8-2018

Essays in Education & Economics

Mary Elizabeth Glenn
University of Tennessee, mglenn6@vols.utk.edu

To the Graduate Council:

I am submitting herewith a dissertation written by Mary Elizabeth Glenn entitled "Essays in Education & Economics." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Economics.

Celeste K. Carruthers, Major Professor

We have read this dissertation and recommend its acceptance:

Donald J. Bruce, Robert F. Kronick, Georg Schaur

Accepted for the Council:
Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Essays in Education & Economics

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Mary Elizabeth Glenn
August 2018

ACKNOWLEDGEMENTS

Thank you to my mother, for her boundless love and patience, particularly throughout this process.
Thank you to my father, who has always believed in me and pushed me to think analytically.
Thank you to all of the faculty and staff at the Boyd Center, who have always supported me to the utmost. Y'all are the best.
Thank you to Celeste Carruthers who has provided invaluable guidance and mentorship throughout this long process. Thank you to Don Bruce, who has remained steadfast in his conviction I can do this, even when my own conviction has wavered.
Thank you to Jill Welch and Becky Davis, for all of the lunches and Diet Coke we have shared. In addition, thank you to Dale Cuva and Jamie Osborne, roommates who helped make Knoxville a place to call home.

ABSTRACT

Within this dissertation, I examine how behaviors are affected by and affect education decisions. Within the first chapter, I examine how infectious disease impacted school enrollment behaviors in a historical context. Both the second and third chapters focus on college financing decisions. The second chapter explores how student loan uptake impacts the retirement behaviors of *parents* of students. The third chapter examines if and how large-scale merit aid scholarships affect student loan uptake.

In the first chapter, I analyze the effect of poliomyelitis outbreaks on school enrollment choices. This chapter adds to a growing literature on avoidance behavior within health economics, which focuses on how the fear of a disease changes behavior and creates additional costs. I find support for the idea that polio outbreaks resulted in lower likelihood of school enrollment for kindergarten-aged children and particularly for kindergarten-aged children with stay-at-home mothers. This result implies the channel for avoidance behavior is the *ability* to change behavior, here as a result either of a child's age or a family's income structure.

In the second chapter, coauthored with Celeste Carruthers, we look at an unstudied topic within the student loan literature: how student loans influence parents of students. We examine the effect of student loan presence on several dimensions of retirement behaviors. We find that student loan presence results in significantly fewer dollars in retirement savings, later expected retirement age, and a higher likelihood of being employed and in the labor force in some specifications. However, there is evidence these results are driven by unobservables.

In the third chapter, I examine the effect of large-scale merit aid programs on a broad measure of household debt. This debt measure includes student loan debt, credit card debt, medical bills, legal bills, and loans from relatives. Previous literature found large-scale merit aid programs result in lower a likelihood of student loan uptake and lower amounts of student loan debt. I employ a difference-in-difference strategy to exploit differences over time and states in the introduction of the programs. I do not find that these programs result in changes in student loan uptake or amount.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

SOMETHING TO FEAR OTHER THAN FEAR ITSELF

"... first of all, let me assert my firm belief that the only thing we have to fear is fear itself..."
-- Franklin D. Roosevelt

"Everyone was afraid of polio, especially those who saw it all the time ... In at least two instances during polio epidemics, hospital staff were felled by their own contagion of what turned out to be hysterical paralysis, brought on by fear."
-- *Patenting the Sun*, Jane S. Smith

Introduction

Some public health threats may be costly because these threats create anxiety that induces behavioral change. A public health threat with rare but severe outcomes may induce larger behavioral responses than anticipated from a rational and risk neutral populace. The economic epidemiology literature refers to these changes as avoidance or aversion behavior. How large could the avoidance behavior induced by anxiety surrounding a severe public health threat be? I examine this question about avoidance behavior in the context of poliomyelitis in mid-century America by answering a more specific question: did polio induce changes in five-and six-year-olds' school enrollment behavior and, if so, how large were those behavioral changes? I examine behavioral changes in this context because polio was a disease with a striking tail end risk (paralysis) the population at large was documented to be frightened of, and were particularly frightened of it affecting five- to nine-year-olds (Sirkin, 1960). In this case, if parents engaged in aversion behavior to limit their child's exposure to polio, parents would limit their own child's exposure to a peer populace where polio is believed to be circulating. School is the most likely place where a young child would be exposed to their peers, and parents would have the most discretion over five- and six-year-olds' enrollment since kindergarten was optional during the time I examine. Further, families with stay-at-home mothers and kindergarten-aged children would have additional discretion to keep the child at home.

I analyze parents' responsiveness to the uncertainty represented by a polio outbreak through the impact of reported state-level polio incidence rates on individual enrollment decisions for kindergarten-aged children. I postulate avoidance behavior will be the most pronounced for families with more discretion over such decisions. I examine this possibility in two ways: (1) I compare the impact of polio on the enrollment decisions for five- and six-year-olds to the impact of polio on seven- and eight-year-olds and nine- and ten-year-olds and (2) I compare the impact of polio on five- and six-year-olds in a family with a stay-at-home mother to the impact of polio on five- and six-year-olds in a family with dual incomes. In both cases, I find evidence that families with additional discretion displayed more avoidance behavior, regardless of if this discretion stemmed from a child's age or family structure. I conclude avoidance behavior can have large impacts at the individual-level and is especially pronounced when such behavior is relatively low cost for the individual or family.

The reasons for studying kindergarteners are twofold: (1) children around this age range were understood to be the prime age to become infected with symptomatic polio,[1] and (2) kindergarten remains optional in most states, suggesting parents had additional discretion over these children's school enrollment.[2]

The purpose of the analysis is to document behavioral change created by fear of a disease rather than to estimate the full welfare effects of this particular behavioral change. Such a welfare estimate would require assumptions about (1) if parents on average were delaying their children's kindergarten entrance or skipping kindergarten enrollment all together and (2) the benefits of attending kindergarten in mid-century America. In a modern context, research has shown that enrollment decisions for this age group may have large impacts on children later in life. Fitzpatrick et al (2010) suggest an additional year of schooling in kindergarten or first grade results in a gain of a standard deviation in both reading and math. Berlinski et al (2008) finds that preprimary attendance reduces levels of grade repetition later. Attending formal kindergarten or remaining at home, rather than attending informal daycare, has been found to have positive impacts on children's test scores in later childhood (Feinstein, Robertson & Symons, 1999). Cascio (2009), the one paper I am aware of that explores the benefits of kindergarten in a historical context, finds that increased kindergarten availability in the 1960s and 1970s decreased the likelihood that White children would be imprisoned later in life and increased the likelihood White children would graduate high school. She does not find similar results for Black children, and she does not find effects on earnings, labor supply, or the receipt of government benefits, possibly because kindergarten during this earlier time period was comparatively low intensity. Whether the results from this literature apply in the context of the 1940s through 1960s is an open question for future research.

What could be the educational impact be of a disease with high profile but rare and severe outcomes, like paralysis from polio? The highest annual polio case count ever recorded in the US in a year was 52,870 in 1952. The likelihood that a person who contracts poliovirus will be paralyzed for any length of time is less than 1% (Center for Disease Control, 2016). Regardless, parents checked newspapers daily to see if there was a polio outbreak in their community or state (Rustein, 1957; Smith, 1990). Schools, churches, movie theaters, and swimming

---

[1] Children five- to nine-years-old were designated as the first priority group for polio vaccination in 1955 (Sirken & Brenner, 1960).

[2] See Table 5.3: Types of state and district requirements for kindergarten entrance and attendance, by state: 2014 (National Center for Education Statistics, 2015).

pools routinely closed for the duration of polio outbreaks (Smith, 1990). Fear of infectious diseases with severe outcomes has not waned since the mid-century: in the 1980s, a school board barred Ryan White, a hemophiliac who contracted AIDS, from attending school even when authorities knew the likelihood White would spread the disease was miniscule (Specter, 1985). Further, during the Zika outbreak of 2016, presidential candidates lauded the idea of quarantining individuals from countries with Zika outbreaks despite calls from public health officials such measures would have little benefit (Maron, 2016). If public health threats cause such anxiety, such indirect impacts only make them costlier.

I look at the impact of uncertainty and associated avoidance behavior surrounding polio on kindergarten enrollment with individual-level pooled ordinary least squares analysis. Using individual-level Census data from 1940, 1950, and 1960, I determine the impact of an increase in the polio incidence rate in a state during the previous twelve months on the likelihood a kindergarten-age child will have attended kindergarten in that same state in a given time span prior to a reference date given by the Census. In 1950 and 1960, this time span is the two months before February 1; in 1940, this time span is one month before March 1. Following Bleakley (2007), I take this mid-year attendance as a signal of enrollment. I use a conservative measure of a state-year's average 1 in 100,000 incidence rate of polio in the twelve months before the reference date for a given Census. I create this average twelve-month incidence rate using weekly incidence rates of polio from Project Tycho (van Panhuis et al, 2013) and only include state-years with 75% or more non-missing state-week observations. In the analyses, I include personal and family controls and state-level controls for education, public sanitation, and crude birth rates from the U.S. Census Bureau (U.S. Census Bureau 1942, 1943, 1946, 1952, 1953, 1962a, 1962b, & 1963).

In these analyses, I assume reported polio incidence rates are exogenous to kindergarten enrollment decisions once I condition on observables. False inference will result in the presence of any endogenous channel impacting both kindergarten enrollment and reported incidence rates. Several aspects of the analyses mitigate this possibility. Of foremost importance is the high variability of polio incidence rates in different years in each state. Polio incidence rates increased from 1940 to 1950, but subsequently declined with the invention of polio vaccines in 1955 and 1960. Although all states display this relationship, they do so with high levels of volatility suggesting that polio incidence rate spikes functioned as exogenous shocks. To guard further against the possibility of endogeneity, I include epidemiological drivers of polio in all analyses. In addition, I present falsification and robustness tests for the main analyses, and I show that polio has the largest impacts for children whose families had the most discretion over the child's enrollment decisions.

I find evidence polio significantly impacted the enrollment decisions for the age group that mid-century parents had the most discretion over (five- to six-year-olds). Moreover, the coefficients of interest are more significant and have greater magnitude when the analyses focus on families with additional flexibility because of the presence of a homemaker to take care of a child who is not enrolled in school. Families with dual incomes do not have significant coefficients of interest at any level in any specification. The results of the analyses suggest polio created a pattern of avoidance behavior resulting in negative impacts on school enrollment for children in the age range to attend kindergarten. The main results hover between marginal significance and insignificance but suggest an additional new case of polio out of 100,000 people in the average week may have decreased kindergarten enrollment by 4 percentage points. Results for families with a stay-at-home mother imply that for this subpopulation, one additional case of population-adjusted polio in the average week decreased the likelihood a child would be enrolled in kindergarten by a little less than 5 percentage points. The implication is that avoidance behavior can be quite strong in response to an illness with severe but rare outcomes especially when such responses are less costly.

Background

The History of Polio

Poliomyelitis is a viral infection that enters through the mouth and multiplies in the gastrointestinal tract, eventually being spread and expelled through feces. Polio is spread primarily through the oral-fecal route so proper sanitation impacts the disease's spread.[3]

Current belief is that polio's method of contagion (the oral-fecal route) is central to understanding the epidemics that arose in developed nations in the late nineteenth century through the mid-twentieth century. A commonly accepted hypothesis is that prior to modern sanitation, polio existed at subclinical levels throughout the population:

> In unsanitary conditions … children are more uniformly infected very early in life and are more likely to experience mild disease. It has been proposed that the late-nineteenth-century invention of modern plumbing and sewage containment led to the shift toward epidemic polio by preventing widespread infantile exposure to mild poliovirus. Once someone has been infected with poliovirus, lifelong immunity develops that prevents future reinfection. The prevention of common infantile polio subsequently allowed children to be infected with more virulent strains later in life. (Kunschner, 2008, p. 547)

This popular theory of the polio's sudden appearance in the late nineteenth century is called the hygiene hypothesis. More recently, increases in the birth rate in developed nations, rather than increases in public sanitation, have been proposed as the mechanism that gave rise to epidemic polio during the late nineteenth century (Martinez-Bakker & Rohani, 2015).

Even in a polio epidemic, paralysis is not a common outcome for infected individuals; polio rarely manifests itself this intensely. Out of the people in the

[3] Some oral-oral transfer is hypothesized but is not viewed to be the main method of the disease's spread (Center for Disease Control, 2016).

population who contract poliovirus, most (up to 72%) exhibit no symptoms (Center for Disease Control, 2015). Twenty-four percent of people who contract poliovirus experience mild symptoms including “fever, headache, and sore throat” (Kunschner, 2008; Center for Disease Control, 2015). Fewer than 1% of people with poliovirus experienced paralysis (Kunschner, 2008; Center for Disease Control, 2015). Further, most paralytic polio patients in mid-century America recovered sufficiently to live normal lives:

> Even in the relatively rare group of paralytic cases, 50 to 75 per cent of those afflicted recover completely without any treatment at all. An additional 10 to 20 per cent make fairly satisfactory recoveries. This leaves only a small fraction of patients who sustain grave aftereffects. (Sulkin, 1946, p. 384)

Notably, of those who contracted paralytic polio, 5% died from the disease (Kunschner, 2008).

Regardless of the rates of mortality or morbidity, polio epidemics were particularly frightening. The first recorded polio epidemic in the United States was in Vermont in 1843, and while there were brief lulls during the winter and some years, as the nineteenth century became the twentieth century, polio only struck more fiercely (Smithsonian Behring Center, 2016).

Even as new technologies, like the iron lung and rocking bed, were developed and lowered the case-to-death ratio, polio cases continued spreading with little hope the medical community could stem the tide of disease without a vaccine. In 1954, both Jonas Salk and Albert Sabin separately developed polio vaccines: Salk created a vaccine with a killed poliovirus and Sabin created a live polio vaccine where the poliovirus was weakened (Smithsonian Behring Center, 2016). These vaccines were not a foolproof panacea, though they were and continue to be effective and were received eagerly by the populace (Smith, 1990). The vaccines had effectiveness rates of 60% to 90% (Sirkin & Brenner, 1960); while this vaccination rate is high enough to create herd immunity, it does not guarantee immunity when only certain individuals in the population are vaccinated as was the case in the early years of the vaccine. In 1957, variation in vaccine penetration was high as evidenced by differences between Census regions: 10.1% of 5- to 9-year-olds in New England were not inoculated, while 22.7% of children in the same age group were not inoculated in the West South Central region. The case for universal vaccination was not helped by early vaccine failures. The most horrific example occurred with Cutter Manufacturing in California, which in 1955 did not perform Sabin’s procedure for deactivating the polio vaccine properly. Tragedy ensued; the Cutter incident “caused 40 000 cases of polio, leaving 200 children with varying degrees of paralysis and killing 10” (Fitzpatrick, 2006).

In fact, in 1957, polio eradication in the United States was far from certain and polio remained a real fear for many parents. A 1957 *Atlantic* article even hinted the true reliability of the Salk polio vaccine remained undetermined:

> The evidence that the presently available polio vaccine does not decrease the number of carriers, and the clear-cut vaccine failures, make it apparent that

> polio will not be eradicated by this means. Unless the Salk vaccine can be improved, other kinds of vaccine must be developed. (Rustein, 1957)

Polio & Uncertainty

Polio's impact on society concerned not only health outcomes but also behaviors and beliefs. How sensitive were mid-century parents to polio threats? Rustein (1957) notes that

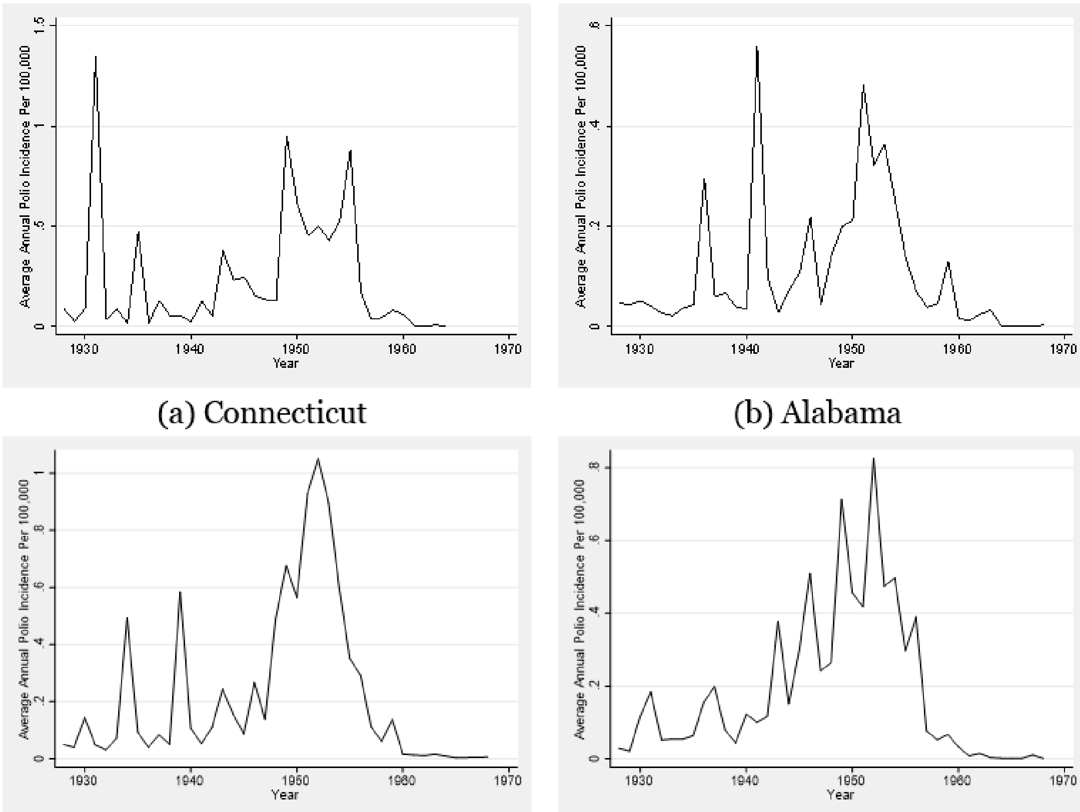
> Polio has its peak occurrence each summer, when parents anxiously note the location of each new case. In the past they have stood by helplessly when the disease struck nearby, watching through the passing months to learn whether this was a "polio year" in their town or state and breathing easily only when cold weather came. (Rustein, 1957)

Rustein's reference to a "polio year" illustrates that it was understood how noisy polio incidence rates could be in the same state over different years. In Figure 1, I present average weekly polio incidence rates for four states to demonstrate how spikes in polio incidence rates functioned as exogenous shocks.

Parental concern over the location of polio infection is central to the identification strategy, as awareness and conditional exogeneity of polio incidence rates are necessary for causal inference. Evidence points to both parents and communities being highly sensitive to polio outbreak location. Entire communities would place themselves under quarantine in the wake of a polio outbreak:

> In 1930 an outbreak centered in Middletown, Connecticut, caused Wesleyan University to cancel its football season and prompted 141 students to quit school. That same fall, local health officials closed schools in Topeka, Kansas, and banned public meetings in Los Angeles, California. ... [In the summer of 1931] Los Angeles witnessed an epidemic so severe the city health services began to break down. Ambulances and stretchers blocked the streets in front of Los Angeles County Hospital, where patients were turned away by frightened hospital employees. In 1935 Boston was hit, the entire city of Annapolis, Maryland, was quarantined, and President Roosevelt, himself a polio victim, called off a national Boy Scout Jamboree in Virginia. In 1936 churches and resorts in Alabama were closed, Chicago was swept by a large epidemic, and Tulsa, Oklahoma, shut down tight. (Smith, 1990, p. 39)

The rare but severe outcome that everyone feared and that caused these disruptions was paralytic polio. Paralytic polio patients were placed behind glass and quarantined – they wrote messages to visitors on chalkboards. If a patient was paralyzed so they could no longer swallow on their own, two respiration tubes were installed: one for feeding and one to syphon fluid out of the lungs. If a patient was paralyzed so they could no longer breathe on their own, they were placed inside the notorious iron lung (Smith, 1990).



(a) Connecticut (b) Alabama

(c) Arizona (d) Illinois

Figure 1. Average Weekly Polio Incidence Rates for One State in Each Census Region.

As previously noted, most paralytic polio patients recovered sufficiently to live normal daily lives. However, those patients who did not recover became emblematic of the horror polio could wreak:

> In 1939 a woman gave birth while in an iron lung; by that time, it had been proven that you could survive in an iron lung for ten years. (Smith, 1990, p. 39)

Parents were afraid of an event with a low likelihood but that rare event was terrifying and strongly motivating toward any behavioral change that might lessen its likelihood.

Previous Literature on Avoidance Behavior

A large literature examines the behavioral changes and direct costs caused by public health threats, including the impact of threats on educational outcomes in historical contexts. However, much of this literature on health threats focuses on the direct effect of the threat, referred to as the "cost of illness" in the health economics literature. Cost of illness includes the impact of having an illness on educational outcomes, rather than the anxiety the disease creates and this anxiety's impact on educational outcomes. Bleakley (2007) examines the impact of hookworm, an intestinal parasite that causes fatigue but has a relatively low publicity profile and low mortality. The large-scale public initiative in the early twentieth century to treat hookworm and raise awareness of preventative measures resulted in significant gains in school attendance, quality of education, and eventual wages for those most likely to be infected with hookworm in the southern United States.

While much of the health economics literature focuses on this and related costs of illness, there is a growing literature concerned with the costs of avoidance behavior in both current and historical contexts. Meyers and Thomasson (2017) in particular examine the impact of the 1916 polio outbreak and its associated school closings (a common form of communal avoidance behavior) on educational attainment measured later-in-life for school-aged individuals who would have been able to opt into the labor force but were not likely to become infected with poliovirus. They find that higher levels of polio mortality (a proxy for the likelihood of education disruptions) in a state resulted in lower later-in-life education attainment; a one standard deviation in polio case counts per 10,000 resulted in 0.07 fewer years of schooling, on average.

Much of the research examining avoidance behavior in contemporary contexts examines such behavior at the macro-level. Avoidance behavior in response to infectious disease outbreaks impacts tourism and trade the most severely. The outbreak of Sudden Acute Respiratory Syndrome (SARS) in Asia in 2003 caused significant declines in GDP in China and reductions in tourism for all of the impacted countries (Hanna & Huang, 2004; Hai et al, 2004). Further, an outbreak of the plague in Surat, India in 1994 resulted in a loss of $2 billion in predicted trade (Cash & Narasimhan, 2000). Some of these macroeconomic impacts are due to the aggregation of individual decisions (for example, individuals deciding not to travel or purchase goods from infected countries) and some are the result of government's decision (Cash & Narasimhan (2000) note that Bangladesh cut off trade with India prior to the official confirmation of a single case of plague). I seek to add to the growing literature on micro-level avoidance behavior by examining the impact of polio on kindergarten enrollment.

Analyses

When I consider the individual impact of polio on kindergarten enrollment, I model the decision as a variable that takes the discrete form $K_i=0$ if a child who is kindergarten-age (5- or 6-years-old) has not attended kindergarten in the time period designated by the Census[4] and $K_i=1$ if a kindergarten-age child has attended kindergarten.[5] I model this decision using the following equation:

$$Prob(K_i=1)=F\left(\alpha_0+\alpha_1\delta_{st}+\boldsymbol{\beta}\omega_{st}+\boldsymbol{\rho}\mu_{st}+\boldsymbol{\lambda}_i'\phi+\sigma_s+\varphi_t+\varepsilon_i\right)$$

[4] Two months prior to February 1 in 1950 and 1960, one month prior to March 1 in 1940.

[5] I take this "attendance" variable to signify enrollment, following Bleakley (2007).

I include fixed effects for time, φ_t, and state, σ_s, to control for state-specific and year-specific effects. I include state fixed effects to control for time-invariant state characteristics (for example, climate, because polio is a disease with a high level of seasonality[6], and overall commitment to education). I include year fixed effects to control for place-invariant year characteristics: kindergarten enrollment rates rose between the beginning and end of the examined timespan, and polio outbreaks varied each year, both as new generations became susceptible and as new vaccines became available. These fixed effects will also control for any state-specific and year-specific heterogeneity in reporting polio case counts. Further, I cluster standard errors at the state-level to account for the likelihood individuals rising within a state may have correlated error terms across years.

The coefficient of interest is α_1, which reflects the impact of reported polio incidence rates in state *s* in year *t* on the probability a child will be enrolled in kindergarten. Further, I include personal and family characteristics in the vector $\boldsymbol{\lambda}_i'\phi$ in the second and third specifications. In the secondary specification, $\boldsymbol{\lambda}_i'\phi$ is a vector of a child's demographic characteristics: age, race, and gender, which are typical to include when examining the impact of exogenous shocks on school attendance (Goodman, 2014; Currie et al, 2009; Fagernäs, 2015). In the third specification, I also include covariates controlling for household head characteristics and family characteristics: the occupational score of the head, if the head is female, if the head is employed in agriculture, if both parents are foreign, if the mother is present in the household, if the father is present in the household, the number of siblings, and if the household is in a rural area. This last set of covariates is based on those included by Fagernäs (2015).[7]

Because head of household education data is largely unavailable for 1950, I include the median school years of residents 25 or older in a state *s* in year *t* as a proxy in vector μ_{st}. Sample statistics for demographic, head of household, and family characteristics, as well as average weekly polio incidence rates and state-year variables, are shown in Table 1 for each Census year. These head of household characteristics are included because female heads of households may have differing preferences for children's education, domestic and foreign parents may differ in education preferences, preferences may vary over occupation, since parents who work in agriculture are more likely to employ their children in labor, and preferences may vary over income. Here income is captured by head occupation score, a variable created by IPUMS to represent "occupational economic standing" (Ruggles et al, 2017).[8] In addition, I include family characteristics because kindergarten availability may be lower in rural areas and family income is likely to be correlated with education uptake.

---

[6] "Seasonality" refers to the fact polio outbreaks were much more common in months with high temperatures than months with low temperatures.

[7] Note that unlike Fagernäs (2015), I do not include a household head's literacy in these analyses because this data is unavailable in the Census years I examine.

[8] Occupational Income Score is the only income measure available for all three Census years I examine.

Table 1. Summary Statistics by Census Year.

	1940	1950	1960
Polio Incidence Per 100,000	0.11	0.63	0.17
Female	0.49	0.48	0.49
Age	5.49	5.50	5.49
White	0.87	0.89	0.87
Black	0.13	0.11	0.12
Other Race	0.01	0.00	0.01
Occupational Score of Head	21.37	23.82	23.79
Female Head	0.06	0.06	0.07
Head Employed in Agriculture	0.25	0.16	0.15
Both Parents Foreign	0.09	0.05	0.06
Father Present in Household	0.91	0.90	0.91
Mother Present in Household	0.96	0.96	0.96
Number of Siblings	2.73	2.14	2.49
Rural	0.56	0.47	0.43
Median Years of Schooling (State-Year)	8.08	9.44	9.37
Birth Rate (State-Year)	18.43	23.82	21.09
Percent of Homes with Running Water (State-Year)	64.78	82.78	80.16

I identify the impact of polio on kindergarten enrollment by relying on variation in reported state-level polio incidence rates. Because of this, I include epidemiological factors that influence polio outbreaks in all analyses (the percent of the population with running water and the crude birth rate in each state-year) with vector ω_{st}. I account for the previously mentioned "hygiene hypothesis" (that polio epidemics were caused by increases in sanitation) by including percent of dwellings in a state-year with access to running water as a proxy for varying sanitation levels. This is an especially important control if both sanitation and public school kindergarten availability are impacted by a state's overall commitment to public good provision. If this is the case, the coefficient of interest would be biased upward, with some states having both a higher likelihood of kindergarten enrollment and more polio outbreaks.

A second hypothesis about polio's rise is that shifting demography drove increases in polio. Martinez-Bakker and Rohani (2015) identify the baby boom as a driver of polio epidemics: rising birth rates created more asymptomatic polio carriers in the form of six-month-old and younger infants who then spread polio to other portions of the population. To control for this possibility, I include birth rates by state of residence for the year *t* in all analyses. I also include birth rates for *t-5* and *t-6* because these lagged birth rates are likely correlated with both birth rates in year *t* and kindergarten enrollment for five- and six-year-olds in *t*. When I examine if impacts are heterogeneous by age, I include lagged birth rates appropriate for each age group (including the birth rate for *t-7* when examining seven-year-olds, for example).

I run several individual-level analyses to examine the impacts of reported polio incidence rates on the likelihood a five- or six-year-old will be enrolled in school.

These analyses include (1) an ordinary least squares regression including all kindergarten-aged children, (2) a base logistic regression including all kindergarten-aged children, (3) a probabilistic regression with all kindergarten-aged children, (4) an OLS regression comparing kindergarten-aged children with older but proximate age groups, and (5) an OLS regression comparing families with a stay-at-home mother and families with two employed parents. I also present a falsification test and robustness checks where (1) missing state-weeks of polio data are included, (2) the effect of one-year polio leads on kindergarten enrollment is the coefficient of interest, and (3) a less flexible model is used with fixed effects excluded and replaced with state-year characteristics.

Data

To explore the impacts of the uncertainty surrounding polio, I obtained reported state-week disease incidence rates from Project Tycho (van Panhuis et al., 2013). Incidence rates are a measure of new disease case counts during a specified timeframe and in the Tycho data take the form:

$$\textit{Incidence Rate per 100,000 in Week W} = \frac{\textit{Number of New Polio Cases in Week W}}{\textit{Estimated Population in 100,000 in Week W}}$$

Using weekly disease incidence rates data, I create an average weekly incidence rate for a given twelve month period. I use the average weekly incidence rate instead of total incidence rate for the given time period because average weekly incidence rate will be less impacted by missing observations within the Tycho data than an aggregated incidence rate. However, average weekly incidence rates are an imperfect measure of either the stock or the flow of population-adjusted polio cases, and instead attempt to capture the average rate of appearance of new cases in the reference period. A parent's decision to enroll their child in school in the wake of a polio outbreak is likely driven by how prevalent the disease appears to be, and this formed belief about prevalence might intensify (or lessen) over time if average disease incidence rates rose (or fell). Average weekly incidence rates then are appropriate to use because they give a sense of how prevalent the disease was over a period of time and on average how many new cases a parent was made aware of throughout the period. I also include base results using two different measures of polio incidence (estimated total yearly incidence rates in state *s*, the maximum weekly incidence rate in the past twelve months in state *s*, and the minimum weekly incidence rate in the past twelve months in state *s*) in Table 26 in Appendix A.

The reference point for the individual-level data is the Census reference date for the individual enrollment question: February 1 for 1950 and 1960 and March 1 for 1940. Because of this, the average polio incidence rate for the analyses is not for the calendar year but the average incidence rate from the February of year *t-1* through the January of year *t* for 1950 and 1960, and the average incidence rate from the March of year *t-1* to the February of year *t* for 1940. The historic polio data that Project Tycho has digitized contains missing observations; I only include observations that have at least three-quarters of a year's data (ie, data that is missing 13 state-week observations or fewer) on incidence rates in order to have a more accurate measure of the average incidence rate of polio. In Table 27 in Appendix A, I present results with varying thresholds for the number of missing state-weeks. There is an apparent tradeoff between observations gained and the model's ability to fit the data (indicated by the R-Squared) when observations in state-years with less data are included in the analysis. Nevertheless, results broadly agree with the results that follow.

I obtained individual-level micro-data for the Census from the Integrated Public Use Microdata Series (IPUMS-USA) website (Ruggles et al., 2017). Heads-of-household are coded as being employed in agriculture if they are "farm, ranch, and other agricultural managers," farmers or ranchers, farm laborers (including farm foremen, farm wage workers, unpaid family workers, or self-service farm laborers), or farm owners or tenants. I code respondents as living in a rural area if they live outside a metro area. I inflation adjust all dollar amounts used to 1960 dollars.

I obtained the percent of reporting homes without running water for the decennial Housing Census (U.S. Bureau of the Census, 1943; U.S. Bureau of the Census, 1953; U.S. Bureau of the Census, 1963). Crude birth rates were taken from the Center for Disease Control's Vital Statistics of the United States (U.S. Bureau of the Census, 1946, 1962b). Crude birth rate by state of residence is used for 1935 and later years; crude birth rate by state of occurrence is used for years prior to 1935 because crude birth rate by state of occurrence is unavailable for these years. Median years of schooling were obtained from Census records made available by the Inter-university Consortium for Political and Social Research (Haines et al, 2010; U.S. Bureau of the Census, 1942; U.S. Bureau of the Census, 1952; U.S. Bureau of the Census, 1962a). In the 1940 Census median years of schooling were calculated separately for men and women; I use the median years of schooling for men in the analyses, but the two are closely related (ρ=0.92).

Results

Main Results

I focus on the question of how an increase in the incidence rate of polio impacted the likelihood a kindergarten-aged child would be enrolled in kindergarten. I report results for all logistic and probabilistic regressions in terms of average marginal effects. I report results for the OLS regression in Table 2, and I report results for the logistic regression in Table 3, and for the probit in Appendix A, Table 28. For all regressions, I present three analyses: an analysis with controls for polio case counts per capita, birth rates, running water controls, and state and year fixed effects; an analysis that adds demographic controls (age, gender, race); and an analysis that further controls for household and family characteristics and median schooling of adults in a state-year.

In the base analysis, shown in column (1), I observe a new case of in polio in the average week per 100,000 people results in a 3 to 4 percentage point reduction in the likelihood a child will attend kindergarten but this result borders on insignificance, especially once a rich panel of covariates is added. The size of this coefficient is comparable to the magnitude of having a female head of household in absolute terms, or approximately half of the size of the impact of living in a rural area where there might be lower kindergarten availability overall.

The main analysis with all coefficients, outside of state and year fixed effects are reported in Appendix A, Table 29. I report OLS as the main analysis because the logistic regression, while it may be better suited to binary outcomes, may suffer from the incidental parameters problem because of the large number of fixed effects in the model.

Logistic regressions are what has been traditionally used within the school enrollment literature. I therefore present these results as well in Table 3. These results imply that a 1.0 increase in the average polio incidence lowers the likelihood a child would attend kindergarten by 4.6 percentage points in column (3). To compare the magnitude of this result to that of other variables, the

Table 2. Pooled OLS Regression of the Impact of Polio Incidence Rate on the Likelihood of Kindergarten Attendance for Five- and Six-Year-Olds.

	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.041	-0.041*	-0.035
	(0.03)	(0.02)	(0.02)
R^2	0.14	0.34	0.35
N	95,839	95,839	95,839
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 3. Pooled Logistic Regression of the Impact of Polio Incidence Rate on the Likelihood of Kindergarten Attendance for Five- and Six-Year-Olds.

	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.062**	-0.051**	-0.046**
	(0.03)	(0.02)	(0.02)
Log Pseudolikelihood	-5,435,139.15	-4,335,417.08	-4,267,705.38
N	95,839	95,839	95,839
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

average marginal effect of a child living in a rural area is a 7.5 percentage point reduction in the likelihood that child would attend kindergarten, all else equal, and the average marginal effect of a child with a head of household employed in agriculture results in a 4.0 percentage point reduction. Thus, the impact of a polio outbreak during the year can be a large impact, comparable to a reduction in the availability of kindergarten or the willingness of a parent to lose an additional farmhand.

Heterogeneous Effects by Age

If the non-mandatory nature of kindergarten is what is driving the results because parents have greater ability to change their behaviors for this particular age group, then other proximately aged children whose schooling is mandatory should show a different pattern of results. Table 4 reports results when looking at different age bands close to five- and six-year-olds: seven- and eight-year-olds and nine- and ten-year-olds. For the other age bands, I find consistently insignificant results. This could indicate polio was a disease of particular concern to parents of young children, and parents were particularly responsive when schooling was not required.

I take these results to point to the idea that parents with additional flexibility because a child was younger (and therefore less likely to be subject to a state's compulsory education law) were more likely to keep the child out of school. Another margin that might indicate additional flexibility to keep a child home from school in the wake of a polio outbreak is the presence of a stay-at-home mother in the family, which I examine in the next section.

Heterogeneous Effects by Maternal Employment

In addition to the age of a child granting parents additional discretion over school enrollment decisions, parental employment may also impact the decision to enroll or not enroll a child in school. To examine this possibility, I analyze the enrollment patterns of two different types of family employment structures: families with a stay-at-home mother who has not had another occupation in at least the last year and families with dual incomes. Focusing on these long-term stay-at-home mothers removes the possibility a working mother quit her job to stay at home with her children because of a polio outbreak. Instead, the implied channel is that some families have more flexibility about school enrollment decisions than others because of the previous presence of a homemaker.

Results shown in Table 5 imply that the impact of polio incidence varies by maternal employment: coefficients for families with long-term homemakers consistently have larger impacts and greater precision than families with two employed parents; the families with an available homemaker seem to be the families driving the analysis. Results imply that one additional case of polio in the average week out of 100,000 people decreased the likelihood a child would be enrolled in kindergarten by more than 4.5 percentage points *if* that child was in a family with a stay-at-home mother. An additional way to view these results is that

Table 4. OLS Regression of the Impact of Polio Incidence Rates on the Likelihood of School Enrollment for Seven- and Eight-Year-Olds, and Nine- and Ten-Year-Olds.

7 & 8 Year Olds			
	(1)	(2)	(3)
Polio Incidence Per 100,000	0.008	0.008	0.004
	(0.02)	(0.02)	(0.02)
R^2	0.06	0.06	0.06
N	92,839	92,839	92,839
9 & 10 Year Olds			
	(4)	(5)	(6)
Polio Incidence Per 100,000	-0.013	-0.012	-0.014
	(0.02)	(0.02)	(0.02)
R^2	0.05	0.05	0.05
N	90,217	90,217	90,217
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5. Pooled OLS Regression of the Impact of Polio Incidence Rates on the Likelihood of Kindergarten Attendance for Five- and Six-Year-Olds, with Separate Analysis for Families with Long-Term Homemakers, and Families with Two Working Parents.

Mother Not in the Labor Force			
	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.060**	-0.053**	-0.046**
	(0.02)	(0.02)	(0.02)
R^2	0.13	0.36	0.37
N	61,518	61,518	61,518
Both Parents in Labor Force			
	(4)	(5)	(6)
Polio Incidence Per 100,000	-0.014	-0.038	-0.027
	(0.04)	(0.04)	(0.04)
R^2	0.14	0.32	0.33
N	10,086	10,086	10,086
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

if the polio incidence rate increased by one standard deviation (.24 for the sample), the likelihood a child with a stay-at-home mother would be enrolled in kindergarten would drop by 1.08 percentage points. These results bolster the argument that that the impact of polio incidence is driven by parental discretion when faced with a health threat: parents with additional flexibility to keep a child at home, either because of a child's age or because there is a person available to take care of the child, are more likely to display behavioral changes in response to polio outbreaks.

Falsification Tests & Robustness Checks

To ensure that results are caused by exogenous shocks to public health uncertainty driven by polio outbreaks and not other mechanisms, I present results in Table 6 where I use one-year leading polio incidence rates instead of current year polio incidence rates. The results are insignificant with smaller coefficients than the main results. I take these findings to indicate that polio case counts are exogenous with respect to kindergarten enrollment once state-level covariates are included in the analyses. Presented in Table 7 are analyses that control for the number of missing weeks in each state-year. A possible relationship between state data quality and kindergarten availability could exist if states with better infrastructure both provide more kindergarten and more reliably report polio cases. Results do not indicate data quality is what drives the results: the number of missing weeks are not significant predictors of kindergarten enrollment, and controlling for the number of missing weeks results in similar coefficients as the main analyses.

In addition to these robustness checks, within Appendix A I present analyses without epidemiological controls (Table 30) and without state fixed effects (Table 31), respectively. Results are more significant when epidemiological controls are excluded. and are insignificant and positive when state fixed effects are excluded.

Conclusion

These analyses indicate polio could have had large impacts on families with the greatest ability to change their behavior – families with children whose schooling was optional and with a person already at home to stay with a child. If this is the case, what explains the increased uncertainty and fear surrounding poliomyelitis?

First, polio's mode of transmission is the oral-fecal route. This differentiates it from other childhood diseases like measles, mumps, or pertussis, which are spread through cough. These diseases' modes of transmission are also their major symptom. Because there is a difference between polio's mode of transmission and its symptoms, parents cannot readily observe who is infected (or contagious) in a child's class or community or even the infection status of

Table 6. Falsification Test: Pooled OLS Regression of Impact of One Year Leading Polio Incidence Rates on Kindergarten Enrollment.

	(1)	(2)	(3)
One Year Leading Polio Incidence Per 100,000	-0.041	-0.030	-0.027
	(0.047)	(0.048)	(0.043)
R^2	0.14	0.34	0.35
N	95,369	95,369	95,369
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 7. Robustness Check: Pooled OLS Regression of Impact of Polio Incidence Rates on Kindergarten Enrollment with Number of Missing Weeks of Polio Observations.

	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.041	-0.042*	-0.035
	(0.025)	(0.023)	(0.021)
Missing Weeks Included	0.000	0.001	0.000
	(0.002)	(0.002)	(0.002)
R^2	0.14	0.34	0.35
N	95,839	95,839	95,839
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

their own child. A parental reaction to keep a child out of school makes sense when thinking of ways individuals choose to delineate boundaries even when such boundaries may not be statistically useful:

> [A disease's] crossing of boundaries is essential to the creation of panic. When the edge of safety cannot be defined, people react in ways that are not necessarily rational—cordoning off suspect populations; creating artificial boundaries that create the illusion of safety; fleeing somewhere, anywhere. (Humphreys, 2002)

Second, as discussed in the above section, polio is a relatively recent disease. This newness may have made polio more threatening: mid-century parents witnessed polio outbreaks growing more frequent and on a larger scale throughout their lifetimes until the introduction of the polio vaccines. Third, polio may have been especially frightening because of the associated paralysis and the fact that polio primarily attacks children, possibly handicapping or scarring them for life (Dredge, 2008). Perhaps as a result, polio became the focus of a massive media campaign for eradication in the 1900's:

> ... in 1938 [Franklin Delano] Roosevelt helped found the National Foundation for Infantile Paralysis (known later as the March of Dimes) that raised millions of dollars for the rehabilitation of those who suffered from paralytic polio and later invested heavily in funding the research that led to effective polio vaccines. (Dredge, 2008, p. 552)

All of these reasons indicate that polio was much more in the public eye and imagination than other childhood diseases. It is possible that the culmination of this publicity and associated notoriety resulted in pronounced avoidance behavior, especially among families where such avoidance behavior was less costly. These results highlight the real impacts that uncertainty surrounding public health can create, and that some populations will be most likely to exhibit this behavior: populations both directly affected by the disease and populations with flexibility to enact such behavior with lower costs.

CHAPTER 2
DEBT PLUS: COULD STUDENT LOANS IMPACT PARENTS' RETIREMENT?

Introduction

Student loans have come to the forefront of national attention in recent years, and a burgeoning economic literature investigates loans' impacts on borrowers' outcomes. This emergence of literature and attention is not surprising: total student loan debt (including Perkins, Stafford, PLUS,[9] and nonfederal loans) has remained above $100 billion in constant 2013 dollars since the 2007-2008 academic year. Inflation-adjusted debt per student rose by 35% from 2004 to 2013, and the number of borrowers increased 86% during the same timeframe (Trends in Student Aid, 2014). These statistics take into account all student loans taken out by both parents of students and students themselves, but the student loan literature thus far has focused mainly on federal Stafford loans or the means-tested Perkins loans and their effect on student borrowers. Little research, economic or otherwise, has been dedicated to how student loan uptake impacts parents. Parents may be directly impacted by taking out student loans to fund their child's education, or they may be indirectly impacted by supporting their children through the loan repayment process. Student loans could then possibly create intergenerational transfers of wealth that may be unplanned and may be particularly risky as a parent approaches retirement age. This possibility opens up economic questions: does student loan uptake delay parents' retirement or decrease their retirement savings? We examine this question by looking at student loan presence within a household (which includes both debt taken out by parents for students and debt taken out by any students that reside within the household) on several measures of retirement behavior, expectations, and savings.

PLUS loans, the most common loan parents take out, made up 12% of all federal and nonfederal loan dollars in the 2016-2017 school year (Trends in Student Aid, 2017). During this time frame the average PLUS loan recipient borrowed 2.4 times more than the average federal Stafford loan recipient (Trends in Student Aid, 2017). Figure 2 illustrates the percent of parents borrowing PLUS loans and the average dollar amount borrowed from a sample of college students' families interviewed by Sallie Mae. Despite the fact that PLUS loan use is relatively common, little information is publicly available about the repayment, default, and deference rates of PLUS loans, unlike federal loans disbursed to students (Fishman, 2018).

In this paper, we examine the effect of student loan presence in the household on parents' employment status, retirement expectations, and retirement preparation. Using a rich data set that links parents to children and

[9] Stafford and Perkins loans are federal student loans made to students; PLUS loans are federal student loans made to parents or graduate students.

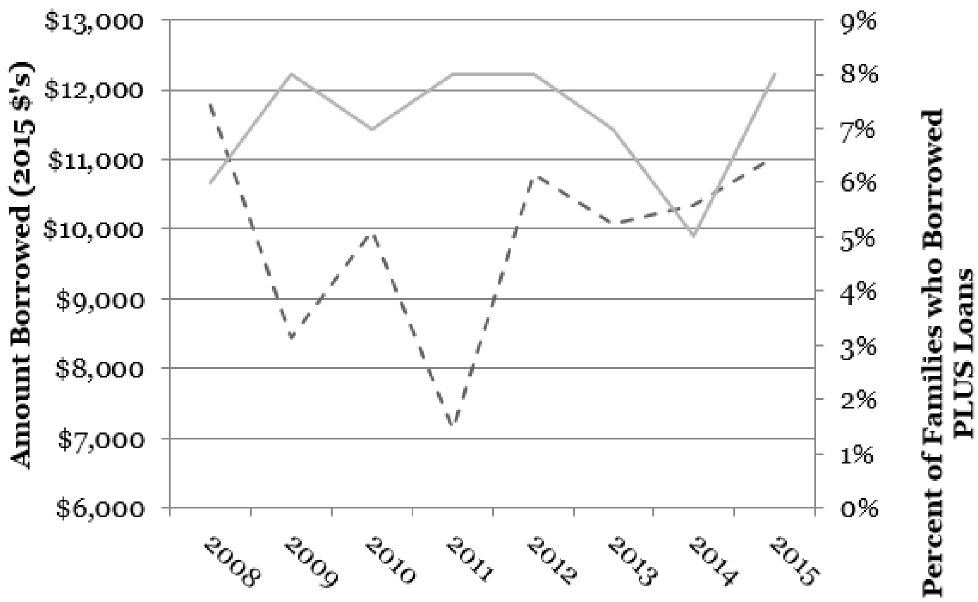


Figure 2. Percent of Families Borrowing PLUS Loans and the Amount of PLUS Loans Borrowed Interviewed in Sallie Mae's "How America Pays for College."[10]

those children to their undergraduate institutions, we measure the effect of student loans on parental labor force participation, employment status, expected retirement age, the gap between the age a parent expects to retire and the age she believes others at her same job retire, and dollars in a family's individual retirement accounts (IRAs). The main concern with a simple regression model to calculate the effect of uptake on these variables is that parents in families where student loans are present may differ from parents in families where student loans are not present in a systematic way that affects outcomes of interest. We account for the possibility of omitted variable bias in three ways: (1) we include a robust set of control variables about the child's institution and control variables correlated with retirement behaviors, (2) we examine if taking out student loans in the present period is correlated with money problems in a past period that would not affect student loan eligibility, and (3) we compare our unconditional results with results that include this robust set of controls and formally evaluate the potential threat from unobservables using a technique developed by Oster (2017).

The question of the relationship between student loan debt and retirement is timely because not only has student loan debt in general come to the forefront of national attention, but there is widespread concern that Baby Boomers retiring will affect the economy at large (Casselman, 2014). Baby Boomers are also more indebted than previous generations, and that indebtedness has led to less financial security (Lusardi, Mitchell, & Oggero, 2017). This paper links these

[10] (Sallie Mae, multiple years)

topics in an attempt to shed light on the ramifications of student loan presence in a household on parents.

The most striking results we find concern dollars in a household's Individual Retirement Accounts. The results indicate a significant reduction in IRA dollars for borrowers: a household with *any* student loan debt on average has $70,000 fewer dollars in IRAs than a household without student loans. Moreover, results imply an additional $10,000 in student loan debt reduces IRA savings by $14,000. Neither the presence nor amount of student loans affect the likelihood a family will have an IRA. Results also indicate an additional $10,000 of student loan debt increases a parent's expected retirement age by 5 months. The presence of any student loan debt increases the likelihood a parent will be employed by 5.5 percentage points and increases the likelihood a parent will be in the labor force by 5.4 percentage points, but the employment status coefficients are only marginally significant. Robustness checks and a formal check for robustness developed by Emily Oster (2017) suggest that these results could be driven by unobservable differences between families who borrow and families that do not.

Background & Previous Literature

PLUS Background

Although we focus on the effect of all student debt within a household, PLUS loan debt, the type of debt taken out directly by parents, and its mechanics may be less familiar. To this end, this section briefly explains this particular type of debt. PLUS loans are available to both parents and graduate students, but we focus on the implications of PLUS loans for parents.

Like other federal student loans, PLUS loans are not collateralized debt nor does an application require an income check. However, PLUS loans differ from other federal loans in two main ways: (1) a PLUS loan requires a separate application that includes a credit check and (2) the amount a PLUS loan recipient can borrow is determined by an institution's cost of attendance rather than a dollar amount set by the federal government.

The credit check required for PLUS loan receipt is only a check for adverse credit history, not a check to examine repayment ability. Moreover, the credit check simply approves or denies a parent for access to the loan rather than for a specific loan amount.

The loan amounts offered to approved PLUS recipients can be extremely high because the loan covers the cost of attending a university less any other aid. Cost of attendance is determined by a university itself, and can include indirect expenses (such as travel) as well as direct expenses (such as tuition, room, and board). PLUS loans thus offer parents an opportunity to fill the gap between the aid their child has received (including any loans taken out by the child) and the cost of attending a university. However, this opportunity also means that an education may make families indebted over multiple generations, with a student taking out the maximum amount of loans they can and then parents borrowing the rest (Fishman, 2018). Dependent students could only borrow $31,000 in Stafford loans *in their total college career* in 2016; no aggregate limit existed for PLUS loans. Further, the interest rates on PLUS loans are typically much higher than Staffords: for the 2016-2017 school year, the interest rate on PLUS loans was 7.00% and the interest rate on Stafford loans was 4.45%. Thus, PLUS loans allow parents to borrow large levels of debt that compound faster than the loans offered to students, with no attempt to estimate a parent's repayment ability.

Previous Literature

Although student loans are frequently portrayed in the media as an unmitigated burden, this debt conforms to economists' lifecycle consumption hypothesis. The lifecycle consumption hypothesis predicts individuals will acquire debt early in life, pay the debt back by midlife, and begin saving for retirement. This consumption-smoothing benefit of Stafford and Perkins student loans is believed to outweigh their costs for most individuals (Webber, 2016). Regardless, numerous studies have documented larger behavioral change than theoretically predicted in the presence of student loan debt (Field, 2009; Minicozzi, 2004; Rothstein & Rouse, 2010; Cooper & Wang, 2014).

On the other hand, parents do not anticipate any consumption-smoothing properties due to student loan uptake. Anecdotally, it seems common for students to promise to pay back the loans themselves even though the debt is in their parent's name (Rhode, 2015). At most, PLUS loans are then a second best solution to Staffords or a supplement to Staffords.

Parental contributions (including PLUS loans and other forms of borrowing, as well as funding education through income or assets) make up the largest portion of student aid (Sallie Mae, 2015). There is a small amount of literature on the interaction of parental aid and student aid or outcomes, but as far as the authors are aware there is no literature on the effect of parental aid or student loans on *parents*. Scott-Clayton and Zafar (2016) find that West Virginia's PROMISE merit aid scholarship decreased the amount of loans taken out by parents by less than $1,000, but this result is statistically significant only for some subpopulations (female students, non-Pell recipients, and public high school students) and not significant for the total population. Stolper (2014) found that parents with access to increased credit, specifically a Homeowner Equity Line of Credit, are more likely to send their child to a more selective school and push renters' children out of these more selective schools. Further, children whose parents are providing some source of their aid are more likely to stay in school but also have lower GPAs than students whose parents are not contributing financially (Hamilton, 2013).

Why would a parent borrowing for education produce different impacts on retirement than parents funding education through their income or savings? We posit the difference is planning. For large portions of borrowing families, borrowing was unanticipated: 30% of borrowing families Sallie Mae (2015) interviewed had not planned to borrow for college. This much unexpected debt late in life may distort parents' labor market choices, especially when many older Americans are uninformed about their own retirement preparedness.[11]

A small literature on later-in-life debt confirms this intuition. Anguelov and Tamborini (2009) emphasize that deviating from the lifecycle consumption path (either from an individual miscalculating their own trajectory or unplanned life events) through accruing late-in-life debt may impact individuals in a variety of ways:

> Servicing high levels of debt while working may hinder a family's ability to save for retirement. … Debt service obligations could lead individuals to work longer. Debt may also reduce the longevity of a household's accumulated financial assets and savings, which would have to be spent down to repay debt when income is limited. Indebtedness, especially from high-interest consumer borrowing, could also leave elderly persons with fewer retirement resources in the face of health and other income shocks. (p.14)

A working paper from the Center of Retirement Research at Boston College finds support for this view. Butrica and Karamcheva (2013) examine the effect of indebtedness and liquidity constraints on nondisabled older adults' (those who are aged 62 to 69) joint decisions of retirement and Social Security uptake through a bivariate probit model. Butrica and Karamcheva look at overall debt, mortgage debt, and credit card debt. They find all three debt measures increase the likelihood an individual will be working and decrease the likelihood of Social Security benefit uptake. Their results imply that having any debt makes an individual 8 percentage points more likely to work, having mortgage debt makes an individual 7 percentage points more likely to work, and having any credit card debt makes an individual 4 percentage points more likely to work. Additionally, they use instruments for mortgage debt to try and determine the effect of debt on retirement and benefit uptake with better identification. Even when using instruments, they find mortgage debt to be a significant predictor of employment.

Mann (2011) estimates the impact of debt on retirement decisions using a multivariate model with individual-level fixed effects. The results confirm intuition: individuals with higher debt levels delay retirement longer. She finds this impact is lessened for those with higher wealth levels.

Economic Model & Methodology

To examine the specific effect of student loan debt, we adapt a simple economic model employed by Butrica and Karamcheva (2013). The model relates liquidity constraints to retirement decisions and includes controls for demographic and economic characteristics. Butrica and Karamcheva use a bivariate probit to model the effect of household debt on labor force participation and Social Security

[11] Munnel et al (2017) estimate that 52% of U.S. households in the *2013 Survey of Consumer Finances* are underprepared financially for retirement, and 19% of U.S. households believe they are prepared for retirement and are not.

uptake as a joint decision.[12] We instead solely examine the effect of student loan debt on employment status, retirement savings, and retirement expectations regressands separately. The model for individual i with children who attend colleges with aggregate characteristics u is thus

$$Y_i = X_i\beta + Lc_i\gamma + K_u\rho$$

Here, Y_i is a parental outcome of interest: employment status (employed, unemployed, not in the labor force, or retired), expected retirement age, distance from parent's own expected retirement age and the age they believe others with their same job retire at, presence of IRA/annuity savings within the household, or dollars in an IRA/annuity for the entire household. For the employment status variables and presence of an IRA account, Y_i is the realized outcome of a latent variable Y_i^* where $Y_i = 1[Y_i^* > 0]$. For the other continuous variables, Y_i itself is the outcome of interest. Following Butrica and Karamcheva (2013), X_i is a vector of economic and demographic data about person i. Lc_i contains information about a household's liquidity constraints, specifically a household's student loan debt. We include aggregate measures of a parent's children's college exposure rather than individual parent-child pairs because we do not want the analysis to be driven by parents with multiple children when the effect we investigate is on the parent rather than the child. We include number of children and total number of years a child has gone to school, however, to account for the fact that some parents are exposed to the choice about borrowing for a child's education multiple times and to account for the number of years a parent could be borrowing to fund.

We include these institutional characteristics because the literature on student loan debt has also found that individuals who attend some school types are more likely to have higher levels of debt. Private and for-profit schools are associated with higher debt levels for students than public schools, and four year colleges are associated with higher debt levels than two year colleges (Chen & Wiederspan, 2015; Cellini & Darolia, 2016). Schools in the private and for-profit sectors also tend to have higher average PLUS loan disbursements and a higher percentage of students receiving parental PLUS loans (Dancy, 2016). In addition, students at historical black colleges and universities (HBCUs) have been found to be take out more loans than students at non-HBCU schools (Hackett, 2016). For these reasons, we include K_u, a vector of data about the colleges that parent i's children attended: number of children that attended college a parent was matched with, the latest child's first entrance year, the total number of years of college a parent's children have attended, and categorical variables for the percent of a parent's children ("all," "some," or "none") that began college at a private school, public school, for-profit school, 4-year school, and historical black college.[13] We include the latest child's first entrance year because it is likely that

[12] Only 29% of our sample is eligible for social security (age ≥ 62) compared to all of Butrica & Karemcheva's sample.

[13] We use categorical variables for these percentages rather than the percentages themselves because data from the restricted data set we use is only allowed to be reported if cells have more than 11 observations. The categorical variables meet this standard; the percentages themselves have too few observations per cell.

when a child first attends college is when a parent makes the decision to fund the child's education with loans or not.

We perform a simple OLS analysis using cross sections of the latest data available for a parent from the Panel Study of Income Dyanmics. The key threat to identification in this case is omitted variable bias: parents in households with student loans could differ systematically from parents in households without. In order for the coefficient of interest to be unbiased and consistent, we require the systematic part of the outcome variable Y_i to be uncorrelated with the unsystematic part of the outcome variable. In practice, this means that if the underlying process generating the outcome Y_i is

$$Y_i = X_i\beta + Lc_i\gamma + + K_u\rho + F_i\vartheta$$

where there is some omitted variable F_i for which $E(Lc_i|F_i) \neq 0$ then the zero conditional mean assumption has been violated. If this happens, the coefficients produced by OLS are inconsistent and biased. In other words, the results produced may not capture the causal effect of loan presence on parents' decisions but spurious correlation between loan presence and retirement behaviors.

This possibility requires serious consideration about how to address it. First, we include control variables for which we believe the conditions above will hold. These controls include those considered by Butrica and Karamcheva (2013): parental demographic characteristics (sex, age, if the respondent has reached the age for Social Security uptake, and race and ethnicity), if a spouse is present, reported health status, wealth and income information (the family income excluding individual i's labor income, and household assets) and state of residence. We differ from Butrica and Karamcheva in three ways: (1) we use the age of Social Security eligibility (62) rather than the Full Retirement Age for Social Security (65) because our sample skews much younger, (2) we do not include data on spousal income or retirement status because these fields are collinear with spouse presence within the PSID data, and (3) we use state rather than Census region because this more granular location data is available in the PSID.

We include additional controls in the robustness checks: a measure of financial literacy and the education of the parent's mother in one robustness check, and a measure of a household's earliest wealth in another.[14] We assume that a parent's maternal education is a proxy for a parent's taste for education and other unobservables that may drive labor market behaviors. A parent with a higher taste for education may be more likely to fund their child's education; further, parental education has also been shown to have a significant effect on labor market outcomes, perhaps because it represents unobservable traits with value in the labor market (Agnarsson & Carlin, 2002; Hudson & Sessions, 2011). Financial literacy has been shown to be positively related to retirement planning behaviors and wealth accumulation (Lusardi & Mitchell, 2007). Financial

[14] These controls reduce the sample size further, so we include them as robustness checks and also only include the employment status regressions; for the other regressions, sample size falls below 400.

literacy may be correlated with PLUS loans in particular in two ways: (1) parents with greater dedication to their child's education, i.e. those who borrow education loans, may prioritize their own education (including financial literacy), and (2) financial literacy may matter especially for financial instruments parents are unfamiliar with and instruments for which there is no income requirement and a comparatively weak credit check, such as PLUS loans. If (1) holds, then the coefficient of interest (the effect of student loan presence on retirement behavior) would be biased downward. If (2) holds, the coefficient of interest would be biased upward. In addition to these controls, we make use of a technique developed by Emily Oster (2017).

Within the economics literature, coefficient stability is frequently taken as a signal that a coefficient is robust and has approached its "true" unbiased value. Oster (2017) points out that only considering coefficient stability and not R^2 movements neglects to consider the possibility that the additional control a coefficient is "robust" to may not be instrumental in explaining variation in the model at all – and thus the coefficient of interest's stability is not indicative of its unbiasedness but of the small effect the additional control has on the model.

To propose formal tests of robustness, Oster (2017) relies on the simple setup

$$Y = \beta X + \Psi \omega^{o} + W_2 + \varepsilon$$

Here, ω^{o} defines a vector of controls, W_2 defines a vector of unobservables, and X is the coefficient of interest ("treatment"). She also denotes $\Psi\omega^{o} = W_1$. Assume W_2 and W_1 are orthogonal. For i = 1, 2, the variance of W_i and the cov(X, W_i) are denoted σ_i^2 and σ_{iX}, respectively. Oster proposes a "coefficient of proportionality," δ, such that

$$\delta \frac{\sigma_{1X}}{\sigma_1^2} = \frac{\sigma_{2X}}{\sigma_2^2}.$$

Oster shows that δ can be interpreted as the degree of selection based on unobservables necessary within an estimation for the effect of treatment (here, the effect of student loan debt) to be explained away. The δ parameter estimated takes into account not only coefficient stability but also R^2 movements. This δ provides an intuitive way to imagine selection bias if we assume a maximum[15] R^2 and β: δ = 3 implies that the selection process would need to rely three times more on unobservable variables rather than observable variables for the treatment effect, β, to equal zero. In addition, it is possible for δ to be negative, but in this case implications about robustness are less clear. In these results, δ < 0 is associated with either (1) the coefficient of interest from the conditional results is in fact further from zero than the coefficient of interest from the naïve results (i.e., those that do not include control variables) or (2) the coefficient of interest is insignificant. The literature surrounding the δ coefficient is still being formed; we thus do not attempt to interpret any δ < 0 as indicating the presence or absence of selection bias.

[15] Oster points out the possible R^2 for a given estimation is unlikely to be 1 "since in many settings there is likely to be some idiosyncratic variation in outcome—and more importantly, the degree of this is likely to vary…" (p. 5).

From this, Oster derives a bias-adjusted β for which δ is assumed to be a specific value (i.e., selection on unobservables and observables is assumed to be a specific ratio). In the following analyses, we follow Oster's suggestions and when calculating δ, we assume $\beta = 0$, i.e. we estimate the degree of selection based on unobservables necessary for the effect of student loan debt on parents to be explained away entirely. When calculating the bias-adjusted β, we assume $\delta=1$ (unobservables drive selection proportional to observables). For both calculations, we must assume a maximum R^2 (the amount of variation within the data that can be explained). We follow Oster's suggestion and assume R^2_{max} is 1.3 times the conditional R^2. Oster suggests two formal standards for determining the robustness of a coefficient of interest when the inclusion of controls moves β further from 0, as happens in our analyses. These formal robustness tests are (1) if δ is greater than 1, and (2) if the bias-adjusted β is within 2.8 standard errors of the controlled β.[16]

Data

Panel Study of Income Dynamics

To perform the analyses, we use individual- and household-level data from the Panel Study of Income Dynamics (PSID). The PSID is a longitudinal study of families in the United States. The PSID began collecting survey data about these families in 1968; originally 5,000 families were surveyed. The PSID surveyed more than 9,000 families in 2015, a result of new populations being added to the PSID (for example, a sample of 2,000 Latino households was added in 1990) and also the formation of new families when children create their own households (Panel Study of Income Dynamics, 2017). Surveys were given annually from 1968 to 1997 and once every two years from 1999 onward.

The PSID survey given to a family always consists of at least two portions: the Main Family survey and the Individual survey. The PSID collects data about members of a Family Unit who reside in the interviewed household. Each member in the Family Unit is identified in relation to the household's Head. The household Head is identified as the main income earner in the household, unless that main income earner is a female with a legal or cohabiting male partner, in which case that male partner is considered the Head and the female partner is considered the legal or cohabiting Spouse.[17] These Family Units can and do change over the course of the survey as people split off from the family and form their own households.

Coverage (i.e., which family members questions are asked about) differs by question, rather than by family, for both the Main Family and Individual survey. Within the Main Family survey, data is collected for the entire Family Unit (for example, the amount of IRA/annuity assets in total for all members of the Family

[16] "+/– 2.8 standard errors is the bounds of the 99.5% confidence interval…" (Oster, 2017).
[17] See the Panel Study of Income Dynamics' Frequently Asked Questions (2018) for more information about the selection process.

Unit) or only for the household Head and the Head's Spouse (for example, the labor income for the Head two years ago, and the labor income for the Spouse two years ago). The Individual survey collects information about all members of the Family Unit (for example, an individual's relationship to the Head) or only about the individual answering the survey questions as well as the designated Head and Spouse (for example, an individual's reported health status).

In addition to these main questionnaires, families may also be given supplemental questionnaires meant to explore specific topics. Of particular relevance to this paper is the Transition to Adulthood (TA) supplement, a rotating survey that began in 2005. Prior to 2017, the survey was given each year to young adults (individuals between the ages of 18 and 28) who had previously responded to the Child Development Supplement (CDS), a supplemental survey about the resources (financial, time, and social-psychological) of children under the age of 18. Beginning in 2017, the Transition to Adulthood survey was given to all young adults in the PSID.

We detail the use and treatment of parent-child pairs, the coefficient of interest (student loan debt), the outcome variables (labor market status, retirement expectations, cash assets, and IRA assets), and important controls (information on the postsecondary institutions a parent was exposed to) below. The treatment of all other covariates is detailed in Table 32 in Appendix B.

In order to create aggregate measures of institutional exposure (for example, the percent of a parent's children whose first college entrance was at a historical Black college), we first link each child to each parent. We only include observations for which a parent-child pairing can be identified using the Main Family survey. The Main Family survey identifies the mother and father of an individual with unique individual identifiers if the mother and father also have been in the PSID. These unique individual identifiers were then used to link other information from the PSID about the parent or child to the parent-child pair. Within the Main Family data, state of residence is collected each survey year. We only include observations whose state of residence is known and within the United States.

The education data the PSID collects is central to this study. Both the Main Family survey and the Transition to Adulthood survey collect education information, including dates of college entry (for the Transition to Adulthood survey) and college exit (for the Transition to Adulthood and Main Family survey). For our study, information on the year a child began attending college is necessary because we assume the first year a child attends college is when a parent makes the decision to fund their child's education with parental loans or not. We use college entry and exit dates from *both* the Main Family and Transition to Adulthood surveys to maximize sample size,[18] but the two surveys ask different questions about college entry and exit dates, so the treatment of the

[18] Individuals whose college entry or exit years are ascertained by one survey are not necessarily ascertained by the other.

data varies by which survey the data comes from. We outline these differences below.

The Transition to Adulthood data explicitly asks for the entrance year for both their recent and prior college program. Only students who have attended at least two colleges list a prior college. We use the year a child began the earliest observed college endeavor as their beginning year. This means that if a child has both information on a prior college entrance year and a recent college entrance year, we treat the prior college entrance year as the child's beginning year.

To determine a child's beginning year of postsecondary education from the Main Family data, imputation is required. The Main Family data collects information on the year a Head or Spouse last attended college, the highest year of college completed (recorded as "completed one year," "completed two years," up to "completed five or more years"), and their highest degree type. There are two caveats to these data. The first caveat is that we are unable to include information from the Main Family data for individuals whose highest degree type is over a Bachelor's because in that case the data available for the year that individual last attended college refers to a degree higher than a Bachelor's (for which a parent could not have borrowed because parents are only eligible to borrow for a dependent child's education and graduate students are considered independents).[19] This means that for individuals with more than a Bachelor's degree, we cannot observe a college entrance year, so we do not include them. The second caveat is that information is not collected about the child's college entrance year. Because of this, we impute a child's beginning college year as the year they last attended college minus the highest year of college completed.

Much of the rich data collected by the PSID is public but some individually identifiable information is restricted and available only by application. For this study, we use restricted PSID data that identifies a child's undergraduate institution. The restricted PSID contains the IPEDS Unit ID for the child's undergraduate institution within both the Transition to Adulthood and the Main Family data. In the Transition to Adulthood data, we match a parent-child pair to the college that the child's college entrance year refers to (i.e., if a child's college entrance year refers to a prior college rather than a recent one, we match the prior college's IPEDS Unit ID to the parent-child pair). In the Main Family data, identifying information is collected on the college that the Head or Spouse (who here is a child of the parent we are interested in) received their highest degree from. (We only include parent-child pairs where the child's highest degree is a Bachelor's or less for reasons outlined in the above paragraph.) Unlike the Transition to Adulthood data, we lack information about college changes, so for the Main Family data we assume that the school a child's Bachelor's degree is from is also the school where a child began their degree. If both Transition to Adulthood and Main Family survey data on undergraduate institution is available, we use the data that has the earliest corresponding beginning year.

[19] This represents about 25% of college-educated of Heads and Spouses in any PSID Main Family survey.

We next detail our treatment of the dependent variables: individual retirement accounts, labor force status, and retirement expectations. The PSID collects information about household wealth, both with and without home equity, as calculated by the sum of seven asset types less the sum of eight debt types. For this paper, we are specifically interested in a family's Individual Retirement Accounts. Note the amount of dollars in IRAs are collected for the family as a whole rather than for individuals.

Labor force status is collected for all individuals. Individuals can be working now (employed), only temporarily laid off, looking for work (unemployed), retired, permanently disabled, keeping house, a student, or other. We only treat those who are looking for work as unemployed; we treat individuals who are retired, permanently disabled, keeping house, or a student as not in the labor force.

If a Head or Spouse has access to a pension or retirement plan at their current job, questions about retirement were asked in the Main Family survey. Respondents were asked at what age the Head or Spouse planned to stop working, with both a continuous response (actual expected age) and a binary response ("never") accepted. Respondents were asked about the usual retirement age of others who performed their job or that they worked with. Using this information, we create a "retirement gap" variable representing the distance between the individual's expected retirement age and the perceived retirement age of others. A positive retirement gap indicates the individual believes they will retire after the average retirement age of others; a negative retirement gap indicates the individual believes they will retire before the average retirement age of others.

Finally, the treatment of the coefficient of interest, student loans, follows. Within the Main Family survey, student loan debt information is collected for a family as a whole. For this study, our main interest is the effect of student loan presence in the household, thus distinguishing between loans taken out by parents and those taken out by students is not our focus and would be difficult to disentangle given the structure of the data available.

Summary statistics are listed below in Table 8. Statistics for dependent variables are listed first, followed by student loan information, and then data for other controls.

While the PSID collects information on many individuals, we end up with 772 observations in our main analysis. This is a result of several factors: (1) of the 27,709 parent-child pairs that the PSID identifies, only 7,640 parent-child pairs have links to children with restricted data available on the child's undergraduate institution, (2) the sample drops from 7,640 parent-child pairs to 5,435 parents once we collapse data to the parent level (2) only 3,510 of these parent-child pairs have information available about student loan presence in the household, (3) only 2,988 of these parent-child pairs have a college that can be matched to information available from IPEDS, and (4) general loss of sample size once

Table 8. Summary Statistics for Main Analyses.

	Mean	Min	Max
Retire Gap	1.37	-32.00	40.00
Usual Employment Status: Retirement Age (Others)	62.31	20.00	97.00
Expected Employment Status: Retirement Age (Self)	65.04	50.00	90.00
Never Expects to Retire	0.14	0.00	1.00
Employment Status: Retired	0.12	0.00	1.00
Employment Status: Not in the Labor Force	0.18	0.00	1.00
Employment Status: Unemployed	0.03	0.00	1.00
Employment Status: Employed	0.81	0.00	1.00
Tens of Thousands of Real 2015 Dollars in IRA/Annuities	11.35	0.00	356.05
Whether Anyone in Family has IRA/Annuities	0.32	0.00	1.00
Whether Anyone in Family has Student Loans	0.18	0.00	1.00
Amount of Student Loans in Family in Tens of Thousands of 2015 Dollars	0.51	0.00	25.00
Age	57.41	38.00	81.00
Social Security Eligibility	0.29	0.00	1.00
Spouse Present	0.69	0.00	1.00
Gender (Female = 1)	0.51	0.00	1.00
Black	0.29	0.00	1.00
Other Race	0.08	0.00	1.00
Hispanic	0.12	0.00	1.00
Health Status	2.56	1.00	5.00
Real Assets in Tens of Thousands of 2015 $s	49.62	0.00	3573.50
Household Income Less Own Labor Income in Tens of Thousands of 2015 $s	5.07	0.00	79.03
Real Assets Less Dollars in IRA/Annuities in Tens of Thousands of 2015 $s	38.26	0.00	3373.50
Survey Reference Year: Employment Status	2014.97	2011.00	2015.00
First entrance year of a parent's latest entering child	2004.69	1987.00	2013.00
Number of Children Matched in College	1.51	1.00	5.00
Total Years of College for All Children	4.56	0.00	19.00

Table 8. Continued.

	Mean	Min	Max
Financial Literacy (Scale = 0-6)	4.76	0.00	6.00
Earliest Wealth in Tens of Thousands of 2015 $s	14.08	-18.36	2059.11
Mother's Education: 0-5 Grades	0.01	0	1
Mother's Education: 6-8 Grades	0.12	0	1
Mother's Education: 9-11 Grades	0.14	0	1
Mother's Education: High School Graduate	0.46	0	1
Mother's Education: High School Graduate + Some Nonacademic Training	0.04	0	1
Mother's Education: Some College or Associate’s Degree	0.11	0	1
Mother's Education: Bachelor’s Degree	0.10	0	1
Mother's Education: Advanced Degree	0.03	0	1

controls are included because many individuals have missing information for at least one control.

Integrated Postsecondary Education Data System

Data on institutions comes from the Integrated Postsecondary Education Data System (IPEDS). IPEDS contains data on a wide variety of college information, including place, financing, and student makeup. For this study, we use data on a college's institutional control (public, private, or for profit); if the college is a Historically Black College; and if the college is a two-year or four-year institution. We merge this information onto the PSID data using the IPEDS Unit ID. Any observation that cannot be matched with an institution through the IPEDS data is dropped. We use dummy variables to represent the percent of a parent's children that have attended a certain type of college. For example, there are three categories for the percent of a parent's children that have attended a privately owned college: "None," "Some," or "All." Reducing the percentages to indicators is necessary because of the small sample size and nature of the restricted PSID data: no data with cell sizes smaller than 11 are allowed to be used in analyses. Institutional characteristics for colleges that parent's children attended are shown below in Table 9.

Results

We present results both for the effect of the presence of student loan debt within the family and the amount of student loan debt in the family in tens of thousands of 2015 dollars. For the main results, we examine the effect of these student loan debt variables on retirement expectations, employment status, and IRAs. In addition, we address the possibility that families who borrow are significantly different from families who do not borrow in three ways: (1) we present two robustness checks with additional controls, one controlling for the earliest wealth a family has and one controlling for the parent's financial literacy and the parent's mother's education, and (2) we look at the likelihood that current student loan debt burden can successfully predict which families had money problems in 1996, and (3) we report Oster's δ statistic and bias-adjusted β to give an idea of how much selection is due to unobservables.

For the robustness checks with additional controls, we present the effect only on employment status because we lose further observations once we introduce these controls and employment status has the most reliably large sample size, however results for the other dependent variables follow the results of the main analysis.

Table 10 shows the effect of having student loans within a family on a parent's retirement expectations. No retirement expectation variable is significantly affected by student loan presence. We do not find a significant effect on the likelihood a parent will never expect to retire, the age a parent expects to

Table 9. College Institution Characteristics for Parent Sample.

	Frequency	Percent
No Children Attended Historical Black College or University	769	93.78
Some Children Attended Historical Black College or University	17	2.07
All Children Attended Historical Black College or University	34	4.15
No Children Attended Private Institution	611	74.51
Some Children Attended Private Institution	76	9.27
All Children Attended Private Institution	133	16.22
No Children Attended For-Profit Institution	759	92.56
Some Children Attended For-Profit Institution	18	2.2
All Children Attended For-Profit Institution	43	5.24
No Children Attended Public Institution	179	21.83
Some Children Attended Public Institution	88	10.73
All Children Who Attended Public Institution	553	67.44
No Children Attended 4-Year or Higher Institution	255	31.1
Some Children Attended 4-Year or Higher Institution	70	8.54
All Children Attended 4-Year or Higher Institution	495	60.37

Table 10. The Effect of Any Student Loan Debt on Retirement Expectations.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Any Student Loans	-0.433	1.13	0.888	0.024
	(0.81)	(0.78)	(0.58)	(0.04)
Bias-Adjusted β (δ=1)	-0.44	1.57	1.389	0.03
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	10.86	-3.48	-1.88	-5.22
δ>1 ?	Yes	No	No	No
R^2	0.27	0.31	0.37	0.18
N	401	418	691	761

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

retire, or the "retirement gap" (the distance between when a parent predicts they will retire and when they believe others' with their same job retire). In addition to being insignificant, the direction of the coefficients does not tell a consistent story: the coefficients for "never" expect to retire and retirement age are positive (implying loans cause parents to expect to work longer) but the coefficient for the retirement "gap" is negative (implying loans cause parents to retire earlier).

Turning to the employment status analysis shown in Table 11, results indicate student loan debt presence shifts workers *into* the labor force and particularly into employment, though both of these coefficients are significant only at the 10% level. The likelihood of a parent being retired is not significantly impacted by student loan debt presence. The likelihood of not being in the labor force for any reason (retirement, student status, housemaker, or disabled) falls by 5.4 percentage points; and the likelihood of being employed increases 5.5 percentage points. The magnitude of the effect of student loan debt presence is in line with effect of mortgage and credit card debt on retirement found by Butrica & Karamcheva (2013). They found that the presence of credit card debt increased the likelihood an individual would be working by 4.3 percentage points and the presence of mortgage debt increased the likelihood an individual would be working by 7.4 percentage points.

The 0.1 percentage point difference we find between the likelihood of being in the labor force and being employed suggests the possibility that some of the increased likelihood of employment may be due to a reduction in unemployment as well, but the unemployment results do not display a significant

Table 11. The Effect of Any Student Loan Debt on Employment Status.

	Retired	Not in Labor Force	Unemployed	Employed
Any Student Loans	-0.035	-0.054*	-0.016	0.055*
	(0.02)	(0.03)	(0.02)	(0.03)
Bias-Adjusted β (δ=1)	0.00	-0.03	-0.02	0.03
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	1.20	1.84	49.06	1.87
δ>1 ?	Yes	Yes	Yes	Yes
R^2	0.37	0.3	0.14	0.3
N	771	771	771	771

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

effect and the R^2 for unemployment is notably lower than the R^2 for the other employment status regressions. The degree of selection on observables for these regressions all exceed the δ = 1.0 threshold, and the bias-adjusted treatment coefficients are all within 2.8 standard errors of the treatment coefficient of the original analyses.

Looking at the retirement savings results presented in Table 12, we do not find an impact of the likelihood of *having* an IRA or annuities within the family, but we find a very large effect on the amount of real 2015 dollars in an IRA or annuity. Results imply that the presence of any student loans within a family decreases the dollars in an IRA by a little more than $70,000 on average. The average per borrower amount of parental PLUS loan in the 2016-2017 school-year was $15,878 (CollegeBoard, 2018), or approximately one-fourth of the decrease in the IRA. If a parent took out this amount over a child's four-year college career, the amount of PLUS loan needing to be serviced would be smaller than the dollars in IRA reduction estimated ($63,512 vs. $70,040). This result also passes Oster's standards for robustness.

Tables 13, 14, and 15 show the effect of the amount of student loan debt on retirement expectations, employment status, and retirement savings, respectively. We lose several observations that reported the student debt in a household as a range (for example, less than $10,000 but more than $5,000) rather than a number. The pattern of results differs from those found when examining the presence of student loan debt. The results indicate an additional $10,000 in student loan debt in a household increases the age a parent expects to retire at by a little less than five months; note that this is only 7.7% of a standard deviation of expected retirement age.[20] Note however that the associated *δ*

[20] The standard deviation for retirement age in the sample is 4.89 years (4 years and 11 months).

Table 12. The Effect of Any Student Loan debt on Retirement Savings.

	Dollars in IRA/Annuities (Tens of Thousands of 2015 $s)	Whether Anyone in Family has IRA/Annuities
Any Student Loans	-7.004**	-0.074
	(3.03)	(0.06)
Bias-Adjusted β (δ=1)	-6.33	-0.05
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes
δ (β = 0)	7.21	2.57
δ>1 ?	Yes	Yes
R^2	0.37	0.35
N	767	767

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 13. The Effect of Amount of Student Loan Debt (In Tens of Thousands of 2015 $s) on Retirement Expectations.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Student Loans (2015 $s, Tens of Thousands)	-0.041	0.159	0.376**	0.001
	(0.15)	(0.17)	(0.16)	(0.01)
Bias-Adjusted β (δ=1)	-0.019	0.199	0.456	0.002
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	1.717	-6.923	-5.594	-1.166
δ>1 ?	Yes	No	No	No
R^2	0.27	0.31	0.39	0.19
N	400	417	687	756

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 14. The Effect of Amount of Student Loan Debt (in Tens of Thousands of 2015 $s) on Employment Status.

	Retired	Not in Labor Force	Unemployed	Employed
Student Loans (2015 $s, Tens of Thousands)	-0.004	-0.005	-0.001	0.005
	(0.00)	(0.01)	(0.00)	(0.01)
Bias-Adjusted β (δ=1)	0.001	0.000	-0.001	0.000
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	No	No	Yes	No
δ (β = 0)	0.856	1.012	3.493	1.717
δ>1 ?	No	Yes	Yes	Yes
R^2	0.37	0.3	0.14	0.3
N	763	763	763	763

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 15. The Effect of Amount of Student Loan Debt (in Tens of Thousands of 2015 $s) on Retirement Savings.

	Dollars in IRA/Annuities (Tens of Thousands of 2015 $s)	Whether Anyone in Family has IRA/Annuities
Student Loans (2015 $s, Tens of Thousands)	-1.411**	-0.013
	(0.59)	(0.01)
Bias-Adjusted β (δ=1)	-1.506	-0.013
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes
δ (β = 0)	-58.934	9.762
δ>1 ?	No	Yes
R^2	0.37	0.35
N	759	759

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

coefficient is less than zero, so the robustness of the result is unclear. An additional $10,000 does not decrease the likelihood of being retired or out of the labor force in general, nor does it increase the likelihood of employment, unlike the results in the analyses that looked at student loan debt presence. Having ten thousand dollars in student loan debt lowers a household's retirement savings by $14,110, or a little less than half of a standard deviation in retirement savings.[21] It is possible this is indicative of selection bias, here too $\delta < 0$, or that there are compounding effects of reducing IRA contributions, such as losing employer matching or interest from the account. However, if a parent owed $10,000 on a PLUS loan with an interest rate of 7.0%[22], the total amount a parent would pay over the *entire repayment period* of 10 years under the most expensive repayment plan is $15,025.[23]

For these base regressions, we report all coefficients outside of fixed effects in Appendix C in Tables 33, 34, 35, 36, 37, and 38. We next turn to the robustness checks.

Robustness Checks: Additional Controls

In the next part of the exploration of the effect of student loan debt, we present the coefficients of interest when a household's earliest level of wealth is included (shown in Tables 16 and 17) and coefficients of interest when a parent's financial literacy (on a scale from 0 to 6, 6 being the highest) and a parent's mother's education is included (shown in Tables 18 and 19). Financial literacy is measured by the number of questions concerning financial literacy an individual answered correctly on a supplemental survey in 2016. Maternal education is included in a categorical variable that ranges from 1 (elementary education) to 6 (advanced degree). We only examine employment status because the inclusion of the additional controls results in the loss of around a hundred observations and employment status is the dependent variable with the most observations.

Interestingly, in the robustness check with earliest wealth included, the additional control results in larger coefficient estimates in absolute value and no loss in significance. However, the estimate when parent's financial literacy and mother's education are included is insignificant and the δ falls below one, suggesting selection into borrower status may be driven by unobservables, perhaps unobservable individual characteristics best proxied for by maternal education and financial literacy.

Robustness Checks: Student Loan Debt and Money Problems

The key issue within this analysis remains the possibility of selection bias, which the robustness checks suggest may be problematic. One such possibility is that people with different levels of financial literacy or concern over financial problems may select into borrower status at different rates. One way that we addressed this issue is by including a measure of financial literacy in the

[21] The standard deviation for IRA dollars is $33,180.
[22] An interest rate of 7.0% is the PLUS loan interest rate for the 2017-2018 school year.
[23] Calculations using Federal Student Aid's Repayment Estimator (Federal Student Aid, 2018).

Table 16. The Effect of Any Student Loan Debt on Employment Status, Controlling for Earliest Wealth Available.

	Retired	Not in Labor Force	Unemployed	Employed
Any Student Loans	-0.034	-0.063*	-0.019	0.063*
	(0.024)	(0.034)	(0.014)	(0.034)
Bias-Adjusted β (δ=1)	-0.006	-0.033	-0.017	0.033
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	1.194	1.978	5.735	1.979
δ>1 ?	Yes	Yes	Yes	Yes
R^2	0.37	0.34	0.18	0.34
N	596	596	596	596

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17. The Effect of Amount of Student Loan Debt on Employment Status, Controlling for Earliest Wealth Available.

	Retired	Not in Labor Force	Unemployed	Employed
Student Loans (2015 $s, Tens of Thousands)	0.000	-0.002	-0.001	0.002
	(0.004)	(0.004)	(0.002)	(0.004)
Bias-Adjusted β (δ=1)	0.004	0.003	-0.001	-0.003
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	0.034	0.393	2.572	0.392
δ>1 ?	No	No	Yes	No
R^2	0.37	0.35	0.18	0.35
N	590	590	590	590

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18. The Effect of Any Student Loan Debt on Employment Status, Controlling for Parent's Financial Literacy and Parent's Mother's Education.

	Retired	Not in Labor Force	Unemployed	Employed
Any Student Loans	-0.023	-0.052	-0.020	0.052
	(0.031)	(0.041)	(0.019)	(0.041)
Bias-Adjusted β (δ=1)	0.024	0.000	-0.018	0.000
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	0.510	0.997	5.926	1.000
δ>1 ?	No	No	Yes	No
R^2	0.45	0.36	0.24	0.36
N	461	461	461	461

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19. The Effect of Amount of Student Loan Debt on Employment Status, Controlling for Parent's Financial Literacy and Parent's Mother's Education.

	Retired	Not in Labor Force	Unemployed	Employed
Student Loans (2015 $s, Tens of Thousands)	0.002	0.002	0.000	-0.002
	(0.004)	(0.005)	(0.002)	(0.005)
Bias-Adjusted β (δ=1)	0.009	0.011	0.001	-0.011
Bias-Adjusted β (δ=1) within 2.8 Standard Errors of Controlled β?	Yes	Yes	Yes	Yes
δ (β = 0)	-0.377	-0.256	0.304	-0.255
δ>1 ?	No	No	No	No
R^2	0.45	0.37	0.24	0.37
N	455	455	455	455

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

robustness check above. Below, we also address it by seeing if student loan debt *in the current period* is a significant predictor of financial problems *in the past*.

The measure of financial problems used here is from the Panel Study of Income Dynamics's 1996 Main Family survey. Within the 1996 survey, the PSID asked several questions about levels of financial distress within a family. From this financial distress data, we create an indicator for if an individual is a member of a family in 1996 that experienced one or more of six types of financial distress: the family (1) was unable to pay bills, (2) obtained a loan to pay off debts, (3) had creditors call, (4) had wages garnished, (5) had a lien on property, or (6) had property repossessed. In Tables 20 and 21 we present the relationship between student loan debt and the presence of money problems in the past. To attempt to ameliorate the possibility that the presence of historical money problems could decrease a family's ability to borrow student loans or increase the need to borrow, we only include observations where the latest child's college entrance year is at least five years *after* 1996.[24]

We find evidence that families with a history of financial problems are more likely to have student loan debt present in the household although we do not find an effect on the amount of student loan debt within the household. Taken with the robustness checks presented above, this indicates there may be systematic differences between parents in borrowing families and parents in non-borrowing families. In particular, we find evidence of a positive selection bias: families that have had money problems in the past are more likely to take out student loans, and these families may then be more likely to work longer, save less, and expect to retire later.

Conclusion

This paper is a first attempt to identify the effect of student loans on parents' outcomes, while also engaging in formal tests of robustness to determine if the results can be construed as causal. We examine the effect of student loan presence in the household on parents' retirement behaviors, expectations, and savings. Results indicate the presence of loans in the household significantly impacts retirement decisions along a number of dimensions: loan presence increases the likelihood a parent is in the labor force and employed and decreases the dollars in a household's IRA. An additional $10,000 in student loan debt amount increases expected retirement age by 5 months and decreases IRA dollars by $14,000. However, we find evidence a household's loan status is driven by previous financial hardship status. Further, a technique developed by Oster raises the possibility that results may be driven by a selection issue.

Given that our results indicate significant effects of student loans on parents on a number of margins, but that we cannot establish causality with certainty, it is an area ripe for more research. More research, however, calls for more and better data. While parent PLUS loans are included in the Title IV

[24] The credit check used for PLUS loans includes a five-year lookback for adverse credit history.

Table 20. Presence of Any Student Loan Debt as a Predictor for Historical Money Problems.

	Any 1996 Money Problems
Any Student Loans	0.114*
	(0.062)
R^2	0.23
N	635
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Table 21. Amount of Student Loan Debt as a Predictor for Historical Money Problems.

	Any 1996 Money Problems
Student Loans (2015 $s, Tens of Thousands)	0.011
	(0.011)
R^2	0.22
N	628
Robust standard errors in parentheses	
*** p<0.01, ** p<0.05, * p<0.1	

Volume Reports[25], they are not included in College's ScoreCards, unlike loans taken out by students. Further, surveys frequently used to examine the effect of post-secondary funding on later outcomes (like Baccalaureate & Beyond) solely examine students. We hope that this first effort will lead to greater attention and study focused on the parental PLUS loan program, both by policymakers and by researchers.

[25] Found at https://studentaid.ed.gov/sa/about/data-center/student/title-iv.

CHAPTER 3
MERIT AID & FAMILY DEBT IN LOTTERY SCHOLARSHIP STATES

Introduction

During the past few decades, two consistent trends in college financing have arisen. The first trend is that student loans now make up a much larger portion of the average student's college financing portfolio: from 1993 to 2014 there was a 40% increase in the number of college graduates who had borrowed student loans to fund their education. During the same time period, many states rolled out large scale merit-based scholarships. The interaction of these concurrent changes is the subject of this paper: I examine the effect of lottery scholarship eligibility on a broad measure of debt within a family, which includes student loans and credit card bills as well as other debt, using a simple difference-in-difference analysis.

Georgia introduced the first lottery-based merit aid scholarship in 1993 (Tennessee Higher Education Commission, 2016). Since the introduction of Georgia's HOPE scholarship, eight states have rolled out lottery-based merit aid scholarships. These lottery scholarships are differentiated from other merit aid scholarships not only by their funding source (a state lottery) but also because of their relatively generous aid packages. Sjoquist & Winter (2014) categorize all states with lottery scholarships as "strong merit aid states," taking both the level of scholarship aid and the percent of students who qualify into account.[26] I list these states, and the date that the state began disbursing lottery scholarships, in Table 22.

Lottery scholarships have a few clear purposes. The Tennessee Hope Scholarship program's purpose is explicitly "to provide access for Tennesseans to post-secondary education, to improve high school and collegiate academic achievement, to keep more of the best and brightest students in Tennessee, and to provide social and economic benefits to the state of Tennessee" (Tennessee Higher Education Commission, 2016). There is a large literature on the first-order impact of lottery scholarships. However, the impact of merit aid scholarships in general on debt uptake and amount is relatively understudied and the subject of only one published paper and two working papers thus far. During this time of rapidly increasing college tuitions, falling levels of state and federal funding for higher education, and rising student loan debt, the impact of merit-based financial aid on debt is worthy of attention.

Previous literature on student loan debt has found that student loans impact borrowers' decisions more than would be predicted by economic models since student loans (even those in tens of thousands of dollars) represent a small

[26] Of the "strong merit aid states" that Sjoquist & Winter (2014) identify, only Nevada does not have a lottery scholarship.

Table 22. Lottery Scholarship States Included in the Analyses & The Scholarship's Implementation Year.[27]

State	Lottery Scholarship Implementation Year
Georgia	1993
Florida	1997
New Mexico	1997
South Carolina	1998
Louisiana	1998
Kentucky	1999
West Virginia	2002
Tennessee	2004

amount of the typical college-going person's lifetime income. Nevertheless, the literature has found that student loans impact career choices, decrease overall levels of wealth, and decrease the likelihood that an individual will be married (Rothstein & Rouse, 2011; Field, 2009; Cooper & Wang, 2014; Gicheva, 2016). In addition, descriptive analyses have found that student loan borrowers have higher levels of credit card debt than non-borrowers and higher debt-to-household-income (Fry, 2014).

Does merit aid lead to lower debt levels for young adults? I seek to answer this question and add to the research in this area using a difference-in-difference analysis to examine the effect of lottery scholarships on a broad measure of debt that includes not only student loan debt but also credit card debt, medical bills, legal bills, and loans from relatives. I define treated individuals as those who were enrolled in a cohort in college during which a lottery scholarship was available in their state of residence. I use data from the Panel Study of Income Dynamics (PSID), a panel that has been running since 1968 and specifically collected information on a broad measure of debt (credit card charges, student loans, medical and legal bills, and loans from relatives) in 1984, 1989, 1994, and 1999 to 2015. The longer time frame of this panel allows me to use a difference-in-difference methodology that takes advantage of variation in the rollout of these programs both across time and across states. The methodology and dataset enable me to examine multiple cohorts and the effect of these scholarships on all treated individuals who attended college, not only those who graduated, and the effect of these scholarships on a broad debt measure.

The results of the analyses suggest that merit aid scholarships do not lower debt levels or the likelihood of debt uptake. These insights add to a literature that has found divergent results through the use of a broader debt measure and an analysis that examines a longer timeframe, both college graduates and non-graduates, and multiple states.

[27] *Source: Sjoquist & Winters (2014), Table 1.*

Literature Review

The other related papers on this topic have all approached this question in slightly different ways. Chen and Wiederspan (2014), the first paper to examine this question, utilize the Baccalaureate & Beyond 2000/2001 data survey of 1999 – 2000 graduates. During this time, only Georgia (whose HOPE program was the first strong merit aid program, as defined by Sjoquist and Winters (2014)) would have treated graduates to be surveyed. Chen and Wiederspen examine the impact of merit aid (in general rather than only large-scale merit aid) on student borrowing. They find that when Georgia is included in the analysis, merit aid significantly lowers student borrowing; however, when Georgia is excluded, this significance disappears. They dub this "the Georgia Effect": "This finding may imply that the relationship between state non-need (merit) aid and graduates' debt burden in Georgia was so strong that it drove the results for all state related covariates" (Chen & Wiederspan, 2014, pg. 585). Why would this be the case? The authors found that Georgia's merit aid level was several times larger in magnitude than the average state's; they propose such a large level of merit aid decreases student loan uptake.

A working paper by Chapman (2015) also utilizes a later Baccalaureate & Beyond survey of the 2008 – 2009 graduating cohort, so the examined cohort contains a larger number of students who would have been exposed to a large-scale scholarship and these students have greater geographic variation than in Chen & Wiederspan (2014). Chapman uses differing ACT score eligibility thresholds for both a difference-in-difference analysis and a regression discontinuity analysis. Because she uses the ACT score eligibility threshold, she restricts the merit aid states she examines to those that use the ACT (or comparable SAT) as a criterion of eligibility. In the difference-in-difference analysis, she exploits variation over state and over ACT score (i.e., a graduate is "treated" if their ACT score is at or above the ACT cutoff and their state has a merit aid scholarship). She finds a consistent negative effect of merit aid eligibility on student loan amount; merit aid significantly reduces the amount of student loan debt the average student will borrow in both analysis types. In the regression discontinuity analysis, aid eligibility reduces total loan amount by between $7200 and $7600; in the difference-in-difference analysis, aid eligibility reduces total loan amount by between $5800 and $6400. She finds evidence for a reduction in the likelihood of loan uptake of around 8 percentage points with the difference-in-difference analysis but no significant effect with the regression discontinuity approach.

Because both Chapman (2015) and Chen & Weiderspan (2016) use the Baccalaureate & Beyond data, they are only able to examine college graduates. Scott-Clayton and Zafar (2016) use a unique data set from West Virginia that links students' academic information with later-in-life credit information. Interestingly, they find greater effects of student loans on the intensive margin rather than the extensive margin, in contrast to Chapman (2015). Like Chapman (2015), they use graduates' ACT scores as the discontinuity in a regression discontinuity analysis and a difference-in-difference analysis with ACT score as one source of treatment variation but the other source they exploit is college entrance year rather than state (i.e. they view a student as "treated" with West Virginia's PROMISE scholarship if the student met the eligibility threshold and began college in a year when the PROMISE was available). Within the regression discontinuity results, Scott-Clayton and Zafar find that the West Virginia PROMISE does decrease the likelihood of student loan debt uptake (by around 7 percentage points). However, they do not find an effect on the average dollar amount of student loan debt that a graduate has because there are countervailing effects: likely PROMISE recipients borrow significantly fewer loans at the undergraduate level but are significantly more likely to attend graduate school and borrow more loans at the graduate level than unlikely PROMISE recipients. The difference-in-difference analysis indicates lower likelihood of student loan uptake but not with the consistency of the regression discontinuity, suggesting the local average treatment effect and the average treatment effect may be dissimilar. In addition, Scott-Clayton and Zafar (2016) also measure the effect of merit aid scholarships on credit card balances. They do not find any significant effect.

Model

The difference-in-difference approach I propose uses both geographic variation and time variation. My estimating equation is as follows

$$Y_i=\alpha+\beta\left(MeritState_s{*}After_t\right)+\theta(StateFE_s)+\vartheta(EntranceYearFE_t)+X_i\varphi+u_i$$

In this estimation, Y_i is a broad debt measure within the household that individual i physically resides in during the most recent survey after the student has stopped attending college. This debt measure includes student loans, credit card bills, medical bills, legal bills, and loans from relatives. I control for state and cohort fixed effects and consider individuals as "treated" with eligibility for a merit aid scholarship if they entered college in a state-year where a lottery scholarship would be available. X_i is a vector of student characteristics, including demographics (age when debt surveyed, gender, and race), the distance between the person's college exit year and year their debt was surveyed, the total years of college education for all individuals within the household, and the individual's father's education level.[28] All analyses use robust standard errors and person-level weights.

The key threat to identification in any difference-in-difference analysis is the validity of the control group. If the control group and the treated group do not have parallel trends in counterfactual outcomes once observables are accounted for, then any estimated impacted of treatment may not be the result of a policy change but of differing paths altogether. In this study, a plausible threat to

[28] Unfortunately, the Main Family PSID does not contain information on student's high school performance or SAT or ACT score to control for a student's academic characteristics prior to college. I include parental education as a proxy for the importance of education within the family.

identification would be if an underlying trend in student financing was occurring in states that introduced lottery scholarships that was not occurring in states that did not introduce them. I examine the results visually and econometrically with an event study design. I do not find any evidence of different pre-existing trends – instead, both pre- and post-treatment years display significant amounts of noise for either amount of debt or debt presence.

An additional worry with this particular analysis is that states' lottery scholarships may have changed the student makeup within the state. Perhaps large scale scholarships lessen "brain drain" (the phenomenon of higher ability students attending out-of-state schools) or encourage students to enroll in college who otherwise would not have attended. I do not address either of these possibilities directly in my analyses, but previous literature in general indicates that large scale scholarships do not affect most students' choice to attend college although it may change their sector choice (Bruce & Carruthers, 2014; Goodman, 2008). The literature does suggest that effects are more pronounced for low income students and thus need-based aid programs may have greater impact on college choice than merit-based: Kane (2003) finds a significant effect on enrollment decisions when examining the effect of a California scholarship program based on both merit and need.

Data

The data I use to explore the effect of large-scale merit aid on a broad measure of debt is the Panel Study of Income Dynamics (PSID). The PSID is a nationally representative survey that has followed families since 1968; currently the PSID is given every other year to participating families. PSID families answer both Individual and Main Family surveys each year.

The PSID is suited for the purposes of this study because the survey contains information both about a family's state of residence each year and debt presence and amount (student loan debt grouped with a small set of other debt types in 1984, 1989, 1994, and 1999 to 2009, and separated subcategories for 2011 to 2015). Below I detail how I treat these debt variables.

The PSID collects information both about the presence of debt ("…do you (or anyone in your family living there) currently have any other debts…?") and about the amount of debt ("If you added up all of these debts (for all of your family living there), about how much would they amount to right now?"). The "other debt" measure the PSID used prior to 2011 explicitly asked about student loan debt, credit card debt, medical bills, legal bills, and loans from relatives *combined*, but as is shown below, within my sample student loan debt is the largest average debt amount and it and credit card debt are the most prevalent. This data comes with two important caveats: (1) the Main Family survey asks the amount of any debt for *all* members of a household rather than individually (i.e., "If you added up all student loans for all of your family living there, about how much would they amount to right now?"), and (2) the Main Family survey data only collects information about college *exit* year, years of college completed, and year of high school graduation, *not* college entrance year. This is problematic because merit aid eligibility rests on assumptions about college entrance year (merit aid programs are implemented in a specific year, and individuals are assumed eligible based on when they enter college). I impute college entrance year using earliest college exit minus the years of college completed in the same survey year. This proxy for college entrance will be accurate if an individual does not take breaks during her college career. To bolster this analysis, I also use an alternative proxy for college entrance year, high school graduation year. High school graduation year may be a more accurate proxy for college entrance, but the PSID has only collected data on high school graduation in 2015, so the sample size is significantly smaller.

Both the advantage and disadvantage of using the broad debt measure data is the ability to examine the effect of large-scale merit aid not on student loan debt in particular but on the effect of large-scale merit aid on a more general measure of debt that includes credit card debt. However, this broad debt measure does not allow me to examine any dynamics that exist *between* types of debt. For example, if individuals treated with merit aid are less likely to take out student loans but more likely to take out credit card debt (or vice versa), the following analyses will be unable to pick up on those effects.

However, using this debt measure allows the creation of an analysis that goes back much further in time than using a student loan specific measure: the question about other debt was asked in 1984, 1989, 1994, and then every survey from 1999 through 2009. For 2011, 2013, and 2015, I construct a proxy variable for this larger debt variable using the separate measures for the five debt types. I assign the debt to a person for the most recent year after their college exit where I observe their debt level. Individuals in the 99th percentile of indebtedness are dropped so that outliers do not drive results. All variables in dollars have been transformed into constant 2015 dollars.

Within the Main Family data, debt is surveyed at the level of the household rather than individual, so I allow fractional measures of merit aid treatment. Families in the analysis for whom I can readily assign as having been exposed to a large-scale merit aid program or not in the family's entirety are simply coded 0 (not treated) or 1 (treated). I allow fractional treatment if members of a family have differing exposure (i.e., if a Head was not treated with aid eligibility but his Spouse was).

I present the summary statistics for the data in Table 23. About 11.5% of the sample was treated and began college in a year when a merit aid scholarship was available. Most of the sample has some form of other debt, and the average amount of debt, including those with no debt at all, is just over $15,100. The average person in the sample began college in about 1998, which indicates that there is a fair amount of variation *within* state because in 1998 five of the eight

Table 23. Summary Statistics for Sample.

	Mean	Standard Deviation	Count	Min	Max
Any Other Debt	.7227666	.4477618	1735	0	1
Amount Real Other Debt (2015)	15118.99	20827.35	1735	0	112477
Scholarship Possible	.1146974	.3001107	1735	0	1
Earliest College Beginning	1997.918	7.433334	1735	1985	2013
College State (FIPS)	24.01787	13.55582	1735	1	51
Distance between College Exit and Survey	2.208646	1.11156	1735	1	5
Age	23.49625	5.3859	1735	17	65
Female	.6069164	.488576	1735	0	1
Years in College (Household)	3.718732	2.244401	1735	1	10
Father's Education: No High School Diploma	.1348703	.3416835	1735	0	1
Father's Education: High School Graduate, no Bachelor's Degree	.5394813	.4985825	1735	0	1
Father's Education: Bachelor's Degree Recipient or Higher	.3256484	.4687517	1735	0	1
Race: White	.6789625	.4670098	1735	0	1
Race: Black	.2945245	.4559601	1735	0	1
Race: Other Race	.026513	.1607013	1735	0	1

large scale merit aid scholarships would have launched. The sample skews more female, but this is reflective of the fact more women attend college than men.[29]

The final analyses end up with a sample size of 1,735. Of the 77,000 individuals within the PSID, only 3,000 have information on debt, college enrollment, and college state. Only a little over half of this number have non-missing information for the rest of the variables included.

Information about subcategories of debt is available in 2011, 2013, and 2015. In Figure 3, I present a breakdown of the percent of individuals that have each of the five subcategories of debt that make up the broad measure of debt. In Figure 4, I present a breakdown of the average amount of debt for those five subcategories for all individuals.

Note that within this sample, student loan debt is both the most likely subcategory of debt to be present and the subcategory of debt with the largest average amount. Credit card debt is a close second to student loan debt in terms of presence, but the average credit card debt never exceeds $5,000 in any of the years subcategory information is available for and the average amount of student loan debt never falls below $15,000. Two of the other debt types, legal bills and loans from relatives, have such a low average amount they are difficult to identify in the graphs.

Results

I analyze the effect of large scale lottery scholarships using a difference-in-difference analysis on the presence and amount of a broad measure of debt that includes credit card bills, student loan debt, medical bills, legal bills, and loans from relatives. Table 24 shows the effect of enrolling in college in a state-year where a lottery scholarship was available; in Table 24 I use an estimate of college entrance year where entrance year is assumed to be the last year a person attended college less the years of college completed.

None of the coefficients of interest are statistically significant. The results do not indicate that lottery scholarship availability when an individual begins college effects the likelihood the individual's family will experience an increase or decrease in the presence of debt, debt amount, or logged debt (which would capture a percent change in debt). These insignificant coefficients are negative, in line with the significant findings within the previous literature. The magnitudes of these estimates would indicate that families with individuals with scholarship eligibility are slightly less likely to have debt (0.6 percentage points less likely), on average have about $1,801 less in debt, and a 17.4% decrease in the amount of debt, all else equal. The magnitude of the amount of debt result are similar to the

[29] See https://nces.ed.gov/fastfacts/display.asp?id=72 for a breakdown of degree recipients both by race and by gender.

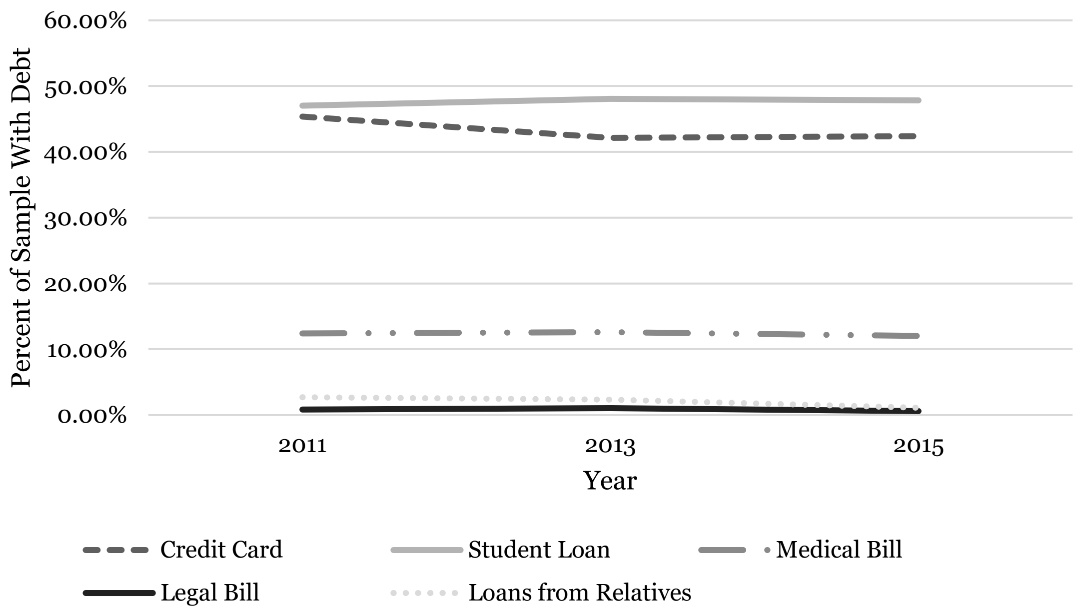


Figure 3. Any Debt for Individuals in Sample, Broken Down by Subcategories and Years.

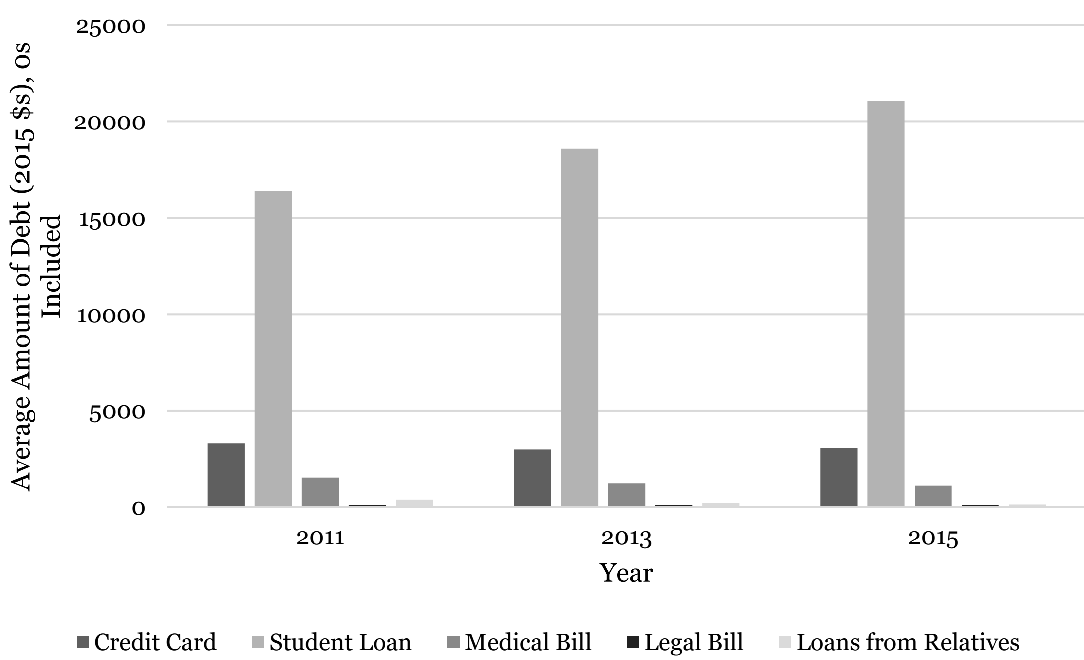


Figure 4. Average Amount of Debt for Individuals in Sample, by Subcategories and Years.

Table 24. Effect of Lottery Scholarship Availability on Debt; College Exit Year Less Years College Completed Used as College Entrance Year.

	Any Other Debt	Amount Real Other Debt (2015)	Log of Amount Real Other Debt (2015) (=0 if Amount of Real Other Debt =0)
Scholarship Available	-0.00666	-1,801	-0.174
	(0.0737)	(3,447)	(0.696)
R^2	0.114	0.170	0.097
N	1,735	1,735	1,735

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

regression discontinuity estimate found by Scott-Clayton and Zafar (2016), who find that West Virginia PROMISE scholarship recipients had $1,400 less *undergraduate* debt on average, but the magnitude is about five times smaller than those found by Chapman (2015). Additionally, effect size of merit aid on debt presence is much smaller than that found by Chapman (2015) or Scott-Clayton and Zafar (2016), both of which found effect sizes of around 8 percentage points.

One caveat to this result is that this analysis uses an imputed college entrance year. Because college entrance year determines if a person is treated with scholarship eligibility, college entrance year is central to the analysis. Below, I present another set of regressions in Table 25 where the individual's college entrance year is treated as the year she graduated high school. The pairing of the two analyses is also informative because the use of an imputed college entrance year is likely to create some false positives (individuals who entered college in lottery scholarship states and took a break during college that the researcher cannot observe actually entered college prior to the implementation of the lottery scholarship, would incorrectly be designated as "lottery scholarship eligible"), whereas the use of the high school graduation year as the college entrance year is likely to create false negatives (individuals who delayed college entrance in lottery scholarship states and graduated high school prior to the lottery scholarship program but did not enter college until after the implementation of the scholarship, would be incorrectly designated "lottery scholarship ineligible").

Although a few hundred observations are lost due to this scholarship eligibility question ("When did you graduate high school?") only being asked in the 2015 survey, this second analysis supports the results of the first. The presence of the broad debt measure, the amount of debt in levels, and amount of debt in logs are not significantly affected by large scale merit aid scholarship eligibility, even with the alternative college entrance year measure, although the magnitudes are uniformly larger. Results with all coefficients except state and year fixed effects for both analyses can be found in Appendix D in Tables 39 and 40.

Table 25. Effect of Lottery Scholarship Availability on Debt; High School Graduation Year Used as College Entrance Year.

	Any Other Debt	Amount Real Other Debt (2015)	Log of Amount Real Other Debt (2015) (=0 if Amount of Real Other Debt =0)
Scholarship Available	-0.0811	-4,920	-0.795
	(0.0992)	(4,474)	(0.994)
R^2	0.140	0.198	0.127
N	1,325	1,325	1,325

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The magnitude of the effect of scholarship eligibility on debt presence is in line with that found in previous literature (Chapman, 2015; Scott-Clayton & Zafar, 2016), and the effect of scholarship eligibility on debt amount is within the fairly large range within the previous literature. Chapman (2015) found a reduction in student loan debt amount between $7,500 and $6,000 when a student was likely to be treated with a merit aid scholarship; Scott-Clayton and Zafar (2016) found a reduction in undergraduate student loan debt of $1,400 but an increase in graduate student loan debt.

In addition to these analyses, I present graphs of event studies in Figure 5 and Figure 6. I do this to examine the possibility there is some type of trend occurring around the implementation year – perhaps debt initially falls with the availability of lottery scholarship funds, but as institutions adjust to the influx of dollars, tuition rises and so does the level of student borrowing. In both event study figures, t = 0 signifies the implementation year of the lottery scholarship, and t = -1 (the year prior to the scholarship's implementation) is the omitted year. Both event studies use imputed college entrance (college exit year less years in college) because of the larger sample size.

The event studies exhibit a significant amount of noise, both before and after the implementation of the lottery scholarship. In addition, the effect of the lottery scholarship is visually centered on zero and zero is always within the 95% confidence intervals for each t, regardless of if the graph is examining the presence of debt or the amount of debt.

In summary, I do not find any evidence that lottery scholarships lower the amount of debt or likelihood of debt uptake. I examine the relationship analytically, using two different measures of college entrance year, and visually and econometrically with an event study. I discuss why these results may diverge from those found in the previous literature in the conclusion.

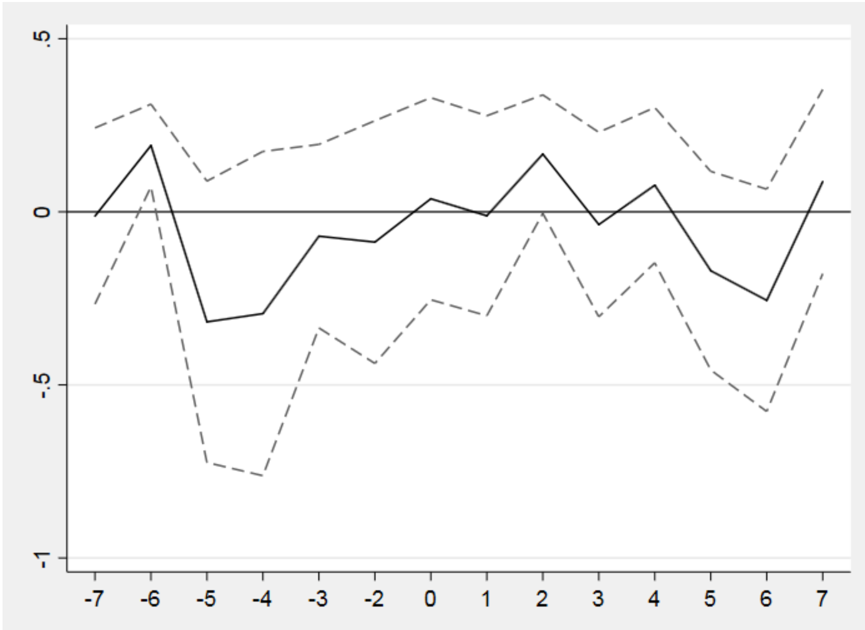


Figure 5. Presence of Broad Debt Measure.

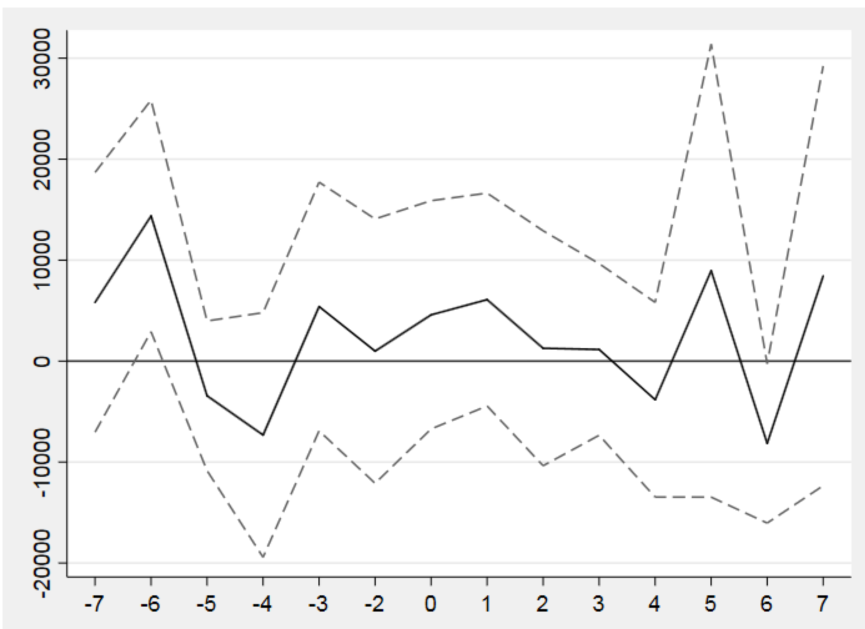


Figure 6. Amount of Debt (2015 $’s).

Conclusion

Previous literature has found some evidence of lower levels and uptake of student loan debt when individuals are treated with large amounts of merit aid. Chen and Wiederspan (2014) found a significant effect of merit aid on student loans only when Georgia, the only state with a large merit aid program in existence during the cohort that Chen and Wiederspan examined, was included. Scott-Clayton and Zafar (2016), meanwhile, find evidence within their regression discontinuity analysis of decreased student loan uptake and lower amounts borrowed at the undergraduate level, and increased student loan uptake and higher amounts borrowed at the graduate level.

It is possible that, because the analyses presented here cannot distinguish between undergraduate and graduate borrowing, the contrasting impact end up masking any true significant effect. It is also possible that the analyses presented here cannot distinguish between the different types of debt within the debt measure and thus any substitution from student loans to credit card debt or vice versa is masked. However, the finding that families where individuals are treated with merit aid do not have lower or higher levels of *total* debt remains important, suggesting that merit aid does not lead to lower debt levels on average.

The diverging results within this literature emphasize that there is still much to be learned about the way individuals shape their college financing portfolios. Both student loan debt and these large merit aid scholarships make up billions of dollars of higher education spending: in the 2016-2017 school-year, over $58 billion dollars in student loans were disbursed to undergraduate students, and South Carolina alone disbursed over $100 million in aid with its Hope scholarship program (College Board, 2018; South Carolina Commission on Higher Education, 2018). Both policymakers and the public would benefit from further research on the way that merit aid scholarships interact with other sources of funding, especially debt.

REFERENCES

Agnarsson, S., & Carlin, P. S. (2002). Family background and the estimated return to schooling: Swedish evidence. *Journal of Human Resources*, 680-692.
Anguelov, C., & Tamborini, C. R. (2009). Retiring in debt? Differences between the 1995 and 2004 near-retiree cohorts. *Social Security Bulletin*,*69*(2), 13-34.
Berlinski, S., Galiani, S., & Manacorda, M. (2008). Giving children a better start: Preschool attendance and school-age profiles. *Journal of Public Economics*, *92*(5-6), 1416-1440.
Bleakley, H. (2007). Disease and development: Evidence from hookworm eradication in the American South. *The Quarterly Journal of Economics*, *122*(1), 73-117.
Bruce, D. J., & Carruthers, C. K. (2014). Jackpot? The impact of lottery scholarships on enrollment in Tennessee. *Journal of Urban Economics*, *81*, 30-44.
Butrica, B., & Karamcheva, N. (2013). Does household debt influence the labor supply and benefit claiming decisions of older Americans? Center for Retirement Research at Boston College Working Paper.
Cascio, E. U. (2009). *Do investments in universal early education pay off? long-term effects of introducing kindergartens into public schools* (No. w14951). National Bureau of Economic Research.
Cash, R. A., & Narasimhan, V. (2000). Impediments to global surveillance of infectious diseases: consequences of open reporting in a global economy. Bulletin of the World Health Organization, 78(11), 1358-1367.
Casselman (2014) What Baby Boomers' retirement means for the U.S. economy. FiveThirtyEight. Retrieved from https://fivethirtyeight.com/features/what-baby-boomers-retirement-means-for-the-u-s-economy/.
Cellini, S. R., & Darolia, R. (2016). Different degrees of debt: Student borrowing in the for-profit, nonprofit, and public sectors. Washington, DC: Brookings Institution.
Center for Disease Control. (2016). Poliomyelitis. In *The Pink Book: Epidemiology and Prevention of Vaccine-Preventable Diseases*. Retrieved from https://www.cdc.gov/vaccines/pubs/pinkbook/polio.html.
Chapman, S. (2015). Student loans and the labor market: Evidence from merit aid programs. Working paper. Retrieved from https://cpb-us-e1.wpmucdn.com/sites.northwestern.edu/dist/7/687/files/2015/09/Download-Job-Market-Paper-PDF-274452u.pdf.
Chen, R., & Wiederspan, M. (2014). Understanding the determinants of debt burden among college graduates. *The Journal of Higher Education*, *85*(4), 565-598.
College Board. (2014). *Trends in Student Aid*. Retrieved from http://trends.collegeboard.org/student-aid
College Board. (2017). *Trends in Student Aid*. Retrieved from http://trends.collegeboard.org/student-aid.

College Board. (2018). *Trends in Student Aid*. Retrieved from http://trends.collegeboard.org/student-aid.

Cooper, D., & Wang, J. C. (2014). Student Loan Debt and Economic Outcomes. Current Policy Perspective No. 14-7. *Federal Reserve Bank of Boston*.

Currie, J., Hanushek, E. A., Kahn, E. M., Neidell, M., & Rivkin, S. G. (2009). Does pollution increase school absences? *The Review of Economics and Statistics,* 91(4), 682-694.

Dancy, K. (2016). Who takes out PLUS loans anyway? Blog post at *New America*. Retrieved from https://www.newamerica.org/education-policy/edcentral/parent-plus-loans/.

Dredge, B. (2008). Poliomyelitis and American Popular Culture. In J. P. Byrne (Ed.), *Pestilence, Pandemics, and Plagues*. (p. 550 – 553). Westport, CT: Greenwood Press.

Fagernäs, S. (2015). Papers, please! The effect of birth registration on child labor and education in early 20th century USA. *Explorations in Economic History,* 52, 63-92.

Federal Student Aid. (2016a). Interest Rates and Fees. Retrieved from https://studentaid.ed.gov/sa/types/loans/interest-rates.

Federal Student Aid. (2016b). Subsidized and Unsubsidized Loans. Retrieved from https://studentaid.ed.gov/sa/types/loans/subsidized-unsubsidized.

Federal Student Aid. (2018). Repayment Estimator. Retrieved from https://studentloans.gov/myDirectLoan/mobile/repayment/repaymentEstimator.action.

Field, E. (2009). Educational Debt Burden and Career Choice: Evidence from a Financial Aid Experiment at NYU Law School. *American Economic Journal: Applied Economics*, 1-21.

Feinstein, L., Robertson, D., & Symons, J. (1999). Pre-school education and attainment in the National Child Development Study and British Cohort Study. *Education Economics*, *7*(3), 209-234.

Fishman, R. (2018). The Wealth Gap PLUS Debt: How Federal Loans Exacerbate Inequality for Black Families. *New America*. Retrieved from https://www.newamerica.org/education-policy/edcentral/wealth-gap-plus-debt/.

Fitzpatrick, M. (2006). The Cutter incident: How America's first polio vaccine led to a growing vaccine crisis. *Journal of the Royal Society of Medicine*, *99*(3): 156.

Fitzpatrick, M. D., Grissmer, D., & Hastedt, S. (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review, 30*, 269-279.

Gicheva, D. (2016). Student loans or marriage? A look at the highly educated. *Economics of Education Review*, *53*, 207-216.

Goodman, J. (2008). Who merits financial aid?: Massachusetts' Adams scholarship. *Journal of Public Economics*, *92*(10-11), 2121-2131.

Goodman, J. (2014). Flaking out: Student absences and snow days as disruptions of instructional time. No. w20221. National Bureau of Economic Research, 2014.

Hackett, B. (2016). HBCU Enrollment, PLUS Loan Awards Decline After Credit Standards Change. *National Association for Financial Aid Administrators*. Retrieved from https://www.nasfaa.org/news-item/484/Report_HBCU_Enrollment_PLUS_Loan_Awards_Decline_After_Credit_Standards_Change.

Hai, W., Zhao, Z., Wang, J., & Hou, Z. G. (2004). The short-term impact of SARS on the Chinese economy. Asian Economic Papers, 3(1), 57-61.

Haines, M. R., and Inter-university Consortium for Political and Social Research. (2010). Historical , Demographic, Economic, and Social Data: The United States, 1790-2002. ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. Retrieved from https://doi.org/10.3886/ICPSR02896.v3.

Hamilton, L. T. (2013). More is more or more is less? Parental financial investments during college. *American Sociological Review*, *78*(1), 70-95.

Hanna, D., & Huang, Y. (2004). The impact of SARS on Asian economies. *Asian Economic Papers, 3*(1), 102-112.

Hudson, J., & Sessions, J. G. (2011). Parental education, labor market experience and earnings: New wine in an old bottle? *Economics Letters*, *113*(2), 112-115.

Humphreys, M. (2002). No safe place: disease and panic in American history. *American Literary History*, *14*(4), 845-857.

Kane, T. J. (2003). *A quasi-experimental estimate of the impact of financial aid on college-going* (No. w9703). National Bureau of Economic Research.

Kunschner, L. J. (2008). Poliomyelitis. In J. P. Byrne (Ed.), *Pestilence, Pandemics, and Plagues*. (p. 547 – 550). Westport, CT: Greenwood Press.

Lusardi, A., & Mitchell, O. S. (2007). Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth. *Journal of Monetary Economics*, *54*(1), 205-224.

Lusardi, A., Mitchell, O. S., & Oggero, N. (2017). *Debt and financial vulnerability on the verge of retirement* (No. w23664). National Bureau of Economic Research.

Mann, A. (2011). The effect of late-life debt use on retirement decisions. *Social Science Research*, *40*(6), 1623-1637.

Maron, D. F. (2016). Why we shouldn't quarantine travelers because of Zika. *Scientific American*. Retrieved from https://www.scientificamerican.com/article/why-we-shouldn-t-quarantine-travelers-because-of-zika/.

Martinez-Bakker, M., King, A. A., & Rohani, P. (2015). Unraveling the transmission ecology of polio. *PLoS Biology*, *13*(6).

Meyers, K., & Thomasson, M. A. (2017). Paralyzed by panic: Measuring the effect of school closures during the 1916 polio pandemic on educational

attainment. NBER Working Paper No. 23890. Retrieved from http://www.nber.org/papers/w23890.
Minicozzi, A. (2005). The short term effect of educational debt on job decisions. *Economics of Education Review*, *24*(4), 417-430.
Munnell, A. H., Hou, W., & Sanzenbacher, G. T. (2017). *Do households have a good sense of their retirement preparedness?* Center for Retirement Research at Boston College Brief #17-4.
National Center for Education Statistics. (2015). Table 5.3. Types of state and district requirements for kindergarten entrance and attendance, by state: 2014. *State Education Reforms*. Retrieved from https://nces.ed.gov/programs/statereform/tab5_3.asp.
Oster, E. (2017). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 1-18.
Panel Study of Income Dynamics (2018). Public use dataset. Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.
Panel Study of Income Dynamics (2017). PSID 2015 Release – Data Highlights. Retrieved from https://psidonline.isr.umich.edu/guide/2015DataHighlights.pdf.
Rhode, S. (2015). How can I payoff massive parent PLUS Loans my mom can't afford? Huffington Post: Get Out of Debt Guy Blog. Retrieved from http://www.huffingtonpost.com/steve-rhode/how-can-i-payoff-massive_b_6987802.html.
Rothstein, J., & Rouse, C. E. (2011). Constrained after college: Student loans and early-career occupational choices. *Journal of Public Economics*, *95*(1), 149-163.
Ruggles, S., Genadek, K., Goeken, R., Grover, J., & Sobek, M. (2017). Integrated Public Use Microdata Series: Version 7.0 [dataset]. Minneapolis: University of Minnesota. Retrieved from https://doi.org/10.18128/D010.V7.0.
Rustein, D. D. (1957). How good is the polio vaccine? *The Atlantic*. Retrieved from https://www.theatlantic.com/magazine/archive/1957/02/how-good-is-the-polio-vaccine/303946/.
Sallie Mae (2015). How America pays for college 2014. *Sallie Mae's National Study of College Students and Parents Conducted by Ipsos Public Affairs (Washington: Salle Mae).*
Sallie Mae (2017). How America pays for college 2017. *Sallie Mae's National Study of College Students and Parents Conducted by Ipsos Public Affairs (Washington: Salle Mae).*
Scott-Clayton, J., & Zafar, B. (2016). *Financial aid, debt management, and socioeconomic outcomes: post-college effects of merit-based aid* (No. w22574). National Bureau of Economic Research.
Sirkin, M. G., & Brenner, B. (1960). Population characteristics and participation in the poliomyelitis vaccination program. Public Health Service Publication, Washington, DC.

Sjoquist, D. L., & Winters, J. V. (2014). Merit aid and post-college retention in the state. *Journal of Urban Economics*, *80*, 39-50.
Smith, J. S. (1990). *Patenting the Sun*. William Morrow and Company, Inc., New York: New York.
Smithsonian Behring Center. (2016). Whatever happened to polio? Retrieved from http://amhistory.si.edu/polio/.
South Carolina Commission on Higher Education. (2018). *South Carolina Higher Education Statistical Abstract 2017*. Retrieved from http://www.che.sc.gov/CHE_Docs/finance/abstract/Abstract-2017-web.pdf.
Specter, M. (1985). AIDS victim's right to attend public school tested in corn belt. *The Washington Post*. Retrieved from https://www.washingtonpost.com/archive/politics/1985/09/03/aids-victims-right-to-attend-public-school-tested-in-corn-belt/8ffb6ac0-93cc-4c6f-88da-f3bb23664294/?utm_term=.275aa9381764.
Stolper, H. (2014). Home Equity Credit and College Access: Evidence from Texas Home Lending Laws. Working Paper.
Sulkin, S. E. (1946). The baffling problem of polio. *Southwest Review, 31*(4), 381-384.
U.S. Census Bureau. (1942). Sixteenth Decennial Census of the United States, 1940 Agriculture, Volume II.
U.S. Census Bureau. (1943). U.S. Census of Population and Housing, 1940: General Characteristics, Tenure and Rent, Income and Rent, and Characteristics of Rural-Farm Families. U.S. Government Printing Office: Washington D.C.
U.S. Census Bureau. (1946). Vital Statistics of the United States: 1944. U.S. Government Printing Office: Washington D.C.
U.S. Census Bureau. (1952). Seventeenth Decennial Census of the United States, 1950. Characteristics of the Population. Volume II, Parts 2-50.
U.S. Census Bureau. (1953). U.S. Census of Housing, 1950. Volume 1, General Characteristics. Part 1, U.S. Summary. U.S. Government Printing Office: Washington D.C.
U.S. Census Bureau. (1962a). Eighteenth Decennial Census of the United States, 1960. Population, Volume I, Characteristics of the Population. U.S. Census Bureau. (1962b). Vital Statistics of the United States: 1960. U.S. Government Printing Office: Washington D.C.
U.S. Census Bureau. (1963). U.S. Census of Housing, 1960. Volume 1, States and Small Areas. Part 1, United States Summary. U.S. Government Printing Office: Washington D.C.
van Panhuis, W. G., Grefenstette, J., Jung, S. Y., Chok, N. S., Cross, A., Eng, H., Lee, B. Y., Zadorozhny, V., Brown, S., Cummings, D., & Burke, D. S. (2013). Contagious diseases in the United States from 1888 to the present. *NEJM*, *369*(22), 2152-2158.
Webber, D. A. (2016). Are college costs worth it? How ability, major, and debt affect the returns to schooling. *Economics of Education Review*, *53*, 296-310.

APPENDICES

Appendix A

Table 26. Base Results with Differing Incidence Measures – No States Dropped Regardless of Number of State-Weeks Available.

Polio Incidence: Average Weekly Incidence In Twelve Month Period			
	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.019	-0.016	-0.01
	(0.02)	(0.02)	(0.02)
R^2	0.14	0.33	0.34
N	125,716	125,716	125,716
Polio Incidence: Total Incidence In Twelve Month Period			
	(4)	(5)	(6)
Polio Incidence Per 100,000	0.00	0.00	0.00
	(0.00)	0.00	0.00
R^2	0.14	0.33	0.34
N	125,716	125,716	125,716
Polio Incidence: Maximum Incidence In Twelve Month Period			
	(7)	(8)	(9)
Polio Incidence Per 100,000	0	0.001	0.001
	(0.00)	(0.00)	(0.00)
R^2	0.14	0.33	0.34
N	125,716	125,716	125,716
Polio Incidence: Minimum Incidence In Twelve Month Period			
	(10)	(11)	(12)
Polio Incidence Per 100,000	-0.17	-0.152	-0.071
	(0.16)	(0.14)	(0.13)
R^2	0.14	0.33	0.34
N	125,716	125,716	125,716
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 27. Main Results with Probabilistic Regressions.

	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.058**	-0.055***	-0.048**
	(0.03)	(0.02)	(0.02)
Log Pseudolikelihood	-5,435,076.45	-4,337,607.97	-4,270,197.76
N	95,839	95,839	95,839
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 28. Base Results with Varying Thresholds for the Number of State-Week Observations Necessary to be Included in the Regression.

Missing Weeks: 10 Weeks (19.2%)			
	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.058**	-0.051**	-0.031
	(0.024)	(0.020)	(0.020)
R^2	0.10	0.36	0.37
N	53,303	53,303	53,303
Missing Weeks: 18 Weeks (34.6%)			
	(4)	(5)	(6)
Polio Incidence Per 100,000	-0.053*	-0.056**	-0.045*
	(0.03)	(0.03)	(0.03)
R^2	0.15	0.36	0.37
N	77,261	77,261	77,261
Missing Weeks: 26 Weeks (50%)			
	(7)	(8)	(9)
Polio Incidence Per 100,000	-0.013	-0.01	-0.005
	(0.02)	(0.02)	(0.02)
R^2	0.14	0.33	0.34
N	120,476	120,476	120,476
Missing Weeks: 39 Weeks (75%)			
	(10)	(11)	(12)
Polio Incidence Per 100,000	-0.019	-0.015	-0.009
	(0.02)	(0.02)	(0.02)
R^2	0.14	0.33	0.34
N	125,027	125,027	125,027
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 29. Main Regression with All Coefficients Except State and Year Fixed Effects.

	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.041	-0.041*	-0.035
	(0.03)	(0.02)	(0.02)
Median Years of Schooling (State-Year)	-0.021***	0.011***	-0.011***
	(0.01)	(0.00)	(0.00)
Female	0.014***	0.444***	0.011***
	(0.00)	(0.02)	(0.00)
Age	0	0.025*	0.445***
	(0.00)	(0.01)	(0.02)
Black	-0.002**	-0.039	0.038***
	(0.00)	(0.03)	(0.01)
Other Race		-0.015**	-0.02
		(0.01)	(0.03)
Occupational Score of Head		0.011***	0.002***
		(0.00)	0.00
Female Head		0	0.041***
		(0.00)	(0.01)
Head Employed in Agriculture		-0.001*	-0.041***
		(0.00)	(0.01)
Both Parents Foreign			0.012
			(0.01)
Number of Siblings			-0.006***
			(0.00)
Rural			-0.083***
			(0.01)
Birth Rate in t (State-Year)			-0.014**
			(0.01)
Birth Rate in t-5 (State-Year)			0.011**
			(0.00)
Birth Rate in t-6 (State-Year)			-0.001
			(0.00)
Percent of Homes with Running Water (State-Year)			-0.002***
			(0.00)
Mother Present in Household			0.013
			(0.01)
Father Present in Household			0.024***
			(0.01)
R^2	0.14	0.34	0.35
N	95,839	95,839	95,839

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 30. No Epidemiological Controls.

	(1)	(2)	(3)
Polio Incidence Per 100,000	-0.078***	-0.070***	-0.081***
	(0.024)	(0.021)	(0.020)
R^2	0.14	0.34	0.35
N	96,834	96,834	96,834
State & Year Fixed Effects	x	X	x
Epidemiological Controls	x	X	x
Demographic Characteristics		X	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 31. No State Fixed Effects.

	(1)	(2)	(3)
Polio Incidence Per 100,000	0.121	0.126	0.123
	(0.088)	(0.088)	(0.077)
R^2	0.09	0.29	0.31
N	95,839	95,839	95,839
State & Year Fixed Effects	x	x	x
Epidemiological Controls	x	x	x
Demographic Characteristics		x	x
Head of Household & Family Characteristics			x
Median Years of School Control			x
Family Characteristics			x

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix B

Table 32. Variables' Descriptions and Creation.

Variable	Description & Creation
Any Money Problems in 1996	=1 if individual member of a family in 1996 where they (1) were unable to pay bills, (2) obtained a loan to pay off debts, (3) creditors called, (4) wages garnished, (5) had a lien on property, or (6) property repossessed
Whether Anyone in Family has Student Loans	If a family has any student loans. From Main Family data.
Amount of Student Loans in Family in Tens of Thousands of 2015 Dollars	Real family student loans in tens of thousands. Transformed from dollars to tens of thousands of dollars and in real 2015 dollars.
First entrance year of a parent's latest entering child	First entrance year of a parent's last college-entering child's college entrance. Uses both Transition to Adulthood (TA) data + Main Family (MF) data. *TA* - college enrollment year, uses Enrollment Year Recent College unless Enrollment Year Prior College is both (a) available and (b) before Enrollment Year Recent College. *MF* - imputed variable for Heads & Spouses; created by subtracting Highest Year In College from Year Last Attended College. If both TA and MF data exist for an individual, the earlier year of the two is used; if earliest year the same for both, MF data is used. For final analysis, we use the LAST entrance year a parent is exposed to as our reference point (ie, if a parent's latest college entrance of a child is 2013, the entrance year included in the analysis is 2013).
Earliest Wealth	Wealth with equity from the earliest data available for a parent. In order for it to be pulled, (a) the year with earliest data availble has to be at least 10 years before the last child's earliest college entrance year, and (b) the parent has to be under 40 for the earliest year of data available. From Main Family File.
Gender (Female = 1)	If this is a mother (ie, female). Parent/Child pairs created using Individual Survey's "1968 ID of Mother" (or Father) and "Person # of Mother" (or Father). If parent-child pair matched on ID of mother, assume individual female/mother. If parent-chld pair matched on ID of father, assume individual male/father.
Usual Employment Status: Retirement Age (Others)	Individual's own Usual Retirement Age of Others at same firm/with same job if individual Head/Spouse. From Main Family File.

Table 32. Continued.

Variable	Description & Creation
Expected Employment Status: Retirement Age (Self)	Individual's own Age Plan to Stop working if "Actual age" given and if individual Head/Spouse. Missing if "Never Retire." From Main Family File.
Never Expects to Retire	Created from Age Plan to Stop Working for Head/Spouse. =1 if individual responds "never" for age plan to stop working; =0 if actual age given. From Main Family File.
Retire Gap	Distance between the usual retirement age at a firm ("*Usual Employment Status: Retirement Age* ") and expected retirement for individual ("*Expected Employment Status: Retirement Age* "). Postive if individual expects to retire before the usual retirement age at a firm; negative if the individual expects to retire after the usual retirement age at a firm. Ex: -6 indicates an individual believes they'll retire 6 years after the typical retirement age at their firm.
Employment Status: Retired	If Individual Retired. From Individual Employment Status. IndEmp =1 if Employment Status =4(retired); IndEmp =0 if Employment Status =1 (Working Now), =2(Only temporarily laid off), =3(looking for work, unemployed), =5(permanently disabled), =6(HouseSpouse), =7(Student) =8(Other). From Individual Data File.
Employment Status: Not in the Labor Force	If Individual Not in the Labor Force. From Individual Employment Status. =1 if Employment Status =4(retired), =5(permanently disabled), =6(HouseSpouse), =7(Student). =0 if Employment Status =1(working now), =2(only temporarily laid off), = 3(looking for work, unemployed) =8(Other). From Individual Data File.
Employment Status: Unemployed	If Individual Unemployed. From Individual Employment Status. IndEmp =1 if Employment Status =3(looking for work, unemployed) ; IndEmp =0 if Employment Status =1 (Working Now), =2(Only temporarily laid off), =4(retired), =5(permanently disabled), =6(HouseSpouse), =7(Student) =8(Other). From Individual Data File.
Employment Status: Employed	From Individual Employment Status. IndEmp =1 if Employment Status =1 (Working Now); IndEmp =0 if Unemployed = 1 or Not in Labor Force = 1. From Individual Data File.

Table 32. Continued.

Variable	Description & Creation
Real 2015 Dollars in IRA/Annuities	From Imputed Value in Annuity/IRA. Made "real IRA Value" by Inflation-adjustment to 2015 dollars. From Main Family File.
Whether Anyone in Family has IRA/Annuities	=1 if Imputed Value IRA > 0; =0 if Imputed Value IRA = 0. From Main Family File.
Age	Individual's age. From Individual Data File.
Social Security Eligibility	=1 if individual's age >= 62; =0 otherwise.
Spouse Present	Spouse Present, derived from Couple Status of Head. Spouse Present=1 if Couple Status of Head = 1 (Head with wife present in FU), =2 (Head with partner present in FU), = 3 (Head (female) with husband present in FU), = 4 (Head with first-year cohabitor present in FU). SpsPrsnt = 0 if Couple Status of Head = 5(Head with no wife, partner, husband, or first-year cohabitor present in FU). From Main Family File.
Real Assets in 2015 $s	Real Imputed Assets. Created by adding "imputed wealth with home equity" and "imputed debts" together. (ie, undoing the transformation of wealth = assets - debt). "Imputed debts" is from the Main Family file; in 2011, 2013, and 2015, when debts were separated out into categories, we add those categories together to create a total as in prior PSID surveys (for 2011, imputed debts = credit card + student loan + medical debt + legal debt + loans from family; in 2013 and 2015, imputed debts = credit card + student loan + medical debt + legal debt + loans from family + farm debt + real estate debt + other debt). Transformed into real 2015 dollars.
Financial Literacy (Scale = 0-6)	Number out of 6 questions about financial literacy that an individual got correct. From 2016 Well-Being Survey.
Mother's Education	Mother's latest education available from survey, categorical (1 = 0-5 grades, 2 = 6-8 grades, 3 = 9-11 grades, 4 = 12 grades, 5 = 12 grades & some nonacademic training, 6 = some college, 7 = BA, 8 = advanced degree)
Health Status	Individual's own Health Status if individual Head/Spouse.

Table 32. Continued.

Variable	Description & Creation
Household Income Less Own Labor Income in 2015 $s	Real household income less individual parent's own labor income. "Household income" is "Total Family Income [previous year]" from Main Family data. Restrict this to positive values only. Because we want Total Family Income in survey years, if the referenced year is between 1996 and 2014, we impute the survey year's Total Famly Income as (1/2)Previous Year's Total Family Income + (1/2)Next Year's Total Family Income. (This means we lose 2015, for which there is no next year of data available.) "Own labor income" is either "Labor Income [2 Years Previous] for Head/Spouse" (2015 - 2003) or "Labor Income [1 Year Previous] for Head/Spouse" (2001 - 1983). This leaves 1999 and 1997 with labor income information because in 2001 and 1999, information was collected for 2000 and 1998, respectfully. Impute these year's as LaborIncome in Current Year = (1/2) Labor Income in Previous Year + (1/2)Labor Income Next Year. (ie, 1999's labor income is (1/2) 2000's labor income and (1/2) 1998's labor income). Subtract individual's own labor income from household income; transform into real 2015 dollars. Because 2015 is unavailable, 2013's is used for individual's whose reference year is 2015.
Number of Children Matched in College	Number of children in college a parent was successfully matched with in the data.
White	Individual's self-identified race for Heads/Spouses, from Main Family Survey. White =1 if Race of Head/Spouse Mention #1 = 1(White). White = 0 if Race of Head/Spouse Mention #1 = 2, 3, 4, 5, 6, or 7.
Black	Individual's self-identified race for Heads/Spouses, from Main Family Survey. Black =1 if Race of Head/Spouse Mention #1 = 2(Black, African-American, or Negro). Black = 0 if Race of Head/Spouse Mention #1 = 1, 3, 4, 5, 6, or 7.
Other Race	Individual's self-identified race for Heads/Spouses, from Main Family Survey. Other Race =1 if Race of Head/Spouse Mention #1 = 3(American Indian or Alaska Native), =4 (Asian), =5 (Native Hawaiian or Pacific Islander), =7 (Other). Other Race = 0 if Race of Head/Spouse Mention #1 = 1 or 2.

Table 32. Continued.

Variable	Description & Creation
Total Years of College for All Children	Total years of college all of a parent's children attended; total of a variable for each child called Years in College. For each child, if "First College Entrance Year" used is from Main Family data, Years In College is Main Family's " Highest Year in College for Head/Spouse" (if child Head/Spouse). If "First College Entrance Year" from Transition to Adulthood data, from TA's "Grade Level Completed" (GLC); "Years In College" =1 if GLC = 14 (One Year college); = 2 if GLC =15 (two years college); = 3 if GLC = 15 (three years college); = 4 if GLC = 16; = 5 if GLC = 17. If "Grade Level Completed" not available for individual where "First College Entrance Year" is from TA data, then a proxy created from TA data of "First College Entrance Year" - 2013 (if individual currently attending school in 2013) or "First College Entrance Year" - 2011 (if individual data not available for 2013 and individual currently attending school in 2011). This proxy is topcoded at 5 so it matches the other data used to create Years In College.

Appendix C

Table 33. The Effect of Any Student Loan Debt on Retirement Expectations, All Coefficients Except State and Year Fixed Effects Included.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Any Student Loans	-0.433	1.13	0.888	0.024
	(0.81)	(0.78)	(0.58)	(0.04)
Mother	0.9	-0.271	-0.541	-0.035
	(0.74)	(0.72)	(0.50)	(0.04)
Black	1.184	-1.765	-0.801	-0.04
	(1.15)	(1.22)	(0.71)	(0.06)
Other Race	-1.064	1.296	-0.645	0.068
	(1.07)	(1.45)	(0.68)	(0.06)
Hispanic	0.541	-1.323	-0.684	-0.072
	(1.21)	(1.25)	(0.77)	(0.05)
Household Income Less Own Labor Income	-0.045	-0.082	-0.057**	-0.002
	(0.06)	(0.05)	(0.03)	(0.00)
Age	0.105	0.169**	0.215***	-0.001
	(0.07)	(0.08)	(0.06)	-0.004
Spouse Present	-0.08	-1.444	-1.809***	-0.045
	(0.94)	(0.97)	(0.61)	(0.05)
Assets in 2015	-0.008***	0.006*	-0.002	0
	(0.00)	(0.00)	(0.00)	0.00
Social Security Eligibility	1.617*	0.221	1.502**	0.098*
	(0.96)	(0.96)	(0.67)	(0.06)
Number of Children in College	0.425	0.132	1.096**	0.018
	(0.83)	(0.78)	(0.49)	(0.04)
Some Children Went to HBCU	-4.192***	3.606*	-1.084	-0.181**
	(1.57)	(1.91)	(1.53)	(0.08)
All Children Attended HBCU	-1.349	1.046	0.112	-0.107
	(1.36)	(1.60)	(1.24)	(0.08)

Table 33. Continued.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Some Children Attended Private College	0.04	-1.976*	-1.011	0.528***
	(0.96)	(1.03)	(1.35)	(0.10)
All Children Attended Private College	-0.527	1.838**	1.134*	0.084
	(0.75)	(0.82)	(0.67)	(0.05)
Some Children Attended For-Profit College	0.361	0.533	1.002	0.383***
	(2.72)	(2.32)	(1.11)	(0.09)
All Children Attended For-Profit College	-2.304*	0.663	-1.302*	0.149*
	(1.29)	(1.21)	(0.72)	(0.09)
Some Children Attended Four-Year College	-0.895	0.555	-1.317*	-0.06
	(1.37)	(1.18)	(0.74)	(0.06)
All Children Attended Four-Year College	-0.023	0.115	-0.444	0.002
	(0.90)	(0.82)	(0.54)	(0.04)
Total Years Children in College	-0.082	0.105	0.032	-0.003
	(0.15)	(0.15)	(0.10)	(0.01)
Health Status: Very Good	0.367	-0.082	1.072*	-0.025
	(0.89)	(0.96)	(0.63)	(0.05)
Health Status: Good	-0.476	-0.101	0.185	-0.066
	(0.99)	(1.02)	(0.66)	(0.04)
Health Status: Fair	0.469	-0.461	-0.182	-0.006
	(1.16)	(1.09)	(0.75)	(0.06)

Table 33. Continued.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Health Status: Poor	-1.868	0.664	1.128	0.151
	(2.02)	(1.56)	(1.37)	(0.15)
R^2	0.27	0.31	0.37	0.18
N	401	418	691	761

Table 34. The Effect of Any Student Loan Debt on Employment Status, All Coefficients Except State and Year Fixed Effects Included.

	Retired	Not in Labor Force	Unemployed	Employed
Any Student Loans	-0.035	-0.054*	-0.016	0.055*
	(0.02)	(0.03)	(0.02)	(0.03)
Mother	0.027	0.070*	0.018	-0.063
	(0.03)	(0.04)	(0.02)	(0.04)
Black	-0.023	-0.014	-0.02	0.015
	(0.04)	(0.05)	(0.02)	(0.05)
Other Race	-0.146***	-0.123**	0.013	0.111**
	(0.04)	(0.05)	(0.03)	(0.05)
Hispanic	0.026	0.022	0	-0.023
	(0.05)	(0.06)	(0.02)	(0.06)
Household Income Less Own Labor Income	0.003	0.003	0	-0.003
	(0.00)	(0.00)	(0.00)	(0.00)
Age	0.010***	0.006*	-0.001	-0.005
	(0.00)	(0.00)	(0.00)	(0.00)
Spouse Present	-0.04	-0.071*	-0.003	0.076*
	(0.04)	(0.04)	(0.02)	(0.04)
Assets in 2015	0	0	0	0
	0.00	0.00	0.00	0.00
Social Security Eligibility	0.112**	0.121**	0	-0.133***
	(0.05)	(0.05)	(0.02)	(0.05)
Number of Children in College	0	-0.012	-0.007	0.012
	(0.03)	(0.03)	(0.01)	(0.03)
Some Children Went to HBCU	-0.044	-0.115	0.018	0.113
	(0.10)	(0.13)	(0.03)	(0.12)
All Children Attended HBCU	-0.011	-0.028	0.019	0.024
	(0.06)	(0.08)	(0.03)	(0.08)

Table 34. Continued.

	Retired	Not in Labor Force	Unemployed	Employed
Some Children Attended Private College	-0.154*	-0.144	0.055	0.131
	(0.08)	(0.11)	(0.05)	(0.11)
All Children Attended Private College	-0.064*	-0.082*	-0.034**	0.082*
	(0.04)	(0.04)	(0.02)	(0.04)
Some Children Attended For-Profit College	-0.122*	0.046	0.053	-0.049
	(0.07)	(0.10)	(0.05)	(0.10)
All Children Attended For-Profit College	0.077	0.061	-0.014	-0.055
	(0.07)	(0.09)	(0.03)	(0.09)
Some Children Attended Four-Year College	-0.022	-0.104	-0.047**	0.110*
	(0.05)	(0.07)	(0.02)	(0.07)
All Children Attended Four-Year College	0.044	0.024	-0.015	-0.015
	(0.03)	(0.04)	(0.02)	(0.04)
Total Years Children in College	0.007	0.006	0	-0.005
	(0.01)	(0.01)	(0.00)	(0.01)
Health Status: Very Good	-0.055	-0.059	-0.003	0.062
	(0.04)	(0.04)	(0.01)	(0.04)
Health Status: Good	0.018	0.058	0.027	-0.056
	(0.04)	(0.05)	(0.02)	(0.05)

Table 34. Continued.

	Retired	Not in Labor Force	Unemployed	Employed
Health Status: Fair	0.049	0.161**	0.011	-0.170**
	(0.06)	(0.07)	(0.02)	(0.07)
Health Status: Poor	-0.013	0.159	0.099	-0.194
	(0.09)	(0.12)	(0.07)	(0.12)
R^2	0.37	0.3	0.14	0.3
N	771	771	771	771

Table 35. The Effect of Any Student Loan Debt on Retirement Savings, All Coefficients Except State and Year Fixed Effects Included.

	Real 2015 Dollars in IRA/Annuities	Whether Anyone in Family has IRA/Annuities
Any Student Loans	-7.004**	-0.074
	(3.03)	(0.06)
Mother	0.611	-0.027
	(3.43)	(0.05)
Black	-4.649*	-0.129*
	(2.79)	(0.07)
Other Race	-2.133	-0.017
	(5.71)	(0.07)
Hispanic	-8.723*	-0.253***
	(4.52)	(0.07)
Household Income Less Own Labor Income	0.598*	0.011***
	(0.34)	(0.00)
Age	-0.138	0.013***
	(0.30)	(0.01)
Spouse Present	0.897	-0.001
	(2.58)	(0.05)
Assets in 2015 (Less IRA/Annuities)	0.067***	0
	(0.02)	0.00
Social Security Eligibility	10.603*	0.047
	(5.57)	(0.07)
Number of Children in College	-2.32	-0.105**
	(3.52)	(0.05)
Some Children Went to HBCU	-8.928	-0.210**
	(6.99)	(0.11)
All Children Attended HBCU	-1.572	-0.129
	(5.12)	(0.10)
Some Children Attended Private College	-3.601	-0.286**
	(8.90)	(0.12)
All Children Attended Private College	3.681	0.068
	(4.87)	(0.06)

Table 35. Continued.

	Real 2015 Dollars in IRA/Annuities	Whether Anyone in Family has IRA/Annuities
Some Children Attended For-Profit College	-13.350*	-0.336***
	(7.14)	(0.10)
All Children Attended For-Profit College	1.456	-0.036
	(3.98)	(0.08)
Some Children Attended Four-Year College	0.317	0.096
	(4.33)	(0.07)
All Children Attended Four-Year College	4.522	0.059
	(3.39)	(0.06)
Total Years Children in College	1.357**	0.025***
	(0.54)	(0.01)
Health Status: Very Good	5.125	-0.029
	(4.92)	(0.06)
Health Status: Good	-1.258	-0.107*
	(4.88)	(0.06)
Health Status: Fair	-3.851	-0.113
	(4.87)	(0.07)
Health Status: Poor	-12.030*	-0.209*
	(6.25)	(0.11)
R2	0.37	0.35
N	767	767

Table 36. The Effect of the Amount of Student Loan Debt on Retirement Expectations, All Coefficients Except State and Year Fixed Effects Included.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Student Loans (Tens of Thousands of 2015 $'s)	-0.041	0.159	0.376**	0.001
	(0.15)	(0.17)	(0.16)	(0.01)
Mother	0.881	-0.214	-0.532	-0.039
	(0.73)	(0.72)	(0.50)	(0.04)
Black	1.175	-1.717	-0.572	-0.039
	(1.14)	(1.20)	(0.71)	(0.06)
Other Race	-0.957	1.166	-0.572	0.075
	(1.14)	(1.54)	(0.70)	(0.07)
Hispanic	0.606	-1.396	-0.65	-0.064
	(1.28)	(1.34)	(0.77)	(0.05)
Household Income Less Own Labor Income	-0.046	-0.079	-0.053*	-0.002
	(0.06)	(0.05)	(0.03)	(0.00)
Age	0.107	0.167**	0.212***	-0.001
	(0.07)	(0.08)	(0.06)	(0.00)
Spouse Present	-0.14	-1.375	-1.868***	-0.05
	(0.95)	(0.97)	(0.63)	(0.05)
Assets in 2015	-0.008**	0.006*	-0.002	0
	(0.00)	(0.00)	(0.00)	0.00
Social Security Eligibility	1.58	0.288	1.557**	0.100*
	(0.97)	(0.98)	(0.65)	(0.06)
Number of Children in College	0.442	0.143	1.139**	0.012
	(0.84)	(0.78)	(0.48)	(0.04)
Some Children Went to HBCU	-4.099***	3.340*	-1.334	-0.176**
	(1.52)	(1.92)	(1.52)	(0.09)
All Children Attended HBCU	-1.354	1.043	0.028	-0.112
	(1.37)	(1.56)	(1.20)	(0.08)

Table 36. Continued.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Some Children Attended Private College	-0.001	-1.890*	-1.054	0.522***
	(0.97)	(1.02)	(1.35)	(0.10)
All Children Attended Private College	-0.449	1.720**	1.242*	0.092*
	(0.76)	(0.83)	(0.68)	(0.05)
Some Children Attended For-Profit College	0.3	0.575	1.118	0.387***
	(2.76)	(2.29)	(1.10)	(0.09)
All Children Attended For-Profit College	-2.260*	0.552	-1.371*	0.150*
	(1.30)	(1.26)	(0.76)	(0.09)
Some Children Attended Four-Year College	-0.854	0.436	-1.463**	-0.055
	(1.36)	(1.17)	(0.72)	(0.06)
Some Children Attended Four-Year College	-0.854	0.436	-1.463**	-0.055
	(1.36)	(1.17)	(0.72)	(0.06)
All Children Attended Four-Year College	-0.036	0.132	-0.578	0.005
	(0.90)	(0.83)	(0.52)	(0.04)
Total Years Children in College	-0.078	0.093	0.021	-0.003
	(0.15)	(0.15)	(0.10)	(0.01)

Table 36. Continued.

	Retirement Gap	Others' Usual Retirement Age for Same Job	Age Expected to Retire (Self)	Never Expects to Retire
Health Status: Very Good	0.364	-0.075	1.178*	-0.026
	(0.89)	(0.96)	(0.63)	(0.05)
Health Status: Good	-0.503	-0.021	0.236	-0.063
	(1.00)	(1.02)	(0.65)	(0.05)
Health Status: Fair	0.471	-0.519	0.038	0.003
	(1.16)	(1.11)	(0.74)	(0.06)
Health Status: Poor	-1.813	0.518	1.286	0.148
	(2.01)	(1.49)	(1.45)	(0.15)
R^2	0.27	0.31	0.39	0.19
N	400	417	687	756

Table 37. The Effect of the Amount of Student Loan Debt on Retirement Expectations, All Coefficients Except State and Year Fixed Effects Included.

	Retired	Not in Labor Force	Unemployed	Employed
Student Loans (Tens of Thousands of 2015 $'s)	-0.004	-0.005	-0.001	0.005
	(0.00)	(0.01)	(0.00)	(0.01)
Mother	0.026	0.070*	0.018	-0.063
	(0.03)	(0.04)	(0.02)	(0.04)
Black	-0.025	-0.017	-0.021	0.018
	(0.04)	(0.05)	(0.02)	(0.05)
Other Race	-0.149***	-0.122**	0.014	0.110**
	(0.04)	(0.05)	(0.03)	(0.05)
Hispanic	0.026	0.022	0	-0.023
	(0.05)	(0.06)	(0.03)	(0.06)
Household Income Less Own Labor Income	0.003	0.003	0	-0.003
	(0.00)	(0.00)	(0.00)	(0.00)
Age	0.010***	0.007*	-0.001	-0.006
	(0.00)	(0.00)	(0.00)	(0.00)
Spouse Present	-0.04	-0.078*	-0.004	0.082*
	(0.04)	(0.04)	(0.02)	(0.04)
Assets in 2015	0	0	0	0
	0.00	0.00	0.00	0.00
Social Security Eligibility	0.110**	0.119**	0	-0.130**
	(0.05)	(0.05)	(0.02)	(0.05)
Number of Children in College	-0.004	-0.011	-0.008	0.011
	(0.03)	(0.03)	(0.01)	(0.03)
Some Children Went to HBCU	-0.037	-0.113	0.021	0.11
	(0.10)	(0.12)	(0.03)	(0.12)
All Children Attended HBCU	-0.015	-0.036	0.016	0.032
	(0.06)	(0.08)	(0.03)	(0.08)
Some Children Attended Private College	-0.155*	-0.142	0.055	0.129
	(0.08)	(0.11)	(0.05)	(0.11)

Table 37. Continued.

	Retired	Not in Labor Force	Unemployed	Employed
All Children Attended Private College	-0.066*	-0.076*	-0.033*	0.075*
	(0.04)	(0.04)	(0.02)	(0.04)
Some Children Attended For-Profit College	-0.118*	0.06	0.056	-0.063
	(0.07)	(0.10)	(0.05)	(0.10)
All Children Attended For-Profit College	0.082	0.075	-0.014	-0.069
	(0.07)	(0.09)	(0.03)	(0.09)
Some Children Attended Four-Year College	-0.02	-0.101	-0.046*	0.107
	(0.05)	(0.07)	(0.02)	(0.07)
All Children Attended Four-Year College	0.044	0.023	-0.015	-0.014
	(0.03)	(0.04)	(0.02)	(0.04)
Total Years Children in College	0.008	0.006	0	-0.005
	(0.01)	(0.01)	(0.00)	(0.01)
Health Status: Very Good	-0.056	-0.062	-0.003	0.064
	(0.04)	(0.04)	(0.01)	(0.04)
Health Status: Good	0.016	0.051	0.027	-0.049
	(0.04)	(0.05)	(0.02)	(0.05)
Health Status: Fair	0.048	0.160**	0.012	-0.169**
	(0.06)	(0.07)	(0.02)	(0.07)
Health Status: Poor	-0.014	0.153	0.099	-0.188
	(0.09)	(0.12)	(0.07)	(0.12)
R^2	0.37	0.3	0.14	0.3
N	763	763	763	763

Table 38. The Effect of the Amount of Student Loan Debt on Retirement Savings, All Coefficients Except State and Year Fixed Effects Included.

	Real 2015 Dollars in IRA/Annuities	Whether Anyone in Family has IRA/Annuities
Student Loans (Tens of Thousands of 2015 $'s)	-1.411**	-0.013
	(0.59)	(0.01)
Mother	0.452	-0.022
	(3.49)	(0.05)
Black	-5.573**	-0.135**
	(2.79)	(0.07)
Other Race	-2.264	-0.02
	(5.81)	(0.07)
Hispanic	-9.301**	-0.258***
	(4.64)	(0.07)
Household Income Less Own Labor Income	0.603*	0.011***
	(0.34)	(0.00)
Age	-0.156	0.014***
	(0.30)	(0.01)
Spouse Present	0.936	0.002
	(2.61)	(0.06)
Assets in 2015 (Less IRA/Annuities)	0.068***	0
	(0.02)	0.00
Social Security Eligibility	10.682*	0.047
	(5.55)	(0.07)
Number of Children in College	-2.213	-0.101**
	(3.66)	(0.05)
Some Children Attended HBCU	-7.844	-0.202*
	(6.87)	(0.10)
All Children Attended HBCU	-2.056	-0.128
	(4.98)	(0.10)
Some Children Attended Private College	-2.922	-0.282**
	(8.87)	(0.12)
All Children Attended Private College	3.61	0.056
	(4.95)	(0.06)

Table 38. Continued.

	Real 2015 Dollars in IRA/Annuities	Whether Anyone in Family has IRA/Annuities
Some Children Attended For-Profit College	-12.815*	-0.339***
	(7.11)	(0.10)
All Children Attended For-Profit College	1.503	-0.036
	(3.89)	(0.08)
Some Children Attended Four-Year College	0.65	0.098
	(4.31)	(0.07)
All Children Attended Four-Year College	4.876	0.06
	(3.40)	(0.06)
Total Years Children in College	1.315**	0.025***
	(0.55)	(0.01)
Health Status: Very Good	5.104	-0.029
	(4.90)	(0.06)
Health Status: Good	-1.264	-0.108*
	(4.87)	(0.06)
Health Status: Fair	-3.832	-0.126*
	(4.92)	(0.07)
Health Status: Poor	-12.120*	-0.207*
	(6.28)	(0.11)
R^2	0.37	0.35
N	759	759

Appendix D

Table 39. Effect of Lottery Scholarship Availability on Debt; College Exit Year Less Years College Completed Used as College Entrance Year, All Coefficients Except State and Year Fixed Effects.

	Any Other Debt	Amount Real Other Debt (2015)	Log of Amount Real Other Debt (2015) (=0 if Amount of Real Other Debt =0)
Scholarship Available	-0.00666	-1,801	-0.174
	(0.07)	(3447.00)	(0.70)
Distance between College Exit and Survey	0.0104	308.9	0.135
	(0.01)	(576.80)	(0.13)
Father's Education: High School Graduate, No Bachelor's Degree	0.0447	-1,398	0.313
	(0.04)	(1996.00)	(0.38)
Father's Education: College Graduate or Higher	-0.0123	-2,503	-0.392
	(0.04)	(2311.00)	(0.41)
Race: Black	-0.0105	476.7	0.346
	(0.04)	(2092.00)	(0.37)
Race: Other Race	-0.216**	-5,217**	-1.799**
	(0.09)	(2501.00)	(0.80)
Age	0.0017	174.6	0.0205
	(0.00)	(108.10)	(0.02)
Female	0.0158	887.7	0.414*
	(0.03)	(1250.00)	(0.24)
Household Years in College	0.00972*	2,824***	0.301***
	(0.01)	(300.40)	(0.05)
Constant	0.418**	-7,360	2.504
	(0.16)	(6251.00)	(1.61)
Observations	1,735	1,735	1,735
R-squared	0.114	0.17	0.097

Table 40. Effect of Lottery Scholarship Availability on Debt; High School Graduation Year Used as College Entrance Year, All Coefficients Except State and Year Fixed Effects.

	Any Other Debt	Amount Real Other Debt (2015)	Log of Amount Real Other Debt (2015) (=0 if Amount of Real Other Debt =0)
Scholarship Available	-0.0811	-4,920	-0.795
	(0.10)	(4474.00)	(0.99)
Distance between College Exit and Survey	-0.00593	1,112	0.17
	(0.02)	(798.50)	(0.17)
Father's Education: High School Graduate, No Bachelor's Degree	0.0208	-579	0.307
	(0.06)	(2827.00)	(0.52)
Father's Education: College Graduate or Higher	-0.027	-2,552	-0.422
	(0.06)	(3163.00)	(0.53)
Race: Black	-0.0532	747.5	0.0892
	(0.05)	(2666.00)	(0.49)
Race: Other Race	-0.235**	-7,659***	-1.554*
	(0.10)	(2177.00)	(0.92)
Age	-0.00558	1,331***	0.117***
	(0.01)	(282.60)	(0.04)
Female	-0.0153	-102.3	0.0647
	(0.03)	(1382.00)	(0.27)
Household Years in College	0.0159***	2,706***	0.327***
	(0.01)	(367.10)	(0.06)
Constant	0.797***	-42,330***	1.02
	(0.22)	(9839.00)	(2.10)
Observations	1,325	1,325	1,325
R-squared	0.14	0.198	0.127

VITA

Mary Elizabeth “Beth” Glenn was born to Tammy and James Glenn in Knoxville, Tennessee on February 12, 1990. She was raised in Easley, South Carolina and graduated from Westside High School in 2008. She attended Winthrop University and graduated magna cum laude in May 2013 with a Bachelor of Arts double-major in Economics and Psychology. She moved to Knoxville, Tennessee in August 2013 to begin her graduate studies at the University of Tennessee. She received her Master of Arts degree in Economics in January 2015 and her Doctor of Philosophy degree in Economics in August 2018. She worked as a graduate research assistant at the Boyd Center for Business and Economic Research for four years and contributed to reports for the Tennessee Valley Authority and the Tennessee Legislature. She was awarded the *Charles B. Garrison Award for Best Graduate Assistant* for her dedication to the center. She has co-authored two papers with professors at the Boyd Center, and she was awarded the *J. Fred and Wilma Holly Fellowship for Excellence in Research* for her second year research paper. This fall, she will move to New Orleans, Louisiana, where she accepted a position as a Postdoctoral Teaching and Research Fellow at the Education Research Alliance at Tulane University.