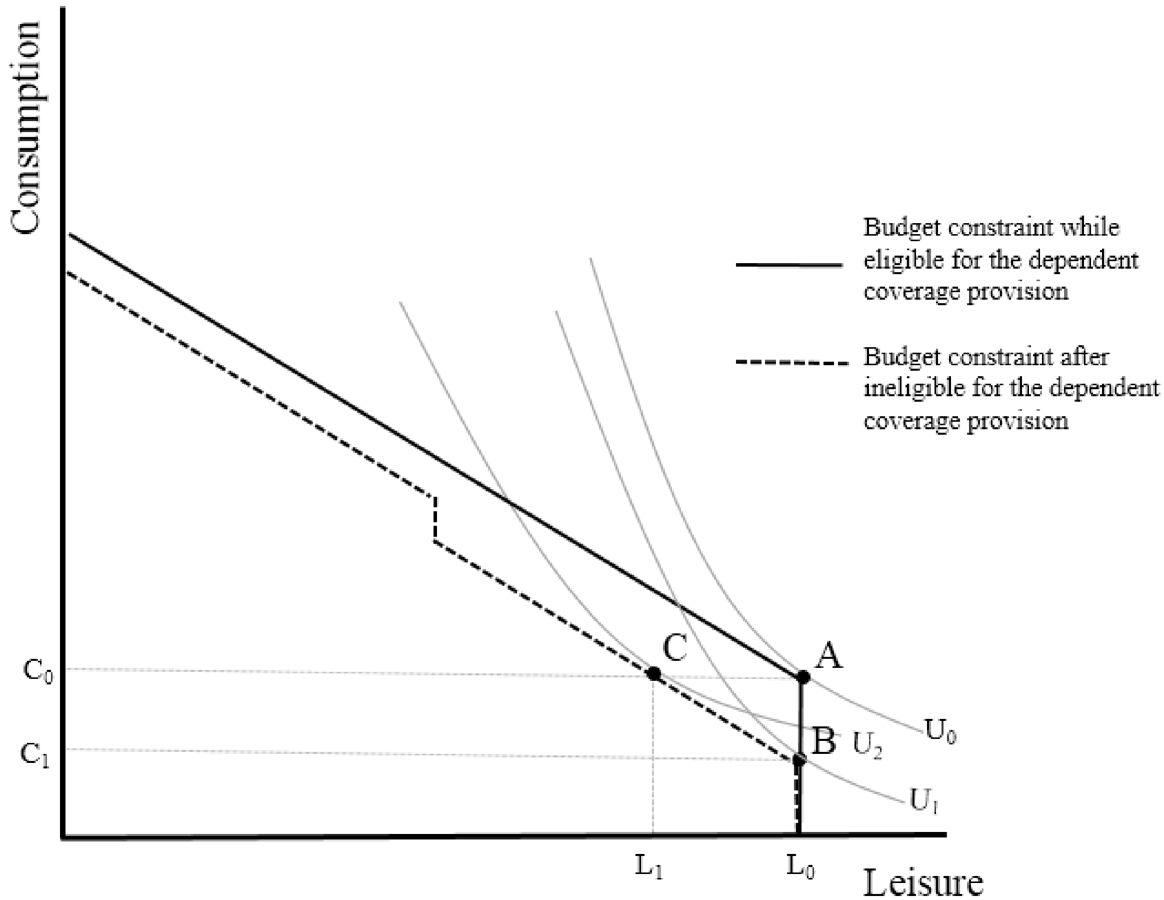


Figure 4.3: Distribution of the number of observations around age 26

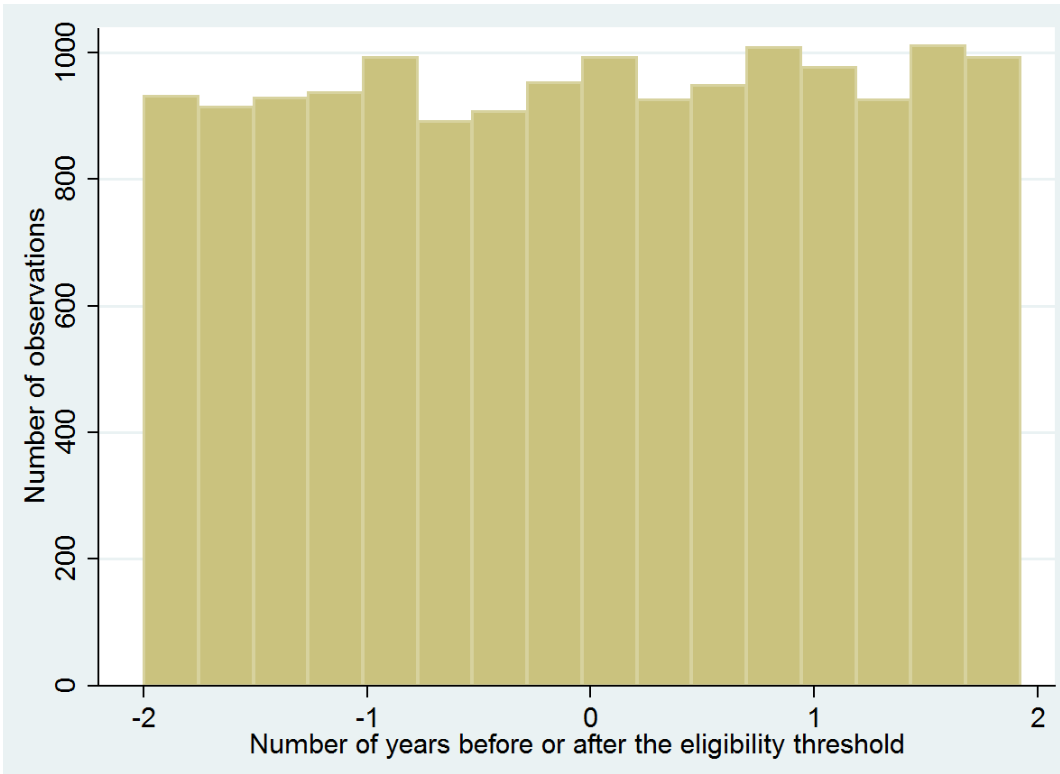
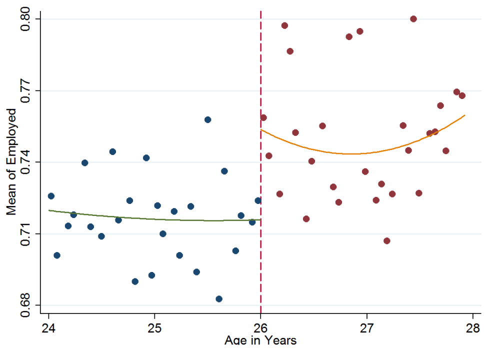
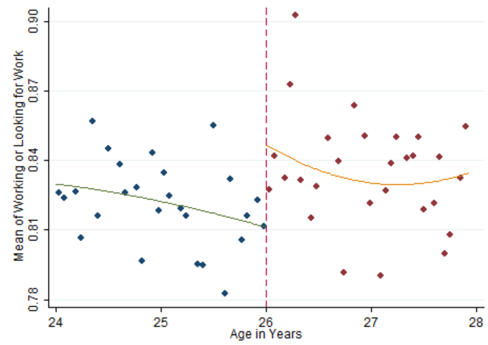


Figure 4.4: Trends in labor and insurance outcomes before and after the Dependent Coverage Provision age threshold for young adulds (aged 24-28)

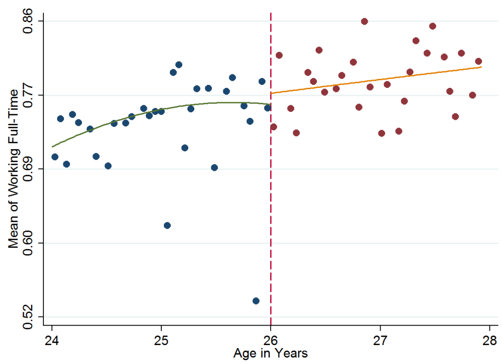
A. Employed



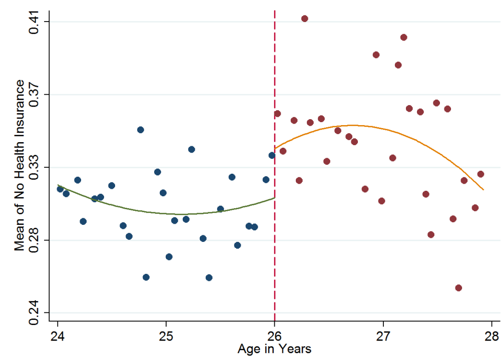
B. Labor Force Participation Rate



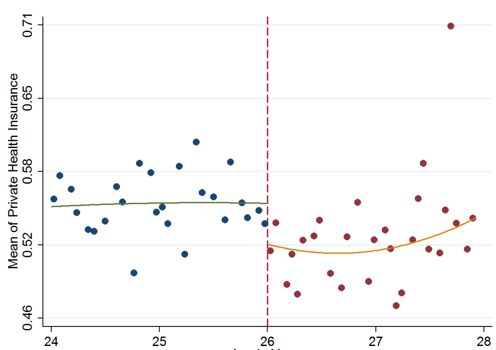
C. Full-Time Employment



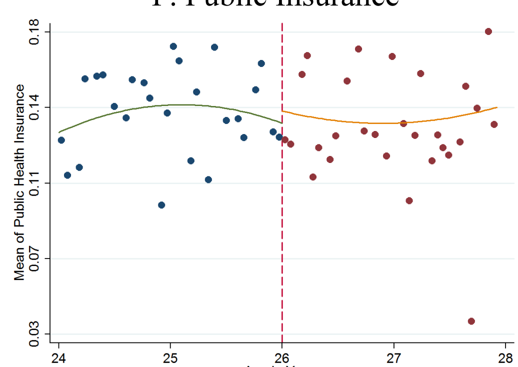
D. Uninsured



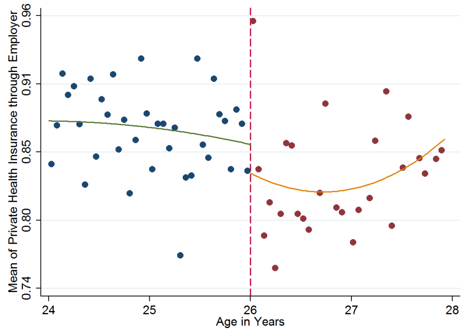
E. Private Insurance



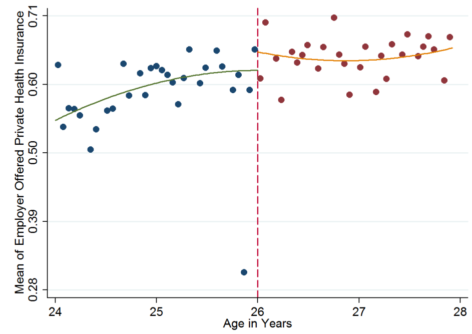
F. Public Insurance



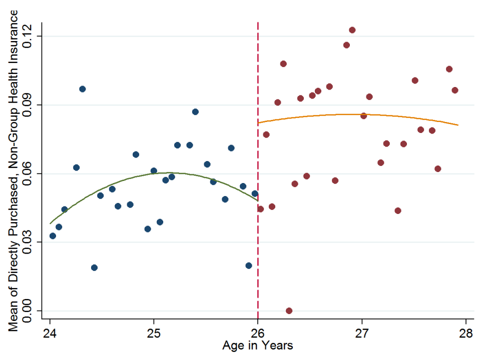
G. Insured through Employer



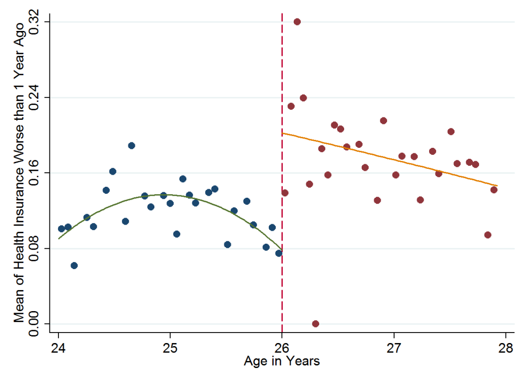
H. Employer Offered Health Insurance



I. Direct Purchase



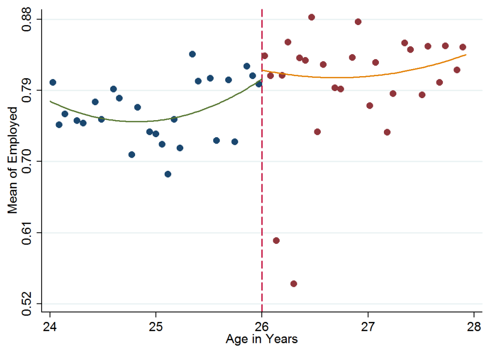
J. Insurance is Worse



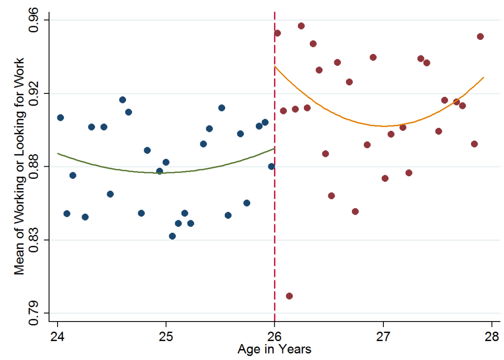
Notes: Mean of the outcome variables for 45 day intervals are plotted. The solid lines on either side of the age-26 cutoff are second order polynomials fitted on individual observations.

Figure 4.5: Trends in labor and insurance outcomes before and after the Dependent Coverage Provision age threshold for young men (aged 24-28)

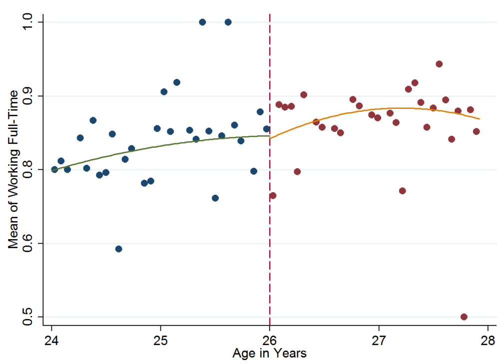
A. Employed



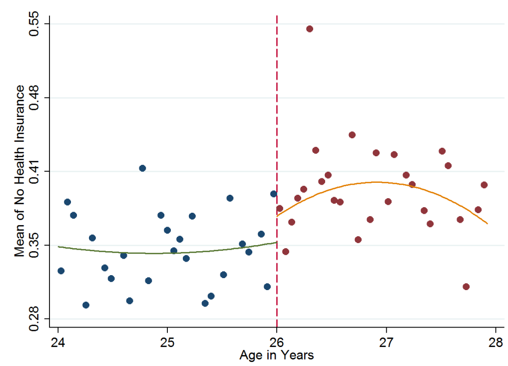
B. Labor Force Participation Rate



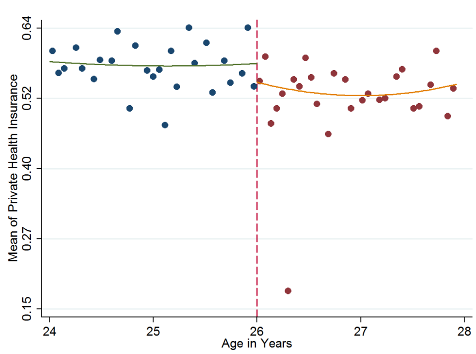
C. Full-Time Employment



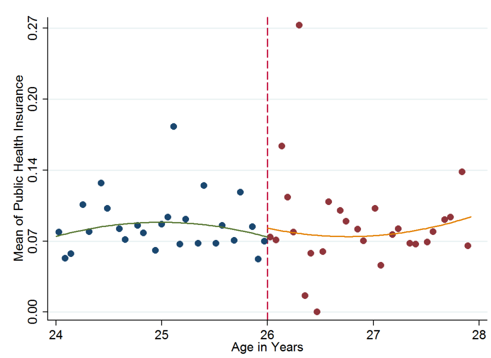
D. Uninsured



E. Private Insurance



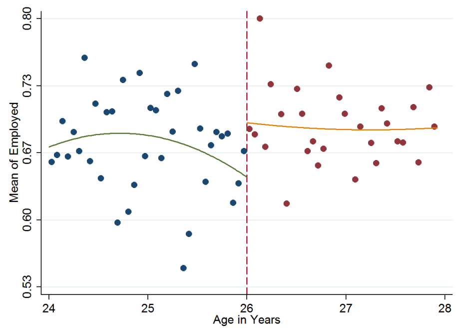
F. Public Insurance



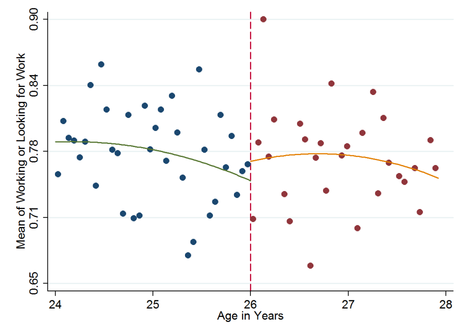
Notes: Mean of the outcome variables for 45 day intervals are plotted. The solid lines on either side of the age-26 cutoff are second order polynomials fitted on individual observations.

Figure 4.6: Trends in labor and insurance outcomes before and after the Dependent Coverage Provision age threshold for young women (aged 24 to 28)

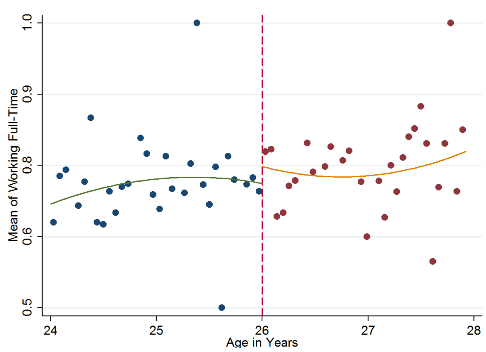
A. Employed



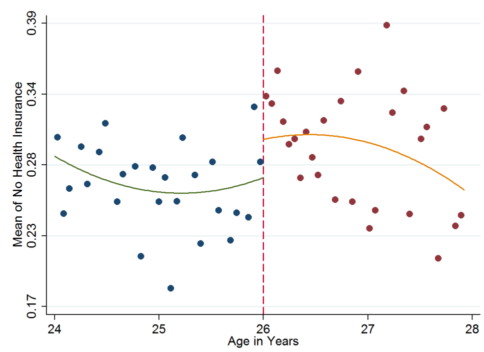
B. Labor Force Participation Rate



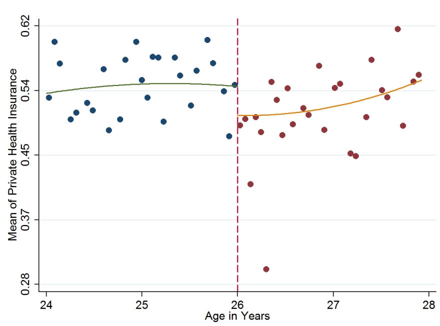
C. Full-Time Employment



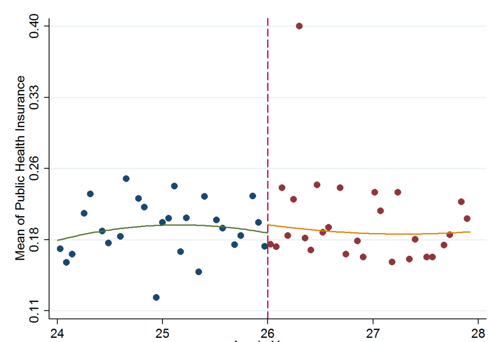
D. Uninsured

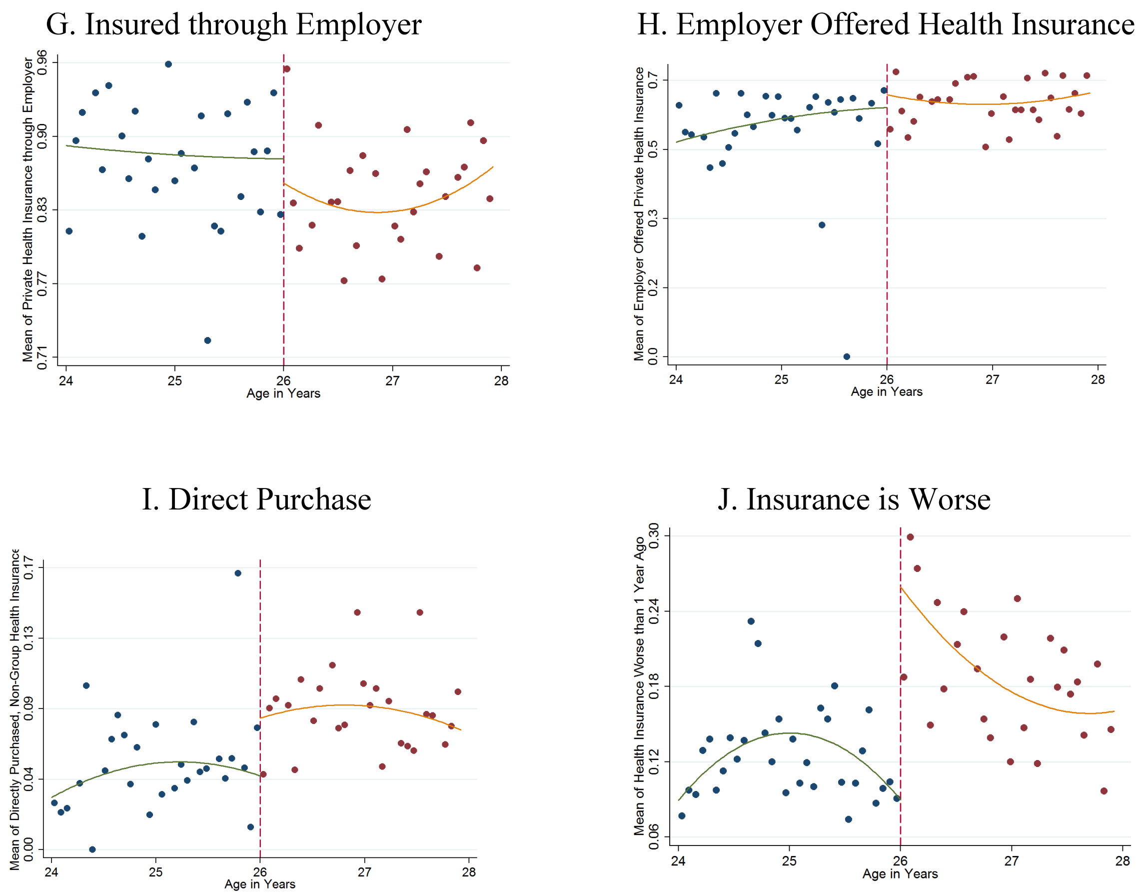


E. Private Insurance



F. Public Insurance





Notes: Mean of the outcome variables for 45 day intervals are plotted. The solid lines on either side of the age-26 cutoff are second order polynomials fitted on individual observations.

Chapter 5

Conclusion

Results from this collection of essay make important contributions to several disciplines within economics—health, labor, and public policy. By using nationally representative data and applying rigorous economic models, I am able to reach policy-relevant conclusion concerning human capital formation of children and young adults.

In Chapter 2, I estimate the impact of maternal depression on academic and non-cognitive outcomes of elementary school-aged children. Results from this essay show that even after addressing omitted variable bias, maternal depression has negatively affects test scores and leads to reduced socioemotional outcomes in children, measured by reductions in self-control, ability to learn, and interpersonal skills, and increases in externalized and internalized problem behavior. In addition, chronicity and severity of maternal depression exacerbate the magnitudes of many of these findings. By applying a bounding methodology, an econometric technique that uses information on observed characteristics to make inferences about the unobserved characteristics not accounted for in the models, these findings address a common concern in the literature—endogeneity. Therefore, this essay moves the literature away from associative relationships and produces meaningful estimates of the effect of maternal depression on non-cognitive human capital measures of children.

Moving forward, I plan to continue using the ECLS-K panel and extend the research surrounding chronicity of maternal depression through a difference-in-differences framework. This methodology takes advantage of the individual-level panel data structure and accounts for multiple time periods and arbitrary treatment patterns. Results from these models can then be compared to the bounding estimates produced in this dissertation.

In Chapter 3, I examine how earlier school start time affects fifth grade cognitive and non-cognitive outcomes. Most school start time research focuses on older students, but this paper demonstrates that changing school start time has negative implications for younger students, too. When schools begin only a small amount earlier (1-29 minutes), there was no effect on children, but when the start time was pushed up 60 or more minutes earlier, there were negative impacts on both test scores and non-cognitive outcomes in children. By estimating the model using an earlier period of time, I am able to confirm that changes in child outcomes were not simply the result of differences in the children whose schools began earlier. Future work will involve seeking permission for use of restricted-level ECLS-K data that contains exact start and end date of the school year so that I can control for such changes that may have occurred concurrently with changes in start time.

Lastly, in Chapter 4, I estimate the effect of aging out of the dependent coverage provision on labor market and health insurance outcomes for young adults. I find large increases in labor force participation and purchase of non-group private coverage for males when they turn 26, as well as increases in health insurance plan dissatisfaction for both males and females concurrent with aging out of the provision. By running a number of robustness checks measuring sample appropriateness and empirical model fit, I am able to state that discontinuities in labor and insurance coverage outcomes occurring at age 26 are due to aging out of the dependent coverage provision and not simply due to changes in demographic and/or economic conditions. While many studies have examined how the provision improved health insurance coverage, health care access, and health, this study fills a gap in the literature by measuring the effect of aging out of the provision.

The significant jump in health insurance plan dissatisfaction at age 26 will be the springboard for future research. Since the NHIS includes a rich set of variables that measure health care barriers (e.g., unmet need for care due to cost, trouble accessing a health care provider, and emergency room use), I plan to analyze the relationship between aging out of the provision and barriers to health care access, use, and affordability. Another extension of this paper is to examine if the dependent coverage provision led to labor market changes in the parents. For example, did parents remain in a job or delay retirement so that their child(ren) might have access to health insurance? For this paper, I plan to utilize the Survey of Income and Program Participation (SIPP), as it allows for distinction between ESI dependent coverage and own ESI and is a longitudinal panel, providing more information over time.

In summary, this dissertation demonstrates that parents, schools, and federal policy affect human capital formation of children and young adults. In order to come up with solutions and recommendations to help children and young adults, policymakers need reliable estimates. This research attempts to support officials in the fields of health care, public policy, and education in their efforts to draft viable, effective policies.

Chapter 6

Bibliography

Access Economics (Firm). & Sane Australia. (2003). *Bipolar disorder: costs: an analysis of the burden of bipolar disorder and related suicide in Australia.* [South Melbourne, Vic.]: SANE Australia.

Altonji JG, Conley, T, Elder T, Taber C. (2011). Methods for Using Selection on Observed Variables to Address Selection on Unobserved Variables. *Yale University Department of Economics Working Paper*.

Altonji JG, Elder TE, Taber CR. An evaluation of instrumental variable strategies for estimating the effects of catholic schooling. *J Hum Resour.* 2005; 40(4):791-821.

American College Health Association. National College Health Assessment: Undergraduate Reference Group Executive Summary, Spring 2013. Available at: http://www.acha-ncha.org/docs/ACHA-NCHA-II_UNDERGRAD_ReferenceGroup_ExecutiveSummary_Spring2013.pdf. Accessed December 9, 2014.

American College Health Association. Undergraduate Reference Group Executive Summary, Spring 2010. Available at: http://www.acha-ncha.org/docs/ACHA-NCHA-II_ReferenceGroup_ExecutiveSummary_Spring2010.pdf. Accessed December 24, 2014.

Antwi YA, Asako SM, Simon K. Effects of federal policy to insure young adults: evidence from the 2010 Affordable Care Act's Dependent-Coverage Mandate." *American Economic Journal: Economic Policy.* 2013; 5(4): 1-28.

Aronson J, Zimmerman J, Carlos L. Improving student achievement by extending school: is it just a matter of time? *Office of Educational Research and Improvement.* 1999.

Automatic Data Processing. "ADP's 2012 study of large employer health benefits: benchmarks for companies with 1,000+ employees." Available at: http://www.adp.com/tools-and-resources/adp-research-institute/research-and-trends/research-item-detail.aspx?id={656665F4-5D63-47D1-8114-57D11D837F60}._Accessed on June 3, 2015.

Banarjee S, Chatterji P, Lahiri K. Effects of psychiatric disorders on labor market outcomes: a latent variable approach using multiple clinical indicators. *CESifo Working Paper Series No. 4260.*

Beardslee WR, Bemporad MD, Keller MB, Klerman GL. Children of parents with major affective disorder: a review. *AM J Psychiat.* 1983; 140(7):825-32.

Becker G (1964) Human Capital, 2nd edn. Columbia University Press, New York, 1975 and 3rd 1994.

Becker G (1981). "A Treatise on the Family," Cambridge, MA: Harvard University Press.

Behrman JR, Wolfe BL. How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation? J Econometrics. 1987; 36(1-2):185-204.

Berger MC Fleisher BM. Husband's health and wife's labor supply. *J Health Econ.*1984; 3(1):63-75.

Bernstein AB, RA Cohen, KM Brett, MA Bush. Marital status is associated with health insurance coverage for working-age women at all income levels, 2007. NCHS Data Brief. 2008; 11.

Bertakis KT, R Azari, JL Helms, EJ Callahan, JA Robbins. Gender differences in the utilization of health care services. *J Fam Pract.* 2000; 49(2):147-152.

Biro FM, Greenspan LC, Galvez MP, et al. Onset of breast development in a longitudinal cohort. *Pediatrics.* 2013;132:1019-1027.

Blau D, Gilleskie D. Health insurance and retirement of married couples. *Journal of Applied Econometrics*. 2006; 21: 935-953.

Brennan PA, Hammen C, Andersen MJ, Bor W, Najman JM, Williams. Chronicity, severity, and timing of maternal depressive symptoms: relationships with child outcomes at age 5. *Dev Psychol*. 2000; 36: 759-766.

Campbell SB, Cohen JF, Meters T. Depression in first-time mothers: Mother-infant interaction and depression chronicity. *Dev Psychol.* 1995; 31(3):349-357.

Carpenter C, Dobkin C. The effect of alcohol consumption on mortality: regression discontinuity evidence from the minimum drinking age. *American Economic Journal: Applied Economics.* 2009;1(1):164-182.

Carskadon MA. Sleep in adolescents: the perfect storm. *Pediatr Clin North Am.* 2011; 58(3): 637-647.

Carskadon MA, Acebo C. Intrinsic circadian period in adolescents versus adults from forced desynchrony. *Sleep.* 2005;28(Suppl.):A21.

Carskadon MA, Acebo C. Regulation of sleepiness in adolescents: update, insights, and speculation. *Sleep.* 2002;25(6): 606–614.

Carksadon MA, Viera , Acebo. Association between puberty and delayed phase preference. *Sleep.* 1993:16 (3).

Carskadon MA, Wolfson AR, Acebo C, Tzischinsky O, Seifer, R. Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. *Sleep*. 1998;21(8):871–881.

Casey P, Goolsby S, Berkowitz C, et al. Maternal depression, changing public assistance, food security, and child health status. *Pediatrics.* 2004; 113(2): 298-304.

Centers for Medicare & Medicaid Services. "Young Adults and the Affordable Care Act." Protecting young adults and eliminating burdens on families and businesses. Available at: http://www.cms.gov/CCIIO/Resources/Files/adult_child_fact_sheet.html. Accessed on December 9, 2014.

Clayton KM, Stewart SM, Wiebe DJ, McConnel CE, Hughes CW, White PC. Maternal depressive symptoms predict adolescent healthcare utilization and charges in youth with type 1 diabetes (T1D). *Health Psychol.* 2013;32(9):1013-22.

COBRA Insurance Headquarters. How much does COBRA insurance cost? Available at: http://www.cobrainsurancehq.com/ Accessed on May 20, 2015.

Cohn JF, Tronick EZ. Three-month-old infants' reaction to simulated maternal depression. *Child Dev.* 1983; 54(1):185-193.

Coile, CC. Health shocks and couples' labor supply decisions. *NBER Working Paper.* 2004; N.10810.

Crowley SJ, Acebo C, Carskadon MA. Sleep, circadian rhythms, and delayed phase in adolescence. *Sleep medicine.* 2007;8(6):602–612.

Cunha F, Heckman JJ, Lochner LJ, Masterov DV. Interpreting the evidence of life cycle skill formation. *Handbook of the economics of education.* 2006; a:697-812.

Cunha F, Heckman JJ. The technology of skill formation. *Am Econ Rev.* 2007; 97(2):31-47.

Curcio G, Ferrara M, De Gennaro L. Sleep loss, learning capacity and academic performance. *Sleep Med Rev.* 2006;10(5):323-337.

Currie J. Early childhood education programs. *J Econ Perspect.* 2001; 15(2):213-238.

Currie J, Madrian BC. "Health, Health Insurance and the Labor Market." Ed. D. Card and O. Ashenfelter. *Handbook of Labor Economics.* Vol. 3. N.p.: Elsevier, 1999. 3309-416. Print

Currie J, Stabile M. Child mental health and human capital accumulation: The case of ADHD. *J Health Econ.* 2006; 25(6):1094-1118.

Currie J, Stabile M. Socioeconomic status and health: Why is the relationship stronger for older children? *Am Econ Rev.* 2003; 93:1813-1823.

Dee TS, Ha W, Jacob BA. The effects of school size on parental involvement and social capital: evidence from the ELS: 2002. *Brookings Pap Ed Pol.* 2006/2007. 77-97.

Dexter D, Bijwadia J, Schilling D, Applebaugh G. Sleep, sleepiness and school start times: a preliminary study. *WMJ : official publication of the State Medical Society of Wisconsin*. 2003; 102(1):44–46.

Ding W, Lehrer SF, Rosenquist JN, Audrain-McGovern J. The impact of poor health on academic performance: new evidence using genetic markers. *J Health Econ.* 2009; 28(3):578-597.

Downey G, Coyne JC. Children of depressed parents: An integrative review. *Psychol Bull.* 1990; 108: 50-76.

Drake C, Nickel C, Burduvali E, Roth T, Jefferson C, Badia P. The pediatric daytime sleepiness scale (PDSS): sleep habits and school outcomes in middle-school children. *Sleep: Journal of Sleep and Sleep Disorders Research.* 2003;26(4):455-458.

D'Souza RM, Stazdins L, Clements MS, Broom DH, Parslow R, Rodgers B. The health effects of jobs; status, working conditions, or both? *Aust NZ Psychiat.* 2007; 29(3):222-228.

Duggan M, NB Ellison, C Lampe, A Lenhart, M Madden. "Social media update 2014," Pew Research Center, January 2015. Available at: http://www.pewinternet.org/2015/01/09/social-media-update-2014/. Accessed on April 3, 2015.

Edwards F. Early to rise? The effect of daily start times on academic performance. *Economics of Education Review.* 2012;31:970-983.

El-Sheikh M, Bagley EJ, Keily MK, Erath SA. Growth in body mass index from childhood into adolescence: the role of sleep duration and quality. *Journal of Early Adolescence.* 2014;34(8):1145-1166.

Employee Benefit Research Institute, November 2012. Self-insured health plans: state variation and recent trends by firm size. Available at: http://www.ebri.org/pdf/notespdf/ebri_notes_11_nov-12.slf-insrd1.pdf.

Employee Retirement Income Security Act of 1974 (ERISA), Pub. L. No. 93-406, 88 Stat. 829 (codified as amended in scattered sections of 5 U.S.C., 18 U.S.C., 26 U.S.C., 29 U.S.C., and 42 U.S.C.).

Epstein R, Chillag N, Lavie, P. Starting times of school: effects on daytime functioning of fifth-grade children in Israel. *Sleep*. 1998;21(3):250–256.

Fendrich M, Warner V, Weissman MM. Family risk factors, parental depression, and psychopathology in offspring. *Dev Psychol.* 1990; 26:40-50.

Field T. Infants of depressed mothers. *Infant Behav Dev.* 1995; 18(1):1-13.

Fletcher JM. Adolescent depression: Diagnosis, treatment, and educational attainment. *J Health Econ.* 2008; 17(11):1215-1235.

Fletcher JM. All in the family: Mental health spillover effects between working spouses. *B.E. J of Econ Analysis and Pol.* 2009; 9(1): Article 3.

Fletcher JM, Lehrer SF. Using genetic lotteries within families to examine the causal impact of poor health on academic achievement. *NBER Working Paper.* 2009; N.15148.

Fletcher JM, Wolfe B. Child mental health and human capital accumulation: The case of ADHD revisited. *J Health Econ.* 2008; 27:794-800.

Frank RG, Meara E. The effect of maternal depression and substance abuse on child human capital development. *NBER Working Paper.* 2009; N.15314.

Frederiksen K, Rhodes J, Reddy R, Way N. Sleepless in Chicago: tracking the effects of adolescent loss during the middle school years. *Child Dev*. 2004;75:84–95.

Fry R. “A rising share of young adults live in their parents’ home,” Pew Research Center. August 2013. Available at: http://www.pewsocialtrends.org/2013/08/01/a-rising-share-of-young-adults-live-in-their-parents-home/. Accessed on April 6, 2015.

Fügenschuh A. Solving a school bus scheduling problem with integer programming. *European Journal of Operational Research.* 2009;193(3):867 – 884.

Gau SF, Soong WT. Sleep problems of junior high school students in Taipei. *Sleep*. 1995;18(8):667–673

Gilliam M, Forbes EE, Gianaros PJ, Erickson KI, Brennan LM, Shaw DS. Maternal depression in childhood and aggression in young adulthood: evidence for mediation by offspring amygdala-hippocampal volume ratio. *J Child Psychol Psychiatry.* 2014, Nov 26. [Epub ahead of print]

Goodman SH, Rouse MH, Connell AM, et al. Maternal depression and child psychopathology: a meta-analytic review. *Clin Child FamPsychol Rev.* 2010; 14(1):1-27.

Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *J Clin Psy.* 2015; 76(2):155-162.

Gruber J, Madrian BC. Health insurance, labor supply, and job mobility: a critical review of the literature. 2002. NBER Working Paper No. 8817.

Gruber J, Madrian B. Health insurance availability and the retirement decision. *American Economic Review*. 1995; 85: 938-948.

Guvenen F, Karahan F, Ozkan S, Song J. What do data on millions of U.S. workers reveal about life-cycle earnings risk? *Federal Reserve Bank of New York Staff Reports.* 2015: 710.

Ham J, Shore-Sheppard L. The impact of public health insurance on labor market transitions", *Journal of Public Economics.* 2003; 66: 349-382.

Hansen M, Janssen I, Schiff A, Zee PC, Dubocovich ML. The impact of school daily schedule on adolescent sleep. *Pediatrics.* 2005;115(6):1551-1561.

Hann D, Winter K, Jacobsen P. Measurement of depressive symptoms in cancer patients: evaluation of the Center for Epidemiological Studies Depression Scale (CES-D). *J Psychosom Res.* 1999; 46:437-443.

Hanushek EA. Measuring investment in education. *J Econ Perspect.* 1996; 10(4):9-30.

Harris Interactive Poll. "National study reveals how teens are shaping and reshaping their wireless world." Available at: http://www.harrisinteractive.com/NewsRoom/PressReleases/tabid/446/ctl/ReadCustom%20Default/mid/1506/ArticleId/191/Default.aspx. Accessed on June 5, 2015.

Herman-Giddens ME, Steffes J, Harris D, et al. Secondary sexual characteristics in boys: data from the pediatric research in office settings network. *Pediatrics.* 2012;130(5):1058-1068.

Healthcare.gov. Health coverage for children under 26. 2015. Available at: https://www.healthcare.gov/young-adults/children-under-26/. Accessed on June 3, 2015.

Heckman JJ. Life cycle consumption and labor supply: An explanation of the relationship between income and consumption over the life cycle. *Am Econ Rev.* 1974; 65(1):188-194.

Heckman JJ. Skill formation and the economics of investing in disadvantaged children. *Science.* 2006; 312(5782):1900-1902.

Heckman JJ. The technology and neuroscience of capacity formation. *Proceedings of the National Academy of Sciences.* 2007; 204(3):13250-13255.

Hinrichs P. When the bell tolls: the effects of school starting times on academic achievement. *Education Finance and Policy.* 2011;6(4):486-507.

Hirano KG, Imbens W, Ridder G. Efficient estimation of average treatment effects using the estimated propensity score. *NBER Technical Working Paper.* 2000; N.251.

Honig M, Dushi I. Household demand for health insurance: price and spouse's coverage. Hunter College. 2005 Working Paper.

Hysenbegasi A, Hass SL, Rowland CR. The impact of depression on the academic productivity of university students. *J Ment Health Policy.* 2005; 8(3):141-151.

Kaiser Family Foundation. 2013 employer health benefits survey. Available at: http://kff.org/report-section/ehbs-2013-section-2/. Accessed on June 4, 2015.

Kaiser Family Foundation. Employer health benefits, 2012 summary of findings. Available at: http://kaiserfamilyfoundation.files.wordpress.com/2013/03/8346-employer-health-benefits-annual-survey-summary-of-findings-0912.pdf. Accessed November 26, 2014.

Karoly LA, Ragowski JA. The effect of access to post-retirement health insurance on the decision to retire early. *Industrial and Labor Relations Review*. 1994; 48: 103-123.

Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush JA, Walters EE,Wang PS. The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *J Americ Med Assoc.* 2003; 289(23):3095-3105.

Klein D, Lewinsohn P, Seeley J, Olino T. Psychopathology in the adolescent and young adult offspring of a community sample of mothers and fathers with major depression. *Psychol Med.* 2005; 35: 353-365.

Kochhar R. "Two years of economic recovery: women lose jobs, men find them." Pew Research Center. July 2011. Available at: http://www.pewsocialtrends.org/2011/07/06/two-years-of-economic-recovery-women-lose-jobs-men-find-them/. Accessed on April 6, 2015.

Krauth BV. Bounding a linear causal effect using relative correlation restrictions. *Working Paper.* 2011.

Lee, DS, Moretti E, and Butler MJ. "Do voters affect or elect policies? Evidence from the U.S. House". *Quarterly Journal of Economics.*2004; **119**(3): 807–859.

Lovejoy CM, Graczyk PA, O'Hare E, Neuman G. Maternal depression and parenting behavior: A meta-analytic review. *Clin Psychol Rev.* 2000; 20(5):561-592.

Lyons-Ruth K, Connell DB, Grunebaum HU, Botein S. Infants at social risk: Maternal depression and family support services as mediators of infant development and security of attachment. *Child Dev.* 1990; 61(1):85-98.

Lyons-Ruth K, Connell DB, Grunebaum HU. The depressed mother and her one-year-old infant. Evnironment, interaction, attachment, and infant development. *New Directions for Child and Adolescent Development.* 1986; 34:61-82.

McLearn KT, Minkovitz CS, Strobino DM, Marks E, Hou W. Maternal depressive symptoms at 2 to 4 months post-partum and early parenting practices. *Arch Pediat Adol Med.* 2006; 160:279-284.

Madrian BC. Employment-based health insurance and job mobility: is there evidence of job-lock?" *Quarterly Journal of Economics.* (1994); 109: 26-54.

Magee CA, Caputi P, Iverson DC. The longitudinal relationship between sleep duration and body mass index in children: a growth mixture modeling approach. *J Dev Behav Pediatr.* 2013:34(3):165-173.

Meyer B, Viscusi WK, Durbin D. Worker's compensation and injury duration: evidence from a natural experiment. *The American Economic Review.* 1995; 83: 322-340.

Minnesota Population Center and State Health Access Data Assistance Center, *Integrated Health Interview Series: Version 5.0.* Minneapolis: University of Minnesota, 2012. For more information, see: http://www.ihis.us.

Moffitt R, Wolfe BL. The effect of the Medicaid program on welfare participation and labor supply. *Review of Economics and Statistics*. 1992; 74(4): 615-626.

Moore PJ, Adler NE, Williams DR, Jackson JS. Socioeconomic status and health: the role of sleep. *Psychosom Med.* 2002; 64: 334-344.

Mulcahy A, Harris K, Finegold K, Kellermann A, Edelman L, Sommers BD. Insurance coverage of emergency care for young adults under health reform. *NEJM.* 2013; 368: 2105-12.

Murray L, Fiori-Cowley A, Hooper R Cooper P. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcomes. *Child Dev.* 1996; 67(5):2512-2526.

National Conference of State Legislatures. Covering young adults through their parents' or guardians' health policy. Available at: http://www.ncsl.org/research/health/dependent-health-coverage-state-implementation.aspx. Accessed November 26, 2014.

Nichols A, Linder S. Why are fewer people in the labor force during the Great Recession? The Urban Institute. 2013: 16.

Olsen C. Part-time work, health insurance coverage, and the wages of married women. In: Alpert, William T., Woodbury, Stephen A. (Eds.), Employee Benefits and Labor Markets in Canada and the United States. *W.E. Upjohn Institute for Employment Research*. 2006: 295-324.

Orvaschel H. Maternal depression and child dysfunction: Children at risk. *Adv Clin Child Psych.* 1983: 169-197.

Oster E. Unobservable selection and coefficient stability: theory and validation. 2015. Tech. Rep No. 19054, National Bureau of Economic Research, Cambridge, MA.

Owens JA, Belon K, Moss P. Impact of delaying school start time on adolescent sleep, mood, and behavior. *Archives of pediatrics & adolescent medicine*. 2010; 164(7), 608–14.

Parsons DO. Health, family structure, and labor supply. *Am Econ Rev.* 1977; 67:703-12.

Perry CD. Does treating maternal depression improve child health management? The case of pediatric asthma. *J Health Econ.* 2008; 27(1):157-173.

Pew research Center. Modern parenthood: roles of moms and dads converge as they balance work and family. Available from: http://www.pewsocialtrends.org/files/2013/03/FINAL_modern_parenthood_03-2013.pdf. Accessed May 30, 2014.

Posner MA, Ash AS. Comparing weighting methods in propensity score analysis. *Working Paper.*

Randall DK (2012). *Dreamland: Adventures in the Strange Science of Sleep.* New York: W.W. Norton & Company. Radloff LS. The CES-D scale: A self-report depression scale for research in the general population. *Appl Psych Meas.* 1977; 1:385-401.

Raposa E, Hammen C, Brennan P, Najman J. The long-term effects of maternal depression: early childhood physical health as a pathway to offspring depression. *J Adol Health.* 2014; 54(1):88-93.

Reschovsky JD, J Hadley. The effect of tax credits for nongroup insurance on health spending by the uninsured. *Health Affairs,* 2004; W4-113-W4-127.

Rice DP, Miller LS. The economic burden of schizophrenia: Conceptual and methodological issues and cost estimates. *Handb Ment Health Econ Health Pol.* 1996.

Rust J, Phelan C. How Social Security and Medicare affect retirement behavior in a world of incomplete markets. *Econometrica*, 1997; 65(4): 781-831.

Shaw DS, Hyde LW, Brennan LM. Early predictors of boys' antisocial trajectories. *Dev and Psychopathology.* 2012; 24: 871-888.

Siegel M, Bradley E, Gallo W, Kasl S. The effect of spousal mental and physical health on husbands' and wives' depressive symptoms, among older adults. *J Aging Health.* 2004; 16(3):398-425.

Silverstein M, Augustyn M, Cabral H, Zuckerman B. Maternal depression and violence exposure: Double jeopardy for child school functioning. *Pediatrics.* 2005; 119: 792-800.

Section 514 (29 U.S.C. § 1144).

Sommers BD. ASPE issue brief. Number of young adults gaining insurance due to the Affordable Care Act now tops 3 million. June 2012.

Sommers BD, Buchmueller T, Decker SL, Carey C, Kronick R. The Affordable Care Act has led to significant gains in health insurance and access to care for young adults. *Health Affairs.* 2013 (32):165-174.

Sommers BD, Kronick R. The Affordable Care Act and insurance coverage for young adults. JAMA. 2013; 307(9): 913-914.

The Commonwealth Fund Health Insurance Tracking Survey of Young Adults, 2013. Available at: http://www.commonwealthfund.org/. Accessed December 9, 2014.

Tarricone R, Gerzeli S Montanelli R, Frattura L, Percudanci M, Racogni G. Direct and indirect costs of schizophrenia in community psychiatric services in Italy. *Health Policy.* 2000; 51(1):1-18.

Thorleifsdottir B, Björnsson JK, Benediktsdottir B, Gislason T, Kristbjarnarson H. Sleep and sleep habits from childhood to young adulthood over a 10-year period. *Journal of Psychosomatic Research*. 2002;53(1):529–537.

Tomohara A, Lee H. Public Health Insurance Reform and Women's Labor Supply Decisions. *International Business and Economics Research Journal*. 2006; 5: 19-28.

Tourangeau K, Nord C, Lê T, Sorongon AG, Najarian M. *Early Childhood Longitudinal Study, Kindergarten Class of 1998*–99 (ECLS-K): Combined User's Manual for the ECLS-K Eighth-Grade and K-8 Full Sample Data Files and Electronic Codebooks. Washington,DC: US Department of Education; 2009.

Trochim WMK. The regression-discontinuity design: an introduction. 2006. Available from: http://www.socialresearchmethods.net/kb/quasird.htm.

Unlu F. California class size reduction reform: New findings from the NAEP. *Working Paper.* 2005.

United States Census Bureau. America's Families and Living Arrangements: 2012: Children (C table series). http://www.census.gov/hhes/families/data/cps2012C.html. Accessed March 30, 2014.

United States Census Bureau. Table 597. Labor Force Participation Rates by Marital Status, Sex, and Age: 1970 to 2010. Available at: http://www.census.gov/compendia/statab/2012/tables/12s0598.pdf. Accessed December 24, 2014.

United States Department of Education, National Center for Education Statistics, Schools and Staffing Survey (SASS), "Public School Data Files," 2001-2004.

United States Department of Labor. "Young adults and the Affordable Care Act: protecting young adults and eliminating burdens on businesses and families." 2015. Available at: http://www.dol.gov/ebsa/faqs/faq-dependentcoverage.html. Accessed on June 5, 2015.

Unlu F. California class size reduction reform: New findings from the NAEP. *Working Paper.* 2005.

Wahlstrom K. Changing times: Findings from the first longitudinal study of later high school start times. *NASSP Bulletin.* 2002; 86(633): 3-21.

Weissman M, Wickramatratne P, Warner V, John BA, Prusoff A, Merikangas KR. Assessing psychiatric disorders in children: Discrepancies between mothers' and childrens' reports. *Arch Gen Psychiat.* 1987; 44(8):747-753.

Welsch-Allis G, Ye W. Psychopathology in children of parents with recurrent depression. *J Abnorm Child Psych.* 1988; 16: 17-28.

Wilcox-Gok V, Marcotte DE, Farahati F, Borkoski. *Early onset of depression and high school dropout*. Amsterdam: Elsevier Science; 2004.

Wilcox-Gok V, Temple J. Maternal depression and child health: evidence on child obesity. *NIU Working paper.* 2006.

Williamson EJ, Forbes A, White IR. Variance reduction in randomized trials by inverse probability weighting using the propensity score. *Statist Med.* 2014; 33:721-737.

Wolfson AR, Carskadon MA. A survey of factors affecting high school start times. *NASSP Bulletin.* 2005;89:47–66.

Wolfson AR, Carskadon MA. Sleep schedules and daytime functioning in adolescents. *Child development*. 1998;69(4):875–887.

Wolfson AR, Carskadon MA. Understanding adolescents' sleep patterns and school performance: A critical appraisal. *Sleep Medicine Review*. 2003; 7:491-503.

Wooldridge JM. Inverse probability weighted estimation for general missing data problems. *J Econometrics.* 2007; 141(2):1281-1301.

Wooldridge JM. Econometric Analysis of Cross-Section and Panel Data. Cambridge, MA: The MIT Press; 2002.

Yelowitz AS. The Medicaid notch, labor supply, and welfare participation: evidence from eligibility expansions. *Quarterly Journal of Economics*. 1995; 100: 909-940.

Yoruk BK, CE Yoruk. The impact of minimum legal drinking age laws on alcohol consumption, smoking, and marijuana use: evidence from a regression discontinuity design using exact date of birth. *J Health Econ.* 2011; 30(4):740-752.

Chapter 7

Appendices

Appendix A: Coefficients from logit models predicting maternal depression (kindergarten)

	(1) Any depression		(2) Severe depression	
Child characteristics				
White, non-Hispanic	-0.091		0.199	
Black, non-Hispanic	0.121		0.412	
Hispanic	-0.228		-0.068	
Asian, non-Hispanic	-0.596	*	-0.690	
Female	-0.130	*	-0.125	
Age (in months)	0.001		0.010	
Weight at birth	-0.041		-0.084	*
Disabled	0.328	***	0.475	**
English is a second language	0.032		0.090	
Family characteristics				
2 parents	-0.296		1.057	
Single parent household	0.146		1.704	
Mom was a teen at child's birth	0.165		0.275	
No. of nights family has dinner together	-0.079	***	-0.113	***
No. of children under age 18 in the house	0.006		0.055	
Socioeconomic status (mother's education, father's education, income, mother's occupation, father's occupation)	-0.498	***	-0.656	***
School characteristics				
Urban	-0.042		-0.222	
Midwest	0.111		-0.061	
South	0.031		-0.074	
West	0.186		-0.035	
Public school	0.536	***	0.307	
Neighborhood problems index	-0.006		0.014	
Free lunch program	0.301	***	0.195	
Teacher turnover	-0.002		0.006	
Over crowdedness is a problem in the school	-0.018		-0.043	
Less than 10% of school is minority	0.031		-0.069	
Constant	-1.240		-4.047	***

Notes: *p<0.05, **p<0.01, ***p<0.001. N=15,677 (Col.1); N=14,509 (Col.2).

Appendix B: Inverse probability weighted regressions

Panel A: Presence of any maternal depression, test scores, and socioemotional outcomes

	(1) Kindergarten		(2) Third Grade		(3) Eighth Grade	
Dependent Variable	$\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		$\tilde{\beta}$, (S.E.), $[\widetilde{R}]$		$\tilde{\beta}$, (S.E.), $[\widetilde{R}]$	
Math	-1.076		-1.118		-0.344	
	(0.627)	[0.2678]	(1.146)	[0.2425]	(1.324)	[0.2499]
Reading	-1.194	**	-2.098		-1.041	
	(0.440)	[0.1712]	(1.298)	[0.2525]	(1.695)	[0.2784]
Learning	-0.070	**	-0.045			
	(0.022)	[0.1406]	(0.032)	[0.1186]		
Control	-0.046	*	-0.067	*		
	(0.021)	[0.0800]	(0.031)	[0.0688]		
Interpersonal	-0.053	*	-0.064	*		
	(0.022)	[0.0773]	(0.033)	[0.0837]		
Externalizing	0.0415		0.061			
	(0.021)	[0.0925]	(0.032)	[0.0827]		
Internalizing	0.066	***	0.069	*		
	(0.019)	[0.0311]	(0.027)	[0.1323]		

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N=13,978 (Col. 1); 11,831 (Col. 2); 6,255 (Col.3). Models control for child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school).

Panel B: Presence of severe maternal depression, test scores, and socioemotional outcomes

	(1) Kindergarten		(2) Third Grade		(3) Eighth Grade	
Dependent Variable	$\tilde{\beta}$, (S. E.), $[\widetilde{R}]$		$\tilde{\beta}$, (S. E.), $[\widetilde{R}]$		$\tilde{\beta}$, (S. E.), $[\widetilde{R}]$	
Math	-2.225	***	-2.617		-0.289	
	(0.640)	[0.2289]	(1.860)	[0.2315]	(1.983)	[0.2383]
Reading	-1.511		-4.240		-1.024	
	(.784)	[0.1683]	(2.228)	[0.2407]	(2.720)	[0.2724]
Learning	-0.129	***	-0.0833			
	(0.040)	[0.1352]	(0.046)	[0.1129]		
Control	-0.044		-0.040			
	(0.040)	[0.0763]	(0.049)	[0.0634]		
Interpersonal	-0.090	*	-0.127	*		
	(0.040)	[0.0745]	(0.051)	[0.0782]		
Externalizing	0.060					
	(0.351)	[0.0896]				
Internalizing	0.091	**				
	(0.030)	[0.0299]				

Notes: *p<0.05, **p<0.01, ***p<0.001 indicate significant differences between children of depressed mothers and non-depressed mothers. N=13,106 (Col. 1); 11,400 (Col. 2); 6,078 (Col.3). Models control for child characteristics (age, race/ethnicity, gender, disability status, English as a second language, and birth weight), family characteristics (structure of household, if mother was a teenager at child's birth, socioeconomic status, and number of children in the family), and school characteristics (region, urban/rural status, teacher turnover, overcrowding, type of school, if fewer than 10% of students are of minority status, and a neighborhood problems index monitoring crime, drugs, violence, gangs, and tension near the school).

Appendix C: Regression-adjusted impact of changing school start time on child outcomes

(earlier restricted to those beginning before 8:15 in 5th grade)

	1-29 minutes earlier	30-59 minutes earlier	60+ minutes earlier		Any later	
Math score						
All	0.36	-1.88	-2.98		0.17	
	(1.13)	(1.54)	(1.81)		(1.00)	
Boys	-0.04	-0.31	0.69		1.04	
	(1.71)	(1.58)	(2.34)		(0.93)	
Girls	0.92	-3.88	-5.39	*	-0.40	
	(1.24)	(2.08)	(2.34)		(1.21)	
Reading score						
All	0.87	-1.23	0.87		0.32	
	(.77)	(1.68)	(0.77)		(0.97)	
Boys	0.89	-3.49	-4.19	***	1.27	
	(1.07)	(2.44)	(1.22		(1.49)	
Girls	1.09	1.58	1.09		-0.38	
	(1.09)	(2.02)	(1.09)		(1.18)	
Internalizing problem behavior						
All	0.04	0.07	-0.31		-0.01	
	(0.04)	(0.09)	(0.25)		(0.04)	
Boys	0.05	0.02	-0.45		0.01	
	(0.05)	(0.12)	(0.37)		(0.04)	
Girls	0.01	0.13	-0.11		-0.03	
	(0.08)	(0.12)	(0.11)		(0.06)	
Externalizing problem behavior						
All	-0.04	0.01	0.06		-0.07	*
	(0.06)	(0.09)	(0.04)		(0.03)	
Boys	-0.04	-0.02	0.15	*	-0.05	
	(0.07)	(0.11)	(0.06)		(0.05)	
Girls	-0.12	-0.05	0.07		0.01	
	(.09)	(0.07)	(0.08)		(0.06)	
Approaches to learning						
All	-0.11	-0.15	0.32		0.04	
	(0.06)	(0.08)	(0.29)		(0.04)	
Boys	-0.10	-0.22	0.53		0.06	
	(0.07)	(0.11)	(0.44)		(0.05)	
Girls	-0.04	0.08	-0.05		-0.08	*
	(0.06)	(0.08)	(0.05)		(0.04)	
Self-control						
All	-0.71	-0.01	-0.16		0.06	

	(0.06)		(0.06)		(0.21)		(0.04)
Boys	-0.02		-0.11		-0.42	*	0.06
	(0.06)		(0.06)		(0.19)		(0.05)
Girls	-0.14		0.12		0.16		0.05
	(0.09)		(0.09)		(0.16)		(0.04)
Interpersonal skills							
All	-0.17	**	-0.20	***	0.08		0.01
	(0.07)		(0.06)		(0.19)		(0.04)
Boys	-0.14		-0.27	***	0.18		0.02
	(0.08)		(0.06)		(0.28)		(0.04)
Girls	-0.21	*	-0.10		-0.07		
	(0.08)		(0.09		(0.10)		

Notes: N=7,245 (all); N=3,679 (boys); N=3,566 (girls). *(**)(***) Significantly different from the estimate for children whose schools did not change start time at the (.05) (.01) (.001) level, two-tailed test.

Appendix D: Testing of higher-order age terms in a regression discontinuity framework

	Model with age-cubed and its interaction with treatment			Model without age-cubed and its interaction with treatment		
Predictor	Coeff	SE	t-statistic	Coeff	SE	t-statistic
Treatment	1.21	0.404	2.99	1.21	0.295	4.08
age	-0.002	0.004	-0.4	-0.003	0.0014	-2.3
age squared	1.63E-06	0.00001	0.13	-4.32E-06	1.85E-06	-2.33
age cubed	5.34E-09	1.10E-08	0.49	n/a	n/a	n/a
age*T	-0.002	0.004	-0.37	0.002	0.002	1.13
age_sq*T	6.09E-06	2.00E-05	0.4	5.05E-06	2.49E-06	2.03
age cubed*T	-1.19E-08	1.37E-08	-0.87	n/a	n/a	n/a
Marginal effect of treatment	14.8	4.97	2.98	14.8	3.63	4.07

Note: All models adjust for the full set of covariates listed in Table 4.2.

Appendix E: The effect of turning 26 on labor market and health coverage outcomes for young adults (excluding individuals 1 month younger or older than age 26)

	All		Male		Female	
Labor market outcome and labor-related coverage measures						
Employed	5.8	*	4.5		6.1	
	(2.68)		(3.84)		(3.7)	
In the labor force	3.9		7.9	**	0.14	
	(2.16)		(2.54)		(3.45)	
Employed full-time	1.7		-0.4		3.53	
	(3.12)		(4.18)		(4.61)	
Employer-sponsored insurance (ESI)	-6.7	*	-6.2		-7.3	
	(2.70)		(4.10)		(4.46)	
ESI offered	4.8		2.2		7.6	
	(3.56)		(4.77)		(4.83)	
Insurance coverage and insurance-related measures						
Uninsured	2.6		2.1		3.2	
	(2.47)		(4.11)		(3.42)	
Public	-0.9		-0.3		-1.3	
	(1.92)		(2.37)		(3.00)	
Private	-1.3		-1.7		-0.8	
	(2.56)		(4.02)		(3.61)	
Directly purchased private insurance	4.9	*	5.2		4.9	
	(2.42)		(2.99)		(4.07)	
Insurance coverage is worse (than prior year)	17.0	***	11.4		22.0	***
	(3.86)		(6.05)		(5.25)	

Notes: Estimates report the coefficient for T, a binary treatment variable equal to one if the respondent is at least 26 years old. In addition to the set of control variables, all regressions include age, age-squared, and their interactions with the treatment variable. *p<0.05, **p<0.01, ***p<0.001. Standard errors are in parenthesis.