ESSAYS ON THE DETERMINANTS OF HEALTH AND LABOR MARKET OUTCOMES

A Dissertation

by

ADAM ROBERTS

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Andrew Barr
Committee Members,	Mark Hoekstra
	Steve Puller
	Laura Dague
Head of Department,	Steve Puller

May 2022

Major Subject: Economics

ABSTRACT

There is increasing evidence that an individual's experiences, setting, and resource access during early childhood and education years have a profound impact on productivity, income, and health in adulthood. Understanding the magnitude and direction of early life interventions has important policy implications for many targeted public spending programs. We contribute along several dimensions to better understanding the determinants of long run labor market and health outcomes.

In Chapter I, I leverage variation in the county level variation in the timing of fluoride adoption within a county to estimate the causal effect of childhood fluoride exposure. Despite being named one of the 10 greatest public health achievements of the 20th century, I find negative long run impacts as a result of community water fluoridation programs. Specifically, I find that children exposed to community water fluoridation from age zero to five experience a 1.9 percent of a standard deviation decrease in their adult economic self-sufficiency and a 1.2 percent of a standard deviation decrease in physical ability and health.

In Chapter II, we examine the effect of providing benefits in-kind versus in cash. We leverage a policy in Puerto Rico that converted cash benefits to in-kind nutritional assistance, holding benefit generosity constant. Using a difference-in-differences strategy, we find that providing the benefits in-kind led to significant increases in food consumption and decreases in maternal anemia, and resulted in long run improvements in adolescent health.

Shifting the focus to post-secondary education years, in Chapter III we estimate the causal impact of access to financial aid. We overcome endogeneity concerns by leveraging the elimination of the Social Security Student Benefit Program. We use a large sample of administrative Social Security records to precisely identify individuals impacted by the elimination of the program and link these individual records to Census Bureau survey data. Preliminary results show that eligible

students received an average of over $7,000 dollars in financial assistance which resulted in 0.185 increased years of education but no statistically significant causal impact on long run labor market or health outcomes.

DEDICATION

It was all for Jessi

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Andrew Barr (advisor) and Steve Puller and Mark Hoekstra of the Department of Economics as well as Professor Laura Dague of the Department of Public Service Administration.

Chapters I and III were made possible by restricted Census Bureau data. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under Project Numbers FSRDC 1888 and DMS 7505725.

All work conducted for the dissertation was completed by the student in collaboration with coauthors Andrew Barr, Alex Smith, and Jonathan Eggleston.

Funding Sources

Graduate study was supported by a fellowship from Texas A&M University as well as the CLLA Summer Dissertation Fellowship and a research fellowship from Private Enterprise Research Center. Chapter III was made possible in part by a grant from the Spencer Foundation (202100134).

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

1. INTRODUCTION

Economists have always been interested in the determinants of individual productivity. There is increasing evidence that an individual's experiences, setting, and resource access during early childhood and educational years have a profound impact on the long run productivity, income, and health. A virtually limitless number of factors have the potential to affect long run outcomes, including genetic predispositions, pollution, access to education, access to resources, nutrition, exposure to natural disasters, and more. Understanding the magnitude and direction of these early life interventions has important policy implications for many targeted public spending programs as well as for individuals' private investment decisions in themselves and their children. Each chapter addresses the impact of an additional factor in long run U.S. health and labor market outcomes, specifically covering the impact of early childhood exposure to water fluoridation, cash vs in-kind nutritional assistance, and the impact of post-secondary financial aid. Each of these chapters has significant policy implications which will be discussed in more detail in the individual chapters and the conclusion.

In Chapter I, I leverage variation in the county level variation in the timing of fluoride adoption within a to estimate the causal effect of childhood fluoride exposure. Over 70% of publicly supplied drinking water in the United States is fluoridated and the CDC has named community water fluoridation as one of the 10 greatest public health achievements of the 20th century (Gooch, 2020). Despite strong evidence that exposure to low levels of fluoride is an effective way to strengthen teeth, many individuals, communities, and industrialized countries oppose water fluoridation out of concern for potential negative health risks. The impact of fluoride on health varies based on both the amount and timing of fluoride exposure. This is true for both the positive impacts on dental health as well as the potential negative side effects on teeth, bones, and cognitive function. While it is well established that fluoride exposure makes teeth more resistant to decay, recent controversy has focused on the role of fluoride as a neurotoxin. The meta-study Choi, Zhang and Grandjean (2012) concludes that early exposure to high fluoride levels results in decreased cognitive functions equivalent to nearly one half of a standard deviation in IQ scores. While the majority of reviewed studies focus on subjects with fluoride levels well above recommended levels, some find negative cognitive effects at relative low levels as well. In a follow up meta-study incorporating more recent evidence, Grandjean (2019) concluded that safe exposure levels are likely below the levels currently recommended for water fluoridation.

I provide the first large sample and quasi-experimental evidence of the long run health and labor market effects of community water fluoridation programs. The restricted Census datasets I use allow for vast increases in precision, require much weaker identifying assumptions via inclusion of birth county fixed effects, and include a broad set of outcomes, birth cohorts, and communities relative to previous work.

I find that children exposed to community water fluoridation during early childhood experience a 1.9 percent of a standard deviation decrease in their adult economic self-sufficiency, 1.2 percent of a standard deviation increase in rates of adult disability, and a 1.5 percentage point decrease in high school graduation. These results show that the net effect of fluoride is negative even at relatively low levels of exposure. A net negative impact suggests that, even at levels previously thought to be safe, the tooth strengthening effect of fluoride provides less benefit than fluoride's corresponding health costs.

While access to safe drinking water is one necessary component of early childhood development and health, adequate nutrition through food access is equally important. A significant body of work has demonstrated the short- and long-term effects of nutritional assistance on consumption patterns, nutrition, crime, and health, showing that the availability of nutritional assistance has a meaningful effect on these long run outcomes.[1]

[1]See, for example, Gertler (2004); Fox, Hamilton and Lin (2004); Lee and Mackey-Bilaver (2006); Whitake, Philips and Orzol (2006); Baum (2007); Hoynes and Schanzenbach (2009); Almond, Hoynes and Schanzenbach (2011); Hoynes and Schanzenbach. (2012); Hoynes, Schanzenbach and Almond (2016); Hastings, Kessler and Shapiro

In Chapter II, we contribute to the literature on childhood nutritional assistance by focusing on the importance of the form of the benefits, rather than simple access. There is a long-standing debate over whether benefits should primarily be provided in cash, where the spending decision is left to the recipient, or in-kind, where much of the spending decision is made by the government. Nutritional assistance is a particularly interesting setting for investigating the tradeoffs between different benefit forms because it is a large program, providing in-kind benefits to more than 42 million Americans at a cost of more than $68 billion, in which the benefits are shared among the members of each recipient household (unlike Medicaid for example). In this context, policymakers may be concerned that, under a cash benefit system, those in the household that are not the "spending decision-maker" (e.g. children) may not receive the optimal level of nutrition. This deficit may yield substantial externalities if it translates into immediate or longer-term health effects.

In this chapter, we investigate whether the form of nutritional assistance affects household food consumption patterns and corresponding changes in health outcomes. We take advantage of a previously unexplored change in the Puerto Rican Nutritional Assistance Program (NAP) in 2001 that converted a cash benefit to one in which recipients were required to spend 75% of their benefits on approved food items. As the total value of the benefit remained fixed, this shift provides a unique opportunity to isolate the impact of providing food relative to an equivalent amount of cash in the context of a large-scale program. NAP's high participation rate, with more than one third of residents receiving assistance, make it an ideal program for observing the impact of this type of shift.

We use a difference-in-differences strategy and find that in-kind benefits resulted in dramatic nutritional improvements, with fruit and vegetable consumption increasing by 0.5 servings per day (20 percent). These nutritional improvements, in turn, yielded some contemporaneous improvements in health with maternal anemia falling by 0.3 to 0.7 pp (13 to 23 percent). Those who grew up after the policy change, and were therefore more likely to have had better nutrition in early

(2018); Barr and Smith (2018).

childhood, are taller and more likely to be normal weight as adolescents.

It is important to note that our results do not imply that in-kind benefits are necessarily welfare improving over cash. Rather, they suggest that unconstrained household spending decisions may not prioritize health, particularly of young children. To the extent that recipients of nutritional assistance are also likely to receive government-provided or subsidized healthcare (i.e. Medicaid or individual marketplace subsidies), constraining household decisions with in-kind benefits may yield a substantial fiscal externality for the government in the form of reductions in healthcare spending.

After these key early development periods where access to safe drinking water and good nutrition play such a vital role in long term outcomes, adolescence and early adult life is largely focused on human capital investments in education. Understanding the returns to higher education has relevant implications both for policymakers seeking to optimize resource allocation for education as well as individuals considering large investments in either their own or their children's education. While those with more education have higher incomes and better health, we it is unclear what part of that relationship is causal. Measuring the extent of health effects is critical to thinking about how investments in education may translate to downstream improvements in the health and functioning of society as well as understanding the optimal level of public funding for higher education.

In Chapter III, we explore the impact of educational attainment on long run health and labor market outcomes. We leverage large changes in college attainment generated by the elimination of a large subsidy to higher education, the Student Benefit Program. Under the Student Benefit Program, students age 18-21 who were the children of retired, disabled, or deceased Social Security beneficiaries were eligible to receive monthly payments if they enrolled as full-time college students. At the program's peak in the late-1970s, 12 percent of full-time college students were receiving benefits that averaged roughly $7,500 (2019 dollars). For context, this was roughly equivalent to the student-weighted average total tuition, fees, room, and board charged across all public four-year institutions at the time. In 1981, Congress voted to eliminate the program. We leverage the decrease in college attainment resulting from the elimination to examine the impacts of changes in college attainment.

We use a large sample of administrative Social Security records to precisely identify individuals impacted by the elimination of the program and link these individual records to Census Bureau survey data. Preliminary results show that eligible students received an average of over $7,000 dollars in financial assistance which resulted in 0.185 increased years of education. This change in educational attainment is only one fourth the size of previous estimates, suggesting that the impact of financial aid on educational attainment may have been overstated (Dynarski, 2003). In these preliminary results, the observed change in educational attainment corresponds to no statistically significant causal impact on long run labor market or health outcomes.

2. EVALUATING THE LONG-RUN EFFECTS OF WATER FLUORIDATION: EVIDENCE FROM COMMUNITY WATER FLUORIDATION PROGRAMS

2.1 Introduction

Over 70% of publicly supplied drinking water in the United States is fluoridated and the CDC has named community water fluoridation as one of the 10 greatest public health achievements of the 20th century (Gooch, 2020). Despite strong evidence that exposure to low levels of fluoride are an effective way to strengthen teeth, many individuals, communities, and industrialized countries oppose water fluoridation out of concern for potential negative health risks. While existing research has not shown conclusive evidence of negative health effects from low levels of fluoride exposure, concerns about the safety of fluoride are supported by a body of research that concludes that early childhood exposure to high doses of fluoride can cause a wide variety of health problems including weakened bones and joints as well as cognitive impairment. The lowest safe level of fluoride exposure is unclear. In this paper, I investigate the impact of early childhood exposure to community water fluoridation on both health and labor market outcomes.

The impact of fluoride on health varies based on both the amount and timing of fluoride exposure. This is true for both the positive impacts on dental health as well as the potential negative side effects on teeth, bones, and cognitive function. While it is well established that fluoride exposure makes teeth more resistant to decay, recent controversy has focused on the role of fluoride as a neurotoxin. The meta-study Choi, Zhang and Grandjean (2012) concludes that early exposure to high fluoride levels results in decreased cognitive functions equivalent to nearly one half of a standard deviation in IQ scores. While the majority of reviewed studies focus on subjects with fluoride levels well above recommended levels, some find negative cognitive effects at relative low levels as well. In a follow up meta-study incorporating more recent evidence, Grandjean (2019) concluded that safe exposure levels are likely below the levels currently recommended for water fluoridation.

Recent economics literature has also begun the process of exploring the long run labor market effects of fluoride exposure, which is potentially affected both by improvements in dental health and any impacts on cognitive function. Glied and Neidell (2010) provides the best evidence in the U.S. context, leveraging variation in the timing of community water fluoridation programs to estimate long run wage effects in the National Longitudinal Survey of Youth – 1979 (NLSY79). Due to the narrow group of birth cohorts in their sample (1957-1964) they are unable to make within-county comparisons and their results instead rely on the assumption that unobservable county characteristics affecting labor market outcomes are uncorrelated with fluoridation status. They find positive but insignificant wage effects in the full sample—driven by a statistically significant 4% increase in wages among women. A more recent paper leverages natural variation in fluoride levels in Sweden and finds positive effects of fluoride on labor force participation and income (Aggeborn and Öhman, 2017). The authors also estimate effects on cognitive ability and health, finding no effect on either. While the Swedish data provide significant precision and measurement advantages over the NSLY79, fluoride exposure is relatively low; over 90% of Swedish observations were exposed to fluoride levels less than those typically added in the United States (0.8-1.2 mg/L).

Childhood exposure to community water fluoridation has the potential to influence long run health and labor market outcomes through multiple mechanisms, but its actual impact is unclear. I provide the first large sample and quasi-experimental evidence of the long run health and labor market effects of community water fluoridation programs. The restricted Census datasets I use allow for vast increases in precision, require much weaker identifying assumptions via inclusion of birth county fixed effects, and include a broad set of outcomes, birth cohorts, and communities relative to previous work.[1]

[1]As is shown in Appendix C of Anders, Barr and Smith (2019), this type of large sample administrative data reduces the likelihood that statistically significant results are false positives, improves precision, and reduces publication bias.

I find that children exposed to community water fluoridation during early childhood experience a 1.9 percent of a standard deviation decrease in their adult economic self-sufficiency, 1.2 percent of a standard deviation increase in rates of adult disability, a 0.4 percentage point increase in likelihood of being incarcerated, a 1.0 percentage point decrease in military service and a 1.5 percentage point decrease in high school graduation.[2] To put some of these results in context, I compare them to Bailey et al. (2020) who estimated the beneficial effects of early childhood access to food stamps using a similar set of outcomes. Taking point estimates from both studies at face value, this suggests that early childhood fluoride exposure has the potential to erase approximately two-thirds of the self-sufficiency benefits caused by early childhood utilization of food stamps.

These results show that the net effect of fluoride is negative even at relatively low levels of exposure. A net negative impact suggests that, even at levels previously thought to be safe, the tooth strengthening effect of fluoride provides less benefit than fluoride's corresponding health costs. While it is difficult to disentangle all the mechanisms at play, the observed decrease in high school graduation rates is consistent with evidence of negative cognitive effects.[3] While Aggeborn and Öhman (2017) find that fluoride improves labor market outcomes with no evidence of negative cognitive effects, the lower average fluoride exposure in their sample may reduce negative health costs enough to result in a net positive labor market impact driven by improvements in dental health.

A gradual re-evaluation of water fluoridation policies is already underway. In 2015, motivated both by research suggesting negative health effects of fluoride as well as an increasing prevalence of dental fluorosis in the U.S., the U.S. Public Health Service reduced recommended fluoride levels from 0.7mg/L and acknowledged the need for more research into the risks of low-level flu-

[2]These are the estimated effects of treatment on the treated, found by taking the intent to treat effects from Tables 3.2 and 2.3 and dividing by .37, the population weighted average percent of a county with access to fluoridated water in that counties first treatment year.

[3]Any negative health effect has the potential to impact high school graduation through increased absences or reduced ability to focus. Decreases in cognitive function are still the most likely culprit given the existing evidence of fluoride's effect on IQ scores and the direct impact that would have on academic performance.

oride exposure (DHHS, 2015). In 2019, the American Dental Association issued a statement that reaffirmed their support of water fluoridation while also welcoming additional research into the potential negative cognitive effects (ADA, 2019). Despite the acknowledged need for more research, fluoride is still being added to a majority of public water supplies in the U.S. and regulations for regions with naturally high levels of fluoride allow water to carry up to 4.0 mg/L, six times the recommended water fluoridation level. The results of this study demonstrate the need to accelerate our re-evaluation of water fluoridation policies. The observed negative impacts of fluoride combined with widespread access to the enamel strengthening benefits of fluoride through toothpaste and dental treatments provides a strong argument for ending the practice of water fluoridation and lowering the maximum levels of fluoride allowed by safe drinking water standards. If water fluoridation practices continue, more research is needed to determine the optimal level of fluoride such that the marginal benefits to dental health are not overwhelmed by negative health costs.

In the next section, I provide an overview of the use of fluoride and the current status of fluoride research. This is followed by a summary of Census bureau data I use and an outline of the stacked difference-in-differences approach. I pay particular attention to what separates the stacked difference-in-differences approach from a traditional two-way fixed effects model and how it avoids bias despite of potential heterogenous treatment effects. I will then present my primary results as well as an additional exercise replicating the methods of Glied and Neidell (2010). I conclude with a discussion of my findings and their implications for water fluoridation policy.

2.2 History of Fluoride

In the 1930's two dentists, Dr. Frederick McKay and Dr. G.V. Black, discovered that exposure to fluoride in drinking water caused a visible discoloration of teeth while simultaneously protecting teeth against decay. Additional study revealed that impacts on dental health occur during the early stages of tooth development, which begin in utero and are entirely completed by age 8. This childhood exposure directly affects the tooth structure making it more resistant to decay. Cases of tooth decay decrease as water fluoride levels increase but the marginal benefits shrink above 0.7 mg/L and plateau by 1.2 mg/L (Heller, Eklund and Burt, 1997). The most common negative side effect of fluoride is mild dental fluorosis, a cosmetic defect that is characterized by lacy white markings on teeth but does not negatively impact dental health (DHHS, 2015). Dental fluorosis increases in frequency and severity with exposure level. Severe fluorosis is not only cosmetic but includes pitting and damage to tooth structure in addition to visible discoloration. Risk of severe fluorosis increases significantly at fluoride levels above 2.0 mg/L.[4]

Targeting the potential benefits of low-level fluoride exposure, Grand Rapids Michigan became the first city to artificially add fluoride to their public water supply in 1945. Over time other communities followed their example and water fluoridation became a common, although far from universal, practice across the United States. Despite its low financial cost (as low as $0.11 yearly per capita in large cities) and its prevalence, water fluoridation decisions have been rife with controversy since the beginning of the practice (Ran and Chattopadhyay, 2016). Referendums regarding water fluoridation typically face strong opposition and frequently fail, with much of the increase in water fluoridation over time being driven by administrative decisions rather than public votes (Sapolsky, 1968). Although specific complaints have changed over time, the opposition to water fluoridation has been based on legitimate concerns as well as conspiracy theories. For example, the concern that human error could lead to toxic levels of fluoride being added to public water supplies was validated by a tragic 1992 accident where nearly 300 Alaskans were poisoned, and one died, after excessively high levels of fluoride were added to a community well. On the other hand, the no-longer-popular conspiracy theory that water fluoridation was part of an elaborate communist plot to poison or control America was shared broadly among ant-fluoridation campaigns in the 1950's Johnston (2004).

Although controversy over community water fluoridation has persisted until today, research

[4]Cases of mild fluorosis affect about 23% of people in the U.S. while severe effect less than 1% (Beltrán-Aguilar, Barker and Dye, 2010)

into the effects of fluoride has also progressed. While early research found that fluoride exposure was most beneficial for children, additional studies have shown moderate benefits for adults as well (DHHS, 2015).[5] Researchers have also explored the effects of fluoride beyond its impacts on dental health. Fluoride exposure at levels above 4.0 mg/L (four times standard water fluoridation rates) can cause skeletal fluorosis, resulting in increased joint pain and weakened bones and joints with higher risk of fracture (DHHS, 2015). While additional research has explored potential negative impacts of fluoride on thyroid health or the roll of fluoride as a carcinogen, the majority of research has found no effect of fluoride on those margins (DHHS, 2015). Increasingly, the focus of research into the health effects of fluoride has been concentrated on the potential for fluoride to act a neurotoxin and negatively impact cognitive functions. Early work using high doses of fluoride in rats showed that fluoride both passes through the blood brain barrier and results in behavioral changes, but whether or not these effects would translate to humans exposed to low doses over a long period of time remained unclear (Mullenix et al., 1995).

The metastudy Choi, Zhang and Grandjean (2012) used evidence from a collection of studies in China and Iran and concluded that high levels of fluoride exposure results in decreases of IQ by nearly half of a standard deviation. While many of the studies included in that review had methodological issues and small sample sizes, additional research by Bashash et al. (2017) in Mexico and Green et al. (2019) in Canada found that in-utero exposure to fluoridated drinking water corresponded to meaningful decreases in IQ scores of young children, especially for boys.[6] These studies accounted for individual level fluoride exposure by measuring fluoride levels in urine of expectant mothers. In fact, a fairly large body of recent literature, many of which are reviewed in the follow-up meta study Grandjean (2019), has consistently found that fluoride has negative

[5]Adult fluoride exposure reduces the production of tooth-damaging acid by mouth bacteria and simultaneously fortifies teeth making them more resistant to acid. Some evidence suggests that the benefits of adult exposure are concentrated among individuals who were also exposed to water fluoridation during childhood (Singh, Spencer and Armfield, 2003).

[6]A 1.0 mg/L increase in urine fluoride levels corresponded to a decrease of 5.0 and 3.7 IQ points in the two studies respectively.

cognitive effects with only a few exceptions.

Economists have also recently begun studying the labor market effects of fluoride, which are potentially affected by improved dental health or by any negative health effects—cognitive effects in particular. This research was led by Glied and Neidell (2010) who provide the best existing evidence in the U.S. context by leveraging variation in the timing of community water fluoridation programs to estimate the impact of childhood fluoride exposure on adult wages. Unfortunately, the narrow group of birth cohorts in the National Longitudinal Survey of Youth – 1979 (NLSY79), which includes individuals born from 1957-1964, does not provide sufficient variation for within county comparisons or a difference-in-differences analysis. Their results instead rely on the assumption that unobservable county characteristics affecting labor market outcomes are uncorrelated with fluoridation status. Perhaps due to the small sample size and limited identifying variation available among the NLSY79 cohorts, they find positive but insignificant effects in the full sample. The positive effects are driven by a statistically significant 4% increase in income among women, which the authors interpret as evidence of appearance-based discrimination. The income point estimate for males is zero.

In contrast, Aggeborn and Öhman (2017) leverage natural variation in fluoride levels in Sweden and finds positive effects on labor force participation and income, with larger effects for men. Interestingly, Aggeborn and Öhman (2017) are also able to test for any impacts on cognitive ability or health and find no effect on either outcome. While the Swedish data provide significant precision and measurement advantages over the NSLY79, fluoride exposure is low; over 90% of Swedish observations were exposed to fluoride levels less than those typically added in the United States (0.8-1.2 mg/L).

Water fluoridation remains an important public health topic due it its role as a low-cost way to improve dental health as well as its potential health risks. Despite improving trends in dental health in the U.S. which are frequently accredited to water fluoridation programs,[7] tooth decay is still one

[7] The prevalence of any tooth decay in adult teeth among adolescents decreased from 90% in the 1960's to 60%

of the most common chronic childhood diseases and one in four children below the poverty line have untreated tooth decay (Newacheck et al., 2000; Dye, Li and Thornton-Evans, 2012).

2.3 Data

The primary data source is restricted individual-level U.S. Census and American Community Survey (ACS) data linked to the Numident file (U.S. birth and death records), housed in the Census Research Data Centers. This includes ACS years 2001-2016. The Numident file contains each individual's date and location of birth as well as date of death for those who are deceased. Water fluoridation data comes from the 1992 Fluoride Census (a public record provided by the CDC).[8,9,10] Using data from the 1992 Fluoride Census, Figure 2.12 shows the rollout of community water fluoridation programs by county over time. My analysis sample is limited to individuals born in a U.S. county that was included in the 1992 fluoridation census and successfully linked to its county FIPS code. For computational ease, I collapse the data to birth-year by birth-county by survey-year level separated by both gender and race. Each collapsed cell is weighted by the number of observations in that cell for all analysis.

Summary statistics are shown in Table 2.1 for the full sample, by gender, and by treatment status. These summary statistics include basic demographic variables, components of each outcome index (which are explained in detail in the next section), as well as secondary outcomes. While there are minor differences between treated and untreated counties these differences do not affect the internal validity of the stacked difference-in-differences design. Sample size is presented as the number of unique individuals, the number of collapsed cells, and the number of observations in-

by 2004 and the CDC named community water fluoridation as one of the 10 greatest public health achievements of the 20th century.(DHHS, 2015)

[8]Matthew Neidell and Sherry Glied have also generously shared the cleaned version of the 1992 Fluoride Census used in Glied and Neidell (2010).

[9]Via a Freedom of Information Act request to the CDC, I have obtained current natural fluoride levels for each community water system. While these are not used in the current analysis, they do show that counties with lower levels of natural fluoride in their water supply were more likely to add fluoride and, among counties that added fluoride, counties with low natural fluoride levels tended to add fluoride in earlier years.

[10]The locations from both the Numident file and fluoridation records are recorded as strings at the city or county level. These locations are matched to their county level FIPS codes following Taylor et al. (2016).

cluded in the final sample—which is a function of the stacked differences-in-differences procedure described in Section 2.4 and includes many exact duplicates.

2.3.0.0.1 Defining Fluoride Exposure: Despite access to administrative records, these data sources are still unable to directly identify the amount of fluoride that is consumed during each year of childhood. I define treatment at the county-birth-cohort level as the fraction of childhood years with any potential exposure to community water fluoridation. Childhood here is defined to include the year of an individual's birth through the year that each cohort reaches age five.[11,12] By this definition, a fully treated county-birth-cohort would have been exposed to fluoride the entire calendar year of their birth and for each of the following five calendar years. Because counties may have multiple public water systems with different water fluoridation policies and because some households source drinking water from private wells not all individuals in a treated county will drink fluoridated water. I am unable to identify individual children's water fluoridation exposure within a treated county. As a result, this primary treatment definition fails to account for the resulting variation in the fraction of each county drinking fluoridated water and the resulting estimates can be interpreted as intent to treat (ITT) effects.[13] Event studies exploring the potential for non-linear treatment affects by age at first exposure are described in Sections 2.4 and 4.2.

2.3.0.0.2 Outcome Variables: The purpose of this research is to identify the net labor market and health effects of community water fluoridation. Using a construction similar to Bailey et al. (2020), I examine two indices that best capture these outcomes in the ACS: (i) economic self-sufficiency, and (ii) physical ability and health. These indices average across standardized com-

[11]This treatment definition is consistent with other early childhood interventions where years of exposure is the most relevant parameter. Specifically, both Hoynes et al. (2016) and Barr and Smith (2021) use the fraction of early childhood with access to food stamps in order to estimate long run effects in a difference-in-differences setting.

[12]This definition intentionally does not account for differences in fluoridation levels (parts per million). Fluoride levels were determined at the local level, but CDC guidelines adjusted recommended rates relative to average local temperatures which may affect rates of water evaporation and consumption. Because of this, variation in fluoride level between 0.8-1.2 should not reflect actual increases in individual fluoride intake but simply a difference in the level of water fluoride level necessary to reach an equivalent per-person level of fluoride exposure.

[13]Section 4.2 discusses the implied treatment on the treated (TOT) effects and presents results from an alternative treatment definition that directly incorporates the percent of a county exposed to water fluoridation.

ponent variables, reversing signs when necessary, such that a more positive value implies a better outcome. The Economic Self-Sufficiency Index includes variables indicating whether or not an individual was in the labor force, worked last year, weeks worked last year, usual hours worked per week, labor income, other income not from public sources, income-to-poverty ratio, not in poverty, reverse coded income from welfare, and reverse coded income from supplemental security.[14] The Physical Ability and Health Index includes reverse coded information on the presence of an ambulatory or independent living difficulty, a cognitive difficulty, a vision or hearing difficulty, and a self-care difficulty.[15]

The index approach alleviates concerns about multiple hypothesis testing and improves statistical power (Kling, Liebman and Katz, 2007). In addition to these two primary outcomes, I also estimate effects on the secondary outcomes of high school graduation, military service, survival to 2020, and incarceration. The next section outlines the details of my analytical approach.

2.4 Analyses

I use a stacked difference-in-differences strategy leveraging the staggered adoption of community water fluoridation across the United States. This design compares outcomes of county-birth-cohorts with exposure to fluoridated water those without any, while controlling for county and year of birth. This strategy does not rely on the exogeneity of fluoride levels conditional on observables, but on the weaker assumption that the shift in health and labor market outcomes of untreated individuals across time effectively proxies for the shift in outcomes that would have occurred for individuals drinking fluoridated water in the absence of fluoride treatment. While I am using a stacked differences-in-differences design, the basic non-stacked version is a useful starting point to discuss merits of this approach. That non-stacked reduced form difference-in-differences specification would be:

[14] Dollar values are inflation adjusted to 2016 dollars prior to index creation.

[15] While the variables "ambulatory difficulty" and "selfcare difficulty" are separate in all ACS surveys, they were asked in a single question for the 2000 decennial long form. For consistency, they were combined in all ACS to represent any ambulatory or selfcare difficulty. The variables "any hearing difficulty" and "any vision difficulty" were also available separately after 2007 but were combined into a single variable in all years for consistency.

$$Y_{ct} = \theta_c + \delta_{s(c)t} + \mu X_{ct} + \beta(Exp6)_{ct} + \epsilon_{ct} \tag{2.1}$$

In this specification, $(Exp6)_{ct}$ represents a county-birth cohort's cumulative exposure to fluoridated water during childhood. The long run health and labor market outcomes are represented by Y_{ct}; while θ_c and $\delta_{s(c)t}$ respectively represent birth county and state-by-birth-cohort fixed effects; and X_{ct} contains a vector of covariates including sex, age, age squared, race and survey year. Although not necessary for identification, the controls for sex, age, race, and survey year are included to increase precision.

Recent literature has shown that, in settings with staggered adoption, the two-way fixed effects approach in Equation 2.1 requires the strong assumption of homogeneous treatment effects to remain unbiased (Goodman-Bacon, 2018; Sun and Abraham, 2020; Chaisemartin and d'Haultfoeuille, 2020). Specifically, in the naïve application of equation 2.1, the coefficient β in represents the weighted average of all 2x2 comparisons between counties in my sample. This includes comparisons where previously treated counties are used as controls for later treated counties, despite the fact that these "control" counties are still being affected by dynamic treatment effects themselves. If there are any heterogeneous treatment effects between counties that are treated at different points in time, those differences in the average treatment effect or the dynamic path of treatment effects over time are not accounted for and instead introduces bias into the estimated effects.

In the setting of water fluoridation, heterogeneous treatment effects are likely for a variety of reasons. The prevalence of other sources of fluoride from dental treatments, food, and toothpaste have changed over the many years of birth cohorts included in this sample. Changing access to dental care over time or generational differences in the importance of dental health could also drive heterogeneous treatment effects. Additional heterogeneity comes from the fact that fluoride is adopted at the public water system level and many counties have multiple public water systems as well as individuals who consume drinking water from private wells. This means that the fraction of a county's population receiving fluoridated water after initial adoption of varies significantly across counties which strongly suggests heterogenous county treatment effects as a result.

Given the high potential for heterogeneous treatment effects, it is necessary to adjust Equation 2.1 to prevent previously treated counties from being used as a control for later treated groups. I do this by implementing a stacked design that is robust to heterogeneous treatment effects. For any given year of initial treatment, all never-treated, and not-yet-treated counties are valid controls while all previously treated counties need to be excluded from the control group. This means that the control group changes over time, shrinking as each treated county is removed from the pool of potential controls for counties treated in later years. Because the same county must be included as a control unit, treated unit, or excluded from the comparison depending on the year of initial treatment, I create a separate dataset every year that any county first began water fluoridation. For example, the 1980 dataset includes all counties that first added fluoride in 1980 as treated counties, all counties that are never treated as controls, counties that will be treated after 1980 as controls, and drops any county treated prior to 1980. Within each dataset, referred to as "stacks" in the remainder of the paper, I generate time variables relative to year of initial treatment for treated groups within that stack as well as a variable indicating the year of initial treatment for that stack. Then, each dataset is appended or "stacked" together. In the last three rows, Table 2.1 displays both the original number of collapsed county-birth-cohort observations as well as the total number of observations after the duplication and stacking procedure.

I also exclude from the control group any counties that are not-yet-treated but will be treated within the next eleven years. I do this for two reasons. First, given the definition of early childhood exposure from age 0-5, birth cohorts that experience any water fluoridation in their first five years of life are not clean controls but partially treated groups. Additionally, because I will be presenting dynamic treatment effects figures from Equation 2.3 shown below, I also need to exclude any groups that will become fully or partially treated within that treatment window (birth cohorts born from 15 years before through 6 years after the first year of fluoridation). These exclusions result in each treated cohort being compared only to counties that are either never treated or that remain untreated for at least eleven years after that cohort's initial treatment year.

The stacked design prevents early fluoride adopters from acting as controls for counties that adopted fluoride later. The resulting estimates represent the unbiased average treatment effect even in the presence of heterogeneous treatment affects. The updated regression equation is:

$$Y_{ctg} = \theta_{cg} + \delta_{s(c)tg} + \mu X_{ct} + \beta(Exp6)_{ctg} + \epsilon_{ctg} \tag{2.2}$$

The key difference between this and the naïve two-way fixed effect approach in Equation 2.1 is the saturation of county and state-by-birth cohort fixed effects with g, indicating the dataset or stack that each observation originated from. Standard errors are clustered at the county level which both accounts for serial correlation over time as well as the repeated inclusion of the same county as a part of multiple stacks. To explore how fluoride exposure affects children of different ages, I will estimate an additional specification where $(Exp6)$, the cumulative exposure measure, is replaced with a set of timing variables indicating the first year of water fluoridation relative to a person's birth. This dynamic difference-in-difference specification is as follows:

$$Y_{ctg} = \theta_{cg} + \delta_{s(c)tg} + \mu X_{ct} + \sum_{a=-6\,[a\neq 5]}^{15} \beta_a * 1[Fl_c - b = a] + \epsilon_{ctg} \tag{2.3}$$

In this specification, Fl_c and b represent the first year that an individual's birth county fluoridated their water and that individual's birth year. The timing variable a represents each individual's age in the first year of water fluoridation and covers the period from 6 years before birth through age 15 with age 5 as the omitted year. The dynamic treatment effects are captured in β_a and represent the effect of receiving fluoridated public water beginning at age a. All other terms are equivalent to those in equation 2.1.

These event studies will show the net effects of first fluoride exposure at a particular age, which will include both the known benefits of fluoride exposure during tooth formation, as well as any negative cognitive or health effects during the treatment window.[16]

2.5 Results

2.5.1 Primary Outcome Indices

I find that early childhood exposure to fluoride negatively impacts both health and labor market outcomes. I estimate the average intent-to-treat effect as a 0.45 percent of a standard deviation reduction in physical ability and health as well as a 0.69 percent of a standard deviation reduction in self-sufficiency; the effects are significant at the 1% and 10% level respectively. These results, as well as their robustness to alternative sets of control variables, are shown in Table 3.2. Additional robustness checks are also described in Section 4.6.1 and shown in Table 2.4. These results are estimated using the county-birth-cohort specifications described in Section 2.4, and do not account for heterogeneity in the fraction of each county that is exposed to fluoride. Because counties may have multiple public water systems with different water fluoridation policies, and because some households source drinking water from private wells, not all individuals in a treated county will drink fluoridated water. This means that the estimates shown in Table 3.2 represent the intent-to-treat (ITT) effect, or the average treatment effect in county-birth-cohorts where anyone is exposed to fluoride. These estimates include individuals who were not exposed to water fluoridation and, as a result, understate the true effect of individual fluoride consumption. In order to approximate the average effect of treatment on treated (TOT) individuals, I divide the intent-to-treat effects by 0.37, the population weighted average percent of a county drinking fluoridated water in the first year of water fluoridation.

These TOT estimates imply that drinking fluoridated water during early childhood causes a 1.9 percent of a standard deviation decrease in adult economic self-sufficiency and a 1.2 percent

[16]It is worth noting that only county of birth is observed, not counties of residence throughout childhood. The likelihood of an individual residing in their birth county decreases over time, so estimates will be attenuated toward zero when estimating the effect of exposure in later years.

of a standard deviation decrease in physical ability and health. To put the magnitude of these results in context, I compare them to Bailey et al. (2020) who estimated the beneficial effects of early childhood access to food stamps using two nearly identical indices constructed from a similar dataset. While Bailey et al. (2020) found no statistically significant effects on physical ability and health, they found meaningful benefits for adults' economic self-sufficiency. Taking point estimates at face value, my findings suggests that early childhood fluoride exposure has the potential to erase approximately two-thirds of the self-sufficiency gains from early childhood utilization of food stamps.

One alternative method to account for county level differences in the percent of treated individuals within a county is to directly incorporate this variation into the definition of treatment. In this case, Equations 2.2 and 2.3 from Section 2.4 are adjusted so that the treatment variables ($Exp6_{ctg}$ and $1[Fl-b=a]$ respectively) are divided by the fraction of the county receiving fluoridated public water during the initial treatment period. Essentially, this inflates each county's estimates by the fraction of that county that was treated, rather than inflating the average intent to treat estimate by the average treatment percentage across all counties. Table 2.5 presents these alternative results for both primary outcomes and shows the robustness of these results to various sets of control variables. This method increases the precision of the estimated effects, with effects on both primary indices significant at the 1% level. The estimated 2.0 percent of a standard deviation impact on self-sufficiency is nearly identical to the previously estimated 1.9 percent. The estimated effect on physical ability and health however is only 0.7 percent of a standard deviation, smaller than the 1.2 percent estimated previously.[17]

I also estimate dynamic effects relative to a birth-cohort's age at the time of initial water fluoridation in their county. These event studies are shown in Figures 2.3 and 2.4. These figures show

[17]Unfortunately, this method may introduce bias if counties with different treatment intensities were on different outcome trajectories prior to treatment. For example, this may be the case if urban counties, where private well use is less common, have a higher percentage of individuals drinking fluoridated water and also income and employment trends that are improving faster than those in rural areas. Despite the increased precision of using this method, the result in Table 3.2 remain my preferred specification as outlined in the pre-analysis plan (Roberts, 2021).

level effect sizes across birth cohorts with equal exposure to fluoride which is consistent with differences in cohorts exposed later being driven by water fluoridation rather than some other factor. The slope of the estimated trend line among cohorts treated from birth is included in these figures; slopes closer to zero provide the strongest evidence in support of my identifying assumption. The observed shrinking of marginal effects at older ages is also consistent with existing theory and evidence, as described in Section 2.2, that fluoride is likely to have the strongest effect on young children. These figures are discussed further in Section 4.6.1.

These results show that fluoride has a net negative impact on health and labor market outcomes even at relatively low levels of exposure. A net negative impact suggests that, even at levels previously thought to be safe, the known tooth strengthening effect of fluoride provide less individual benefit than the corresponding costs in the long-run.

2.5.2 Secondary Outcomes

I also explore the effect of childhood fluoride exposure on high school graduation, incarceration, military service, and mortality. These results are shown in Table 2.3. I find that statistically significant decreases in high school graduation, and military service as well as increases in rates of adult incarceration. Dividing these effects sizes by the 0.37, the average percent of a treated county exposed to water fluoridation, suggests TOT effects of a 1.5 percentage point decrease in high school graduation, a 0.4 percentage point increase in likelihood of being incarcerated, and 1.0 percentage point decrease in military service. To once again frame effect sizes relative to Bailey et al. (2020), these point estimates, taken at face value, suggests that early childhood fluoride exposure has the potential to erase four-fifths of the decrease in incarceration caused by early childhood utilization of food stamps.[18] Point estimates also suggests increases in mortality, measured by a decrease in the likelihood of survival to 2020, but this effect is only statistically significant for men.[19]

[18]Similar comparisons of high school graduation and military service are not possible because equivalent outcomes are not included in Bailey et al. (2020)'s study of food stamps.

[19]The TOT effect on men is a 0.3 percentage point decrease in the likelihood of surviving to 2020.

I also estimate dynamic effects on each secondary outcome relative to a birth-cohort's age at the time of initial water fluoridation in their county and these results are shown in Figures 2.5-2.8. As discussed previously, near zero slopes on the left-hand side of the figures show consistently sized effects across birth cohorts with equal exposure to fluoride; consistent effects for cohorts with equal treatment exposure supports the assumption that differences in that specific outcome are being driven by water fluoridation rather than some other factor. The impact on high school graduation is shown in Figure 2.5 and clearly shows that birth cohorts with the most fluoride exposure have the lowest rates of high school graduation. Additionally, the effects are concentration during early childhood (age 0-5) showing that exposure during those years has the largest impact on educational attainment. These results are consistent with the hypothesis that early fluoride exposure negatively impacts cognitive development.

The interpretation of impacts on military service, as shown in Figure 2.7, are less clear. While level effects for fully exposed individuals supports the identifying assumption, the effects on military service are concentrated on later years, from age 5-9. Because these are estimates of the net effect of fluoride including potential health risk as well as improvements in dental health, it is difficult to distinguish what mechanism drives this pattern of effects.[20]

While estimated increases in incarceration rates are meaningfully large, the event study in Figure 2.6 shows that this trend exists even amount cohorts that were exposed from birth (those born in the first full year of water fluoridation or up to six years after). This suggests that, of the outcomes included in this study, my identifying assumption is least likely to hold in the case of incarceration effects; effects on incarceration should be interpreted with an additional degree of caution.

[20] Anecdotally, military service is known for providing high quality medical and dental care, such that marginal individuals may seek out military service as while seeking dental treatment. If this is the case, then improvements in dental health may reduce military service. On the other hand, negative cognitive affects have an ambiguous effect on military service. While some individuals who opt out of additional education may turn to military service as an alternative, others who may have otherwise served in the military might be excluded if they are unable to pass military entrance requirements. Determining the interactions of these mechanisms at each age is beyond the scope of this study.

2.5.3 Replicating Glied and Neidell (2010)

In prior work leveraging the county level adoption of water fluoridation policies, Glied and Neidell (2010) found that water fluoridation increased wages for women. These results were estimated using birth cohorts from 1957-1964 included in the National Longitudinal Survey of Youth (NLSY79).[21] I replicate their analysis using my larger sample. Specifically, I redefine treatment to match their definition (average fraction of county exposed to water fluoridation during an individuals first five years of life), and add controls for 1960 county characteristics, state fixed effects, and fluoride exposure as an adult.[22,23] These results, for the full sample as well as by gender, are shown in Table 2.7. Contrary to the results found using the NLSY, I estimate negative effects on wages for both genders. While these effects are only statistically different from zero at the 10% level, they are sufficient to reject, at the 5% level, the positive effects estimated by Glied and Neidell (2010) for both the full population and female only samples.[24]

To explore how sample size affects the estimated results, I also repeatedly draw 1000 random samples equal to the sample sizes used by Glied and Neidell (2010).[25] I estimate treatment effects separately within each random draw following their estimation model. Figures 2.9, 2.10, and 2.11 show histograms of these results, for the full sample as well as by gender. Each histogram also includes a line indicating the effect size estimated by Glied and Neidell (2010) in their equivalent sample. While their model predicts negative effects of fluoride on wages in the my Census sample, these figures show that a non-trivial portion of small sample estimates are positive. The positive

[21] Differences between their model and my own are described in more detail in Section 2.2.

[22] This is not an exact replication of their approach, as the NLSY includes numerous individual level variable that are not available in the ACS or decennial surveys. I also do not control for other county level variables included in their analysis, such as health care and investment measures, as these controls had very little impact on their estimates.

[23] This analysis is conducted on a clean version of the ACS and decennial surveys without any of the transformation used to collapse the data to the county-birth-cohort level or "stacking" used in my preferred specification.

[24] Glied and Neidell (2010) estimate small negative effects on men, which is not statistically different from my estimates.

[25] While their sample consists of roughly 12,000 individuals, they observe these individuals multiple time resulting in observation counts of 37,098 in the male only sample, 35,297 in the female sample, and 72,395 in the combined sample. I separately draw 1000 random samples equal to the respective observation counts from the each of three groups.

effects found by Glied and Neidell (2010) may be the result of the relatively small NSLY sample being a similar outlier.[26,27]

2.5.4 Testing Identifying Assumption

The key identifying assumption is that, conditional on birth cohort and county fixed effects, the non-fluoride factors that influence an individuals' long run health and labor market outcomes are orthogonal to the presence or level of community water fluoridation in their county of birth at a particular age. This means that, conditional on birth cohort and county fixed effects, any difference in outcomes among those exposed to fluoridated water is the result of the fluoride itself and not any other factor. In this setting, the main assumption is that the shift in health and labor market outcomes of untreated individuals across time effectively proxies for the shift in outcomes that would have occurred for individuals drinking fluoridated water in the absence of fluoride treatment. It is impossible to observe the counterfactual outcomes of individuals exposed to water fluoridation, but I conduct several tests to explore how likely this assumption is to hold.

Water fluoridation is endogenously determined at the local level. One potential threat to my identifying assumption is if communities that implemented water fluoridation had outcomes that were already trending away from untreated communities at the time of fluoride adoption; this would violate parallel trends. To explore the relationship between county characteristics and the timing of decisions to adopt water fluoridation, I estimate the impact of various 1960 county characteristics on binary decision to ever adopt fluoride as well as the timing of that fluoride adoption. These results are shown in Table 2.8. I find that the decision to add fluoride is positively correlated with population, urbanicity, homeownership and education but negatively correlated with the percent of a county that voted (in the prior election) and the percent living in rural areas.

[26]It is also possible that these differences are simply a result of failing to exactly match the model used by Glied and Neidell (2010), specifically that my estimation does not contain the breadth of individual level controls included in their study.

[27]This difference is not the result of different sample periods. An alternative version of this replication procedure restricted the sample to the same birth cohorts used by Glied and Neidell (2010) (1957-1964) and found statistically significant negative effects of an even larger magnitude.

Column 2 of Table 2.4 present the result where the sample is restricted to exclude counties that never adopt water fluoridation. This has little impact on estimated effects, suggesting that my results are not driven by differential trends between treated and never treated counties. Additionally, among counties that adopted water fluoridation urban counties adopted fluoride in earlier while counties with a high percentage of the population under the age of five tended to adopt fluoride later. It is worth noting that these differences in the levels of observable characteristics are not a threat to the internal validity of my results unless they also correspond to differential trends between treatment and control counties in my outcome variables. Column 3 of Table 2.4 shows the results of including these predictors interacted with linear time trends as controls while estimating the effect of fluoride exposure on my primary outcomes.[28] While point estimates remain negative for both outcomes, effects on physical ability and health are diminished and lose statistical significance.[29]

I explore the evidence of the parallel trends assumption by generating even studies for each outcome. As shown in Equation 2.3, I estimate dynamic effects relative to a birth-cohort's age at the time of initial water fluoridation in their county. If treatment and control groups have differential trends unrelated to fluoride treatment, then we would expect individuals born after the beginning of water fluoridation to continue to trend apart despite the fact that water fluoridation is not changing for these groups.[30] On the other hand, consistently sized effects for these birth cohorts would provide suggestive evidence that the identifying assumption holds for that outcome. The slope of the estimated trend line among fully treated cohorts is displayed in Figures 2.3-2.8 as evidence for each respective primary and secondary outcome; slopes closer to zero provide the strongest evidence in support of my identifying assumption. Additionally, because treatment is likely to

[28]Only 1960 county characteristics that have a statistically significant relationship (at the 10% level) with the timing of water fluoridation are included in this linear trends specification.

[29]Given the cumulative nature of the impact of fluoride over time, it is possible that these linear trends are overfitting and absorbing some of the true impact of fluoride as well.

[30]While presence of water fluoridation is not changing, there may be minor changes in the fraction of a county that is treated during this time period, but these changes are relatively small on average and unlikely to drive any differential trends.

have the strongest effect on young children, a leveling off of treatment at older ages, due to smaller marginal impacts of fluoride at those ages, is consistent with differences being driven by water fluoridation rather than some other factor.

My identifying assumption also might also fail if there are meaningful shifts in the composition of people being born into treated and untreated counties across the sample period. This would occur if demographic shifts between counties happened simultaneous to water fluoridation or if individuals migrate between counties in response to water fluoridation. Aggeborn and Öhman (2017) suggest that migration in response to water fluoridation is unlikely because fluoride in water is colorless, odorless, and tasteless, meaning that changes in water fluoridation are not salient to the affected populations. Additionally, decisions regarding water fluoridation are frequently made with little or no input from local residents, making it even more unlikely that water fluoridation levels are salient enough to drive migration across counties.[31] It is however still possible that migration patterns happened to coincide with water fluoridation. This test is particularly relevant, given that the great migrations of more than 6 million blacks from the rural south into urban cities continued through the 1950's and 60's, overlapping with a large portion of the fluoride variation included in this study.

I test for demographic shifts that correspond to race by estimating the effect of water fluoridation on race, gender, age at time of survey, and likelihood of living in birth county as an adult, as shown in Table 2.6. This test shows the timing of water fluoridation did coincide with changing racial demographics, specifically a 1 percentage point decline in the fraction of the population that was white. This means that counties that adopted fluoride also tended to outpace control counties in the rate at which racial diversity increased. While any causality between county migration and water fluoridation remains unlikely, it does appear that the timing of the great migration into urban areas coincided with the adoption of water fluoridation which also tended to be adopted early in

[31] While some referendums were held allowing individuals to vote on community water fluoridation, roughly two-thirds of early water fluoridation decisions were made by government administrators without citizen input (Crain, Katz and Rosenthal, 1969).

urban areas as shown in Table 2.8.

These differential trends in racial demographics are a threat to my identifying assumption if control counties are not a reliable counterfactual for treatment counties. I explore the impact of migration and its potentially effect on my estimates in several ways. First, I estimate effects restricted to only white individuals. If my overall effects are driven by migration of non-white individuals around the time of water fluoridation, then restricting the sample to only include white individuals eliminates that source of variation and should result in shrinking effect sizes. In practice, comparing Column 1 in Table 2.4 to my primary results in Table 3.2 shows that effect sizes are nearly identical even when race is restricted to only white individuals.[32] While effects are much smaller in this sample than in Column 1 of Table 3.2 (estimates with no controls), they are quite similar to estimated effects in the full sample when race fixed effects are included, suggesting that these fixed effects are already controlling for any key differences between current county residents and migrating racial groups.

While the balance tests only clearly show changing racial makeup over time, it is possible that there are changes to other unobservable characteristics that coincide with the timing of water fluoridation. In order to account for more generally patterns of migration I also estimate effects on primary outcomes while restricting the sample to only include individuals with strong geographic roots, measured by an individual living in their birth county at the time of their adult survey.[33,34] Additionally, I create a county level measure of both in and out migration where in migration is defined as the fraction of individuals surveyed in a county who were not born there and out migration is defined as the fraction of individuals born in that county who are also surveyed there

[32]Migration may still affect the outcomes of individuals in this sample indirectly through changing peer groups and county characteristics, but I expect these effects to be relatively small compared to potential direct effect of a changing sample.

[33]It is still possible that birth cohort were changing over time as a response to the migration of their parents generation. But, to the extent that first generation residents of a county are less likely to remain in county through adulthood, this measure still captures a sample that is ex-ante less likely to be affected by migration effects.

[34]This sample also serves the dual purpose of focusing on individuals who likely lived in their birth county throughout childhood, giving a more accurate representation of their fluoride treatment status.

as adults. I create several samples, restricting to counties with progressively lower levels of both in and out migration rates. These estimates are shown in Table 2.4 columns 5-7 and represent counties with both in- and out-migration levels below the 90th, 75th, and 50th percentiles respectively. The magnitude and statistical significance of estimates remains consistent across these migration cuts. These results suggest that while significant migration did occur during my sample period, and even coincided with the timing of water fluoridation for some groups, migration is unlikely to be a primary driver of the estimated effects of water fluoridation.

2.6 Conclusion

Tooth decay is one of the most common chronic childhood diseases in the United States and one in four children below the poverty line have untreated tooth decay (Newacheck et al., 2000; Dye, Li and Thornton-Evans, 2012). Water fluoridation has been promoted since 1945 as a simple, cost effective, and egalitarian approach to improving dental health. Today, over 70% of publicly supplied drinking water in the United States is fluoridated. But, despite strong evidence that exposure to low levels of fluoride are an effective way to strengthen teeth, recent evidence has suggested that fluoride may have negatively affect cognitive ability even at these low levels (Choi, Zhang and Grandjean, 2012; Grandjean, 2019). On the other hand, recent studies within economics have also found that childhood exposure to water fluoridation improves adult labor market outcomes (Glied and Neidell, 2010; Aggeborn and Öhman, 2017).

In this paper, I use U.S. Census data linked to childhood fluoride exposure to provide large sample quasi-experimental evidence of the long run health and labor market effects of community water fluoridation programs. This data includes both respondents to the long form 2000 decennial census as well as American Community Survey respondents from 2001 to 2016. I generate a physical ability and health index as well as a self-sufficiency index and estimate the effect of childhood exposure to water fluoridation on these outcomes as well as the secondary outcomes of high school graduation, military service, incarceration, and mortality.

I find that children exposed to community water fluoridation from age zero to five experience a 1.9 percent of a standard deviation decrease in their adult economic self-sufficiency, 1.2 percent of a standard deviation increase in rates of adult disability, as well as a 1.5 percentage point decrease in high school graduation, a 1.0 percentage point decrease in military service, and a 0.4 percentage point increase in likelihood of being incarcerated. There are no statistically significant effects on mortality. These results show that the net effect of fluoride is negative even at relatively low levels of exposure. A net negative impact suggests that, even at levels previously thought to be safe, the tooth strengthening effect of fluoride provides less benefit than fluoride's corresponding health costs.

These findings have important policy implications for water fluoridation policies. Fluoride is still being added to a majority of public water supplies in the U.S. and regulations for regions with naturally high levels of fluoride allow water to carry up to 4 mg/L, four times the level of water fluoridation level evaluated in this study. Many regions around the globe have groundwater that is naturally high in fluoride. The results of this study demonstrate the need for a re-evaluation of water fluoridation policies. The observed negative impacts of fluoride combined with widespread access to the enamel strengthening benefits of fluoride through toothpaste and dental treatments provides a strong argument for ending the practice of water fluoridation and lowering the maximum levels of fluoride allowed by safe drinking water standards. If water fluoridation practices continue, more research is needed to determine the optimal level of fluoride such that the marginal benefits to dental health are not overwhelmed by negative health and labor market costs. Further study is needed to determine the exact biological mechanisms that are driving these negative effects and discover solutions that mitigate them.

2.7 Tables

Table 2.1: Summary Statistics

	(1) Full Sample	(2) Men	(3) Women	(4) Treated Counties	(5) Never Treated
Demographics					
White	0.876	0.886	0.866	0.871	0.882
Male	0.480	1	0	0.478	0.482
Age	51.82	51.69	51.93	54.88	48.38
Resides In Birth County	0.164	0.167	0.162	0.167	0.162
Physical Ability and Health Index					
No Ambulatory Difficulty	0.921	0.927	0.915	0.909	0.933
No Cognitive Difficulty	0.953	0.952	0.953	0.949	0.956
No Independent Living Difficulty	0.944	0.948	0.939	0.937	0.951
No Hearing Or Vision Difficulty	0.971	0.967	0.975	0.968	0.975
Self-sufficiency Index					
In Laborforce	0.698	0.759	0.641	0.658	0.742
Worked Last Year	0.736	0.792	0.684	0.699	0.778
Average Weekly Work Hours	29.11	33.38	25.18	27.56	30.85
Weeks Worked Last Year	10.84	11.93	9.830	9.570	12.27
Labor Income	39460	53410	26600	38140	40950
Other Income	3359	4232	2553	3860	2796
Percent Of Poverty Level	351	360	342	354	347
Not In Poverty	0.920	0.932	0.909	0.922	0.918
Welfare Income	46	34	57	40	53
Social Security Income	230	224	236	249	209
Other Outcomes					
Incarcerated	0.008	0.015	0.001	0.006	0.01
Veteran	0.158	0.300	0.027	0.178	0.135
Graduated High School	0.860	0.848	0.870	0.854	0.867
Currently Married	0.686	0.718	0.656	0.688	0.684
Survived To 2020	0.944	0.934	0.953	0.931	0.958
Sample (Cells)	32,660,000	15,890,000	16,770,000	15,370,000	17,290,000
Unique Piks	29,150,000	13,860,000	15,300,000	24,850,000	4,296,000
Collapsed Cells	3,493,000	1,668,000	1,825,000	2,087,000	1,406,000

Note: This table shows summary statistics for the primary sample in column (1) with additional summary statistics by gender in columns (2-3) and by county treatment status in columns (4-5). Summary statistics for the component parts of the primary two outcome indices are listed separately. The number of observations included in each regression (after collapsing and duplicating data as described in the analysis section) is included as "Sample (Cells)".

Table 2.2: Main Outcomes - By Gender

	(1)	(2)	(3)	(4)	(5)
Full Sample					
Physical Ability and Health Index	-0.0064***	-0.0046***	-0.0046***	-0.0045***	-0.0045***
	(0.0015)	(0.0014)	(0.0014)	(0.0014)	(0.0014)
Self-sufficiency Index	-0.0099**	-0.0054	-0.0057	-0.0054	-0.0069*
	(0.0044)	(0.0038)	(0.0037)	(0.0037)	(0.0037)
Observations	32,660,000	32,660,000	32,660,000	32,660,000	32,660,000
Women Only					
Physical Ability and Health Index	-0.0063***	-0.0043***	-0.0043***	-0.0042**	-0.0042***
	(0.0017)	(0.0016)	(0.0016)	(0.0016)	(0.0016)
Self-sufficiency Index	-0.0066	-0.0032	-0.0032	-0.0029	-0.0041
	(0.0045)	(0.0041)	(0.0041)	(0.0041)	(0.0041)
Observations	16,770,000	16,770,000	16,770,000	16,770,000	16,770,000
Men Only					
Physical Ability and Health Index	-0.0066***	-0.0049***	-0.0049***	-0.0047***	-0.0048***
	(0.0018)	(0.0017)	(0.0017)	(0.0017)	(0.0017)
Self-sufficiency Index	-0.0138***	-0.0083**	-0.0083**	-0.0080**	-0.0099**
	(0.0048)	(0.0040)	(0.0040)	(0.0040)	(0.0040)
Observations	15,890,000	15,890,000	15,890,000	15,890,000	15,890,000
Race FE	N	Y	Y	Y	Y
Gender FE	N	N	Y	Y	Y
Survey Year FE	N	N	N	Y	Y
Age FE	N	N	N	N	Y

Note: This table displays the primary index outcomes as additional controls are added - ending with the preferred specification in column (5). Results are also shown separately by gender. Observations refers to the number of observations used in each regression, after the after the collapsing and duplication procedures outlined described in the analysis section. Significance levels indicated by: * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

Table 2.3: Secondary Outcomes

	(1) HS Diploma	(2) Incarcerated	(3) Alive in 2020	(4) Veteran
Full Sample	-0.0054***	0.0015***	-0.0006	-0.0038***
	(0.0014)	(0.0003)	(0.0004)	(0.0011)
Observations	18,890,000	22,970,000	32,660,000	32,660,000
Women Only	-0.0048***	0.0003***	-0.0001	-0.0002
	(0.0014)	(0.0001)	(0.0004)	(0.0003)
Observations	9,708,000	11,800,000	16,770,000	16,770,000
Men Only	-0.0061***	0.0027***	-0.0011**	-0.0079***
	(0.0016)	(0.0005)	(0.0005)	(0.0021)
Observations	9,183,000	11,170,000	15,890,000	15,890,000

Note: This table displays a set of secondary outcomes with each column representing a different outcome and subsequent rows presenting results separately by gender. Observations refers to the number of observations used in each regression, after the after the collapsing and duplication procedures outlined described in the analysis section. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 2.4: Robustness Checks

	(1) White	(2) Ever Treated Counties	(3) Linear Time Trends	(4) Reside in Birth County	(5) Lowest 90% Migration	(6) Lowest 75% Migration	(7) Lowest 50% Migration
Physical Ability & Health Index	-0.0046***	-0.0049***	-0.0016	-0.0054**	-0.0042***	-0.0050***	-0.0052***
	(0.0014)	(0.0017)	(0.0015)	(0.0022)	(0.0015)	(0.0016)	(0.0019)
Self-Sufficiency Index	-0.0046	-0.0047	-0.0063*	-0.0106**	-0.0095*	-0.0101**	-0.0076**
	(0.0034)	(0.0037)	(0.0038)	(0.0052)	(0.0038)	(0.0042)	(0.0043)
Observations	24,360,000	15,370,000	32,040,000	17,390,000	27,730,000	19,800,000	9,487,000

Note: This table shows results from various robustness checks with each column representing a separate specification and the two rows showing the effect on the two primary outcomes. Column (1) restricts the sample to white individuals. Column (2) restricts the sample to only include counties that were eventually treated within the treatment window. Column (3) shows the results from including demographics controls for each county interacted with linear time trends. Column (4) restricts the sample to only include individuals who resided in their birth county at the time of their survey. Columns (5-7) restrict the sample to exclude counties with high levels of migration. County level migration is defined in two different ways. First, as the fraction of individuals born in a county who were surveyed elsewhere as an adult and secondly as the fraction of adults living in a county who were not born there. Counties with migration rates above the 50th, 75th and 90th percentile in either measure were excluded from the respective samples. Observations refers to the number of observations used in each regression, after the after the collapsing and duplication procedures outlined described in the analysis section. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 2.5: Main Outcomes - Adjusted by Percent of County Treated

	(1)	(2)	(3)	(4)	(5)
Full Sample					
Physical Ability and Health Index	-0.0117***	-0.0069***	-0.0069***	-0.0067***	-0.0069***
	(0.0026)	(0.0024)	(0.0024)	(0.0025)	(0.0024)
Self-sufficiency Index	-0.0292***	-0.0173**	-0.0175**	-0.0171**	-0.0200***
	(0.0085)	(0.0072)	(0.0071)	(0.0072)	(0.0072)
Observations	32,660,000	32,660,000	32,660,000	32,660,000	32,660,000
Race FE	N	Y	Y	Y	Y
Gender FE	N	N	Y	Y	Y
Survey Year FE	N	N	N	Y	Y
Age FE	N	N	N	N	Y

Note: This table displays primary outcomes when the treatment variable has been adjusted to incorporate the fraction of county exposed to fluoride. Observations refers to the number of observations used in each regression, after the after the collapsing and duplication procedures outlined described in the analysis section. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 2.6: Balance Tests

	(1) Age	(2) Male	(3) Resides in Birth County	(4) White
Balance Tests	0.0077	0.0003	0.0002	-0.0101***
	(0.0070)	(0.0007)	(0.0014)	(0.0026)
Observations	32,660,000	32,660,000	32,660,000	32,660,000

Note: This table displays the results of balance tests where my preferred stacked difference-in-differences design was used to estimate any changes in observable demographic characteristics that simultaneously with treatment. Observations refers to the number of observations used in each regression, after the after the collapsing and duplication procedures outlined described in the analysis section. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 2.7: Replication of Glied and Neidell (2010)

	(1) Full Sample	(2) Female	(3) Male
Replication Results			
Log Wage	-0.0105*	-0.0076	-0.0125
	(0.0064)	(0.0068)	(0.0078)
Observations	19,320,000	10,140,000	9,179,000

Note: This table displays the results of replicating Glied and Neidell (2010) by estimating the effect of childhood exposure to water fluoridation on log hourly wages. The details of this replication are outlined in Section 2.7. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

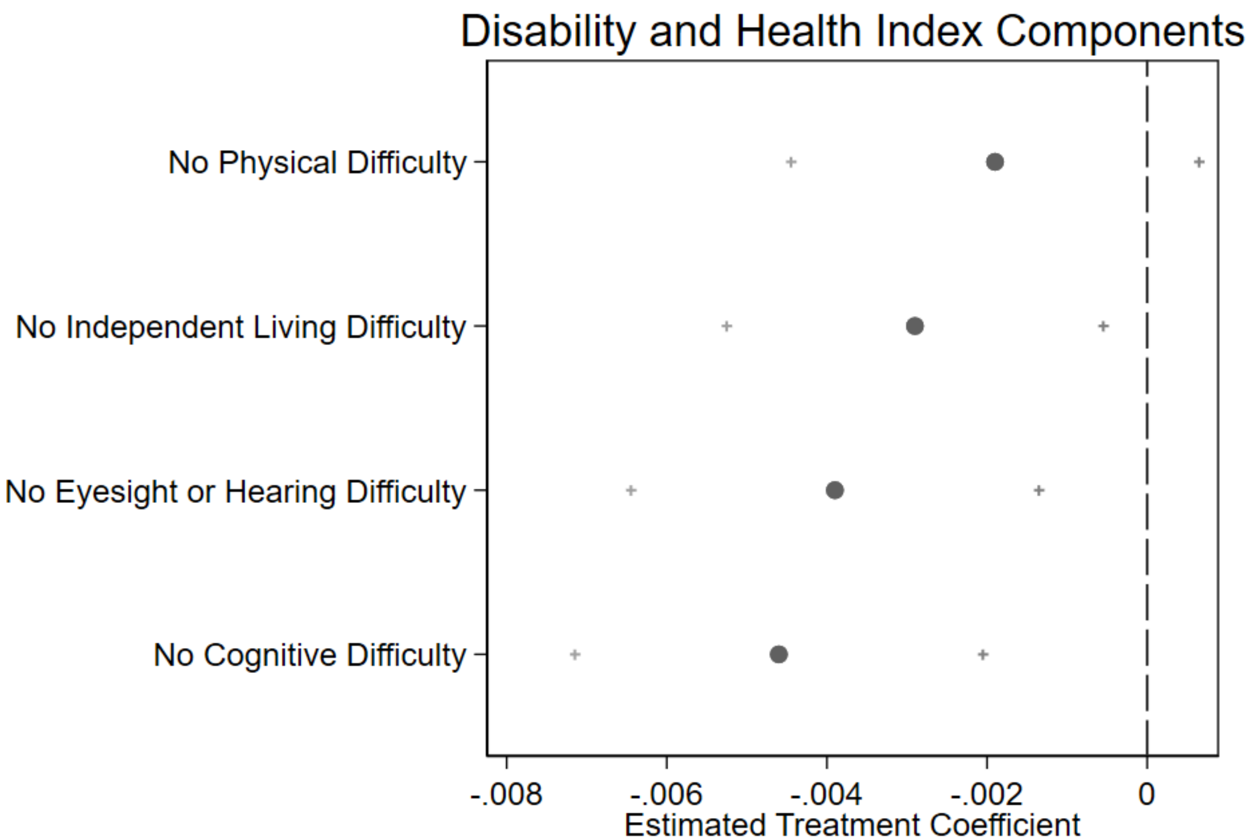
Table 2.8: Water Fluoridation and 1960 County Characteristics

	(1) Ever Treated	(2) Year of First Treatment
1960 County Characteristics		
Population (in 10,000's)	0.0016***	-0.0018
	(0.0001)	(0.8676)
Population Per Mile (in 1,000's)	-0.0018	0.1630
	(0.6822)	(0.1446)
10 Year Population Change (in 1,000's)	0.0284	-20.6297
	(0.9010)	(0.0917)
Percent in Urban Area	0.0012**	-0.0949***
	(0.0028)	(0.0000)
Percent in Rural Area	-0.0038***	-0.0200
	(0.0000)	(0.4673)
Percent Non-White	0.0014	0.0319
	(0.0974)	(0.2175)
Percent Under Age 5	-0.0177	0.8593**
	(0.0635)	(0.0052)
Percent Over Age 65	0.0095	0.3142
	(0.1590)	(0.1721)
Median Age	-0.0078	0.1359
	(0.1398)	(0.4175)
Median Income (in $1,000's)	-0.0103	-0.8151
	(0.5323)	(0.1676)
Median Years Education	-0.0321*	1.3249*
	(0.0143)	(0.0257)
Percent with Less Than 5 Years Education	-0.0078**	-0.0140
	(0.0011)	(0.8537)
Percent with High School Diploma	-0.0000	-0.1868*
	(0.5198)	(0.0185)
Death Rate	-8.9473	58.4335
	(0.1005)	(0.7600)
Marraige Rate	-1.1346	32.6536
	(0.1790)	(0.2263)
Employment Rate	-0.3762	-3.5741
	(0.1735)	(0.6838)
Percent Homeowners	0.6118***	-0.4737
	(0.0000)	(0.8851)
Percent Voted	-0.7391***	13.4338*
	(0.0000)	(0.0234)
Democratic Voteshare	0.0018*	-0.0248
	(0.0119)	(0.2541)
Household Size	-0.0559	2.7432*
	(0.1089)	(0.0105)
Observations	2988	2070

Note: This table shows the relationship between county characteristics and the endogenous decision to adopt water fluoridation. Column (1) shows the relationship between these county characteristics and a binary variable indicating if a county ever adopted water fluoridation (by 1992). Column (2) shows the relationship between county characteristics and the first year of fluoride adoption. Negative values indicate that those types of counties first adopted fluoride in earlier years. All regressions include state fixed effects. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

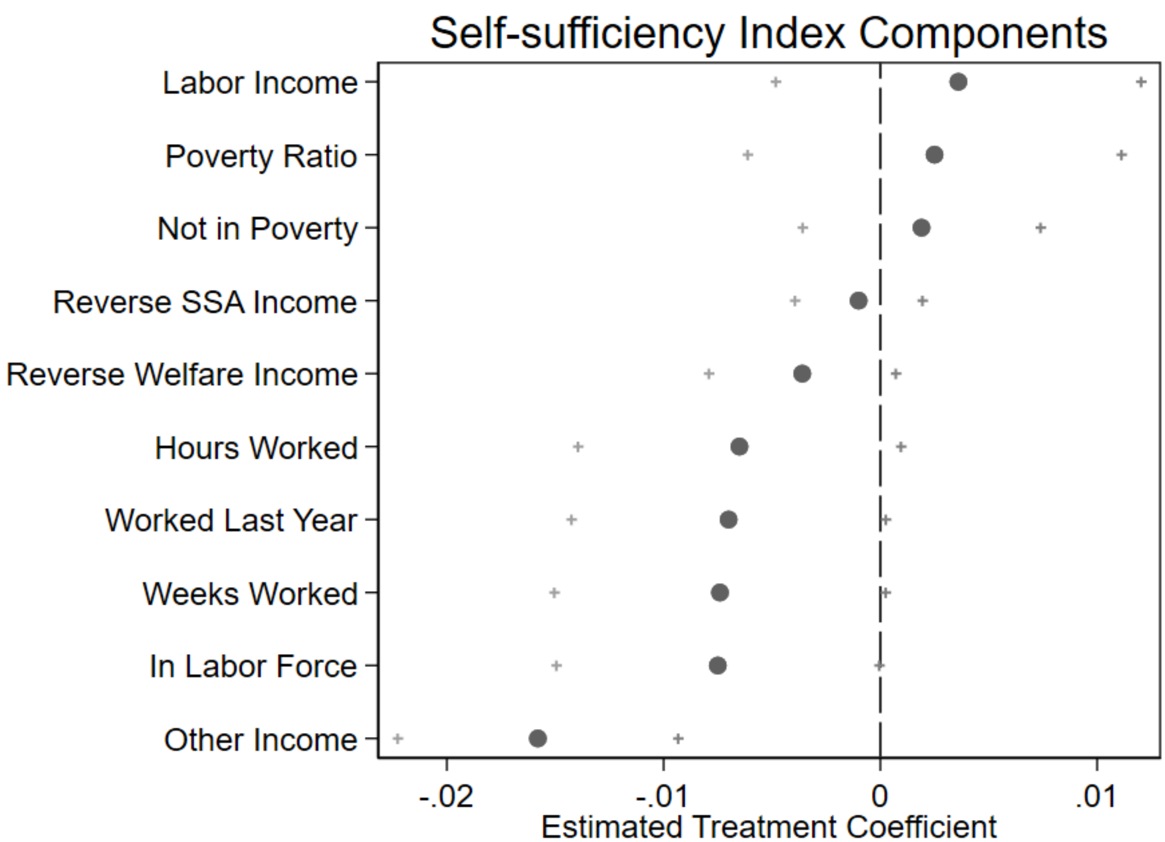
2.8 Figures

Figure 2.1: Disability and Health Index Components



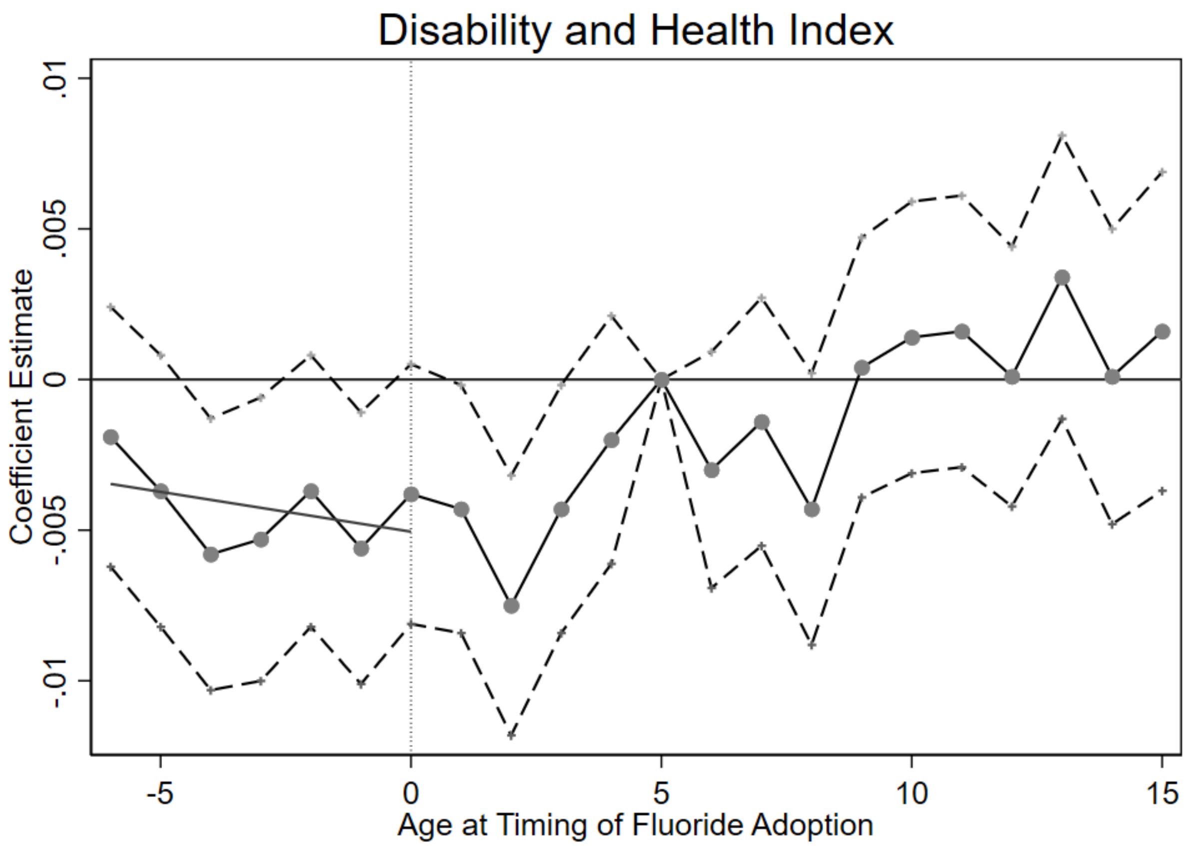
Note: This figure shows the estimated treatment effect on each (normalized) outcome included in the Disability and Health Index.

Figure 2.2: Self-sufficiency Index Components



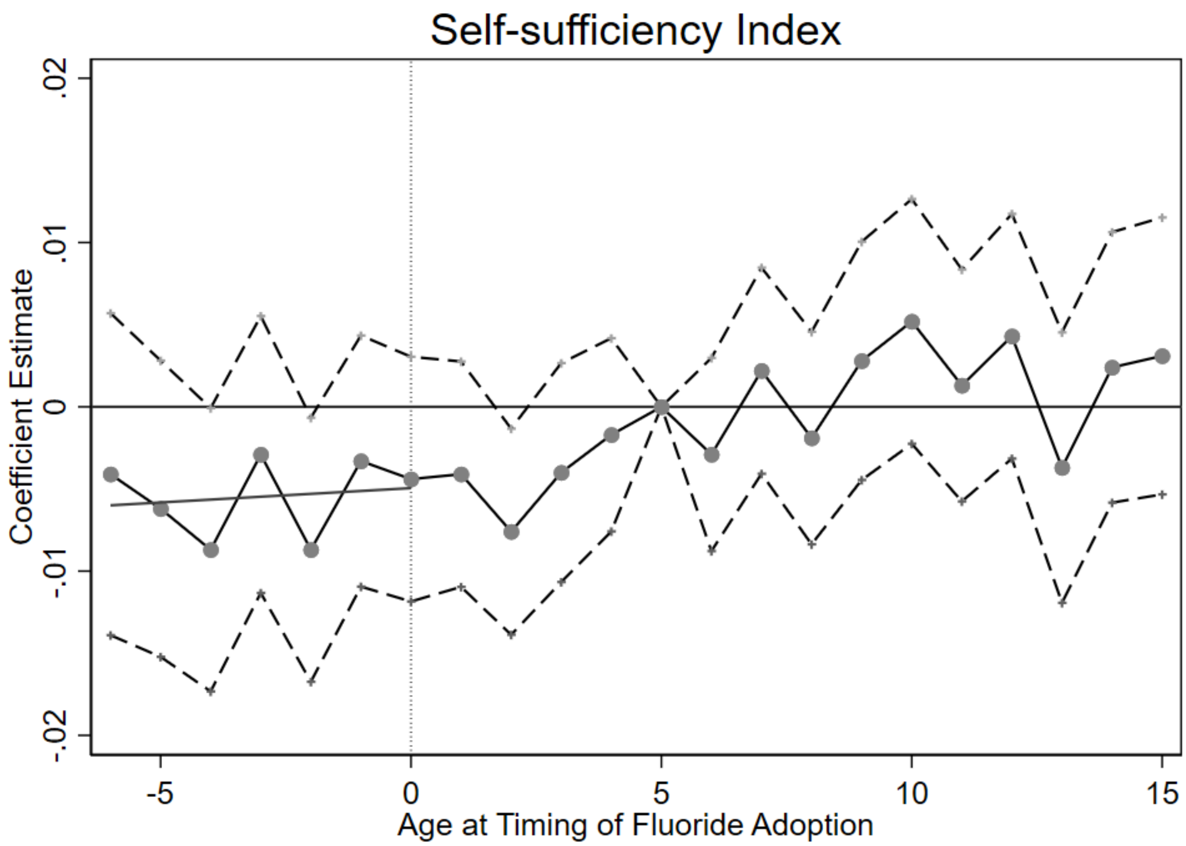
Note: This figure shows the estimated treatment effect on each (normalized) outcome included in the Self-sufficiency Index.

Figure 2.3: Event Study - Disability and Health Index



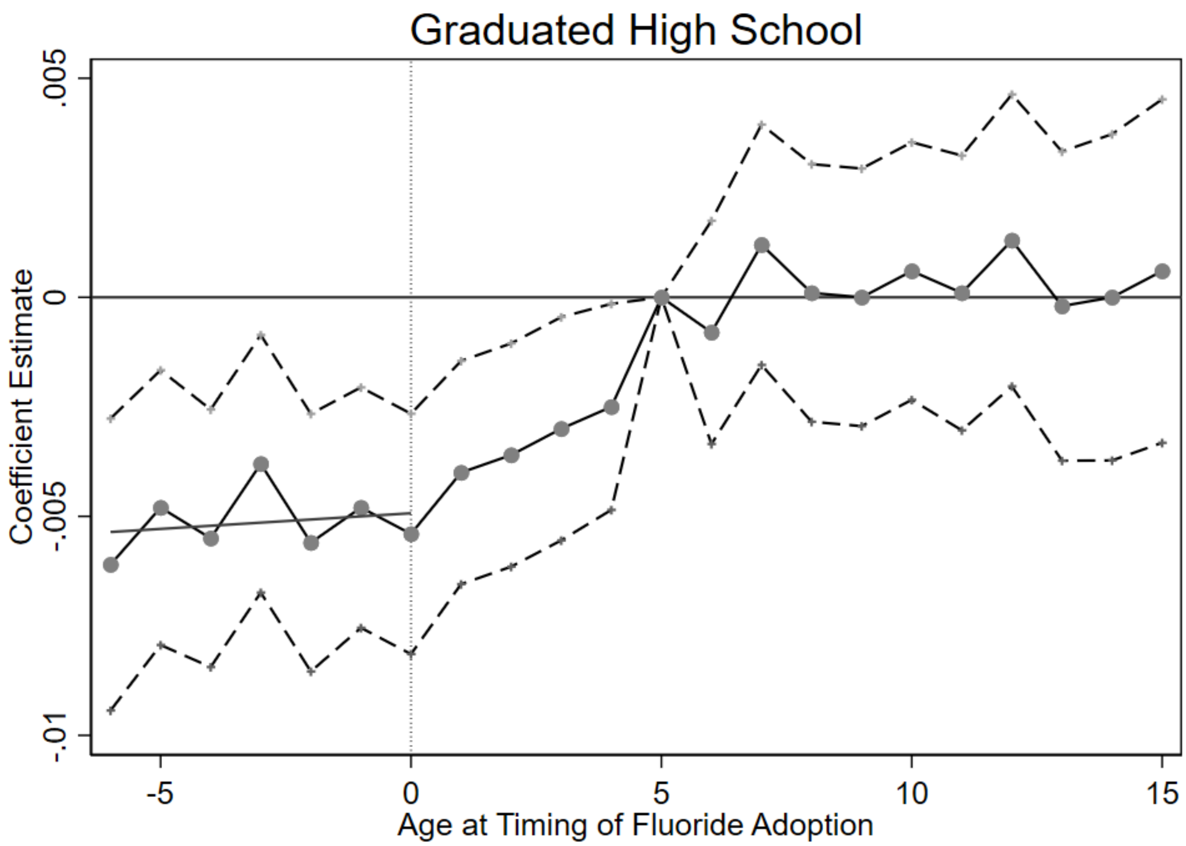
Note: This figure shows the dynamic effects of fluoride exposure on Disability and Health by cohort age at the time of county fluoride adoption. Cohorts left of zero were born after fluoride adoption and were potentially exposed to fluoride for their entire childhood. Cohorts to the right of zero received less childhood exposure depending on their age when fluoride was first adopted.

Figure 2.4: Event Study - Self-sufficiency Index



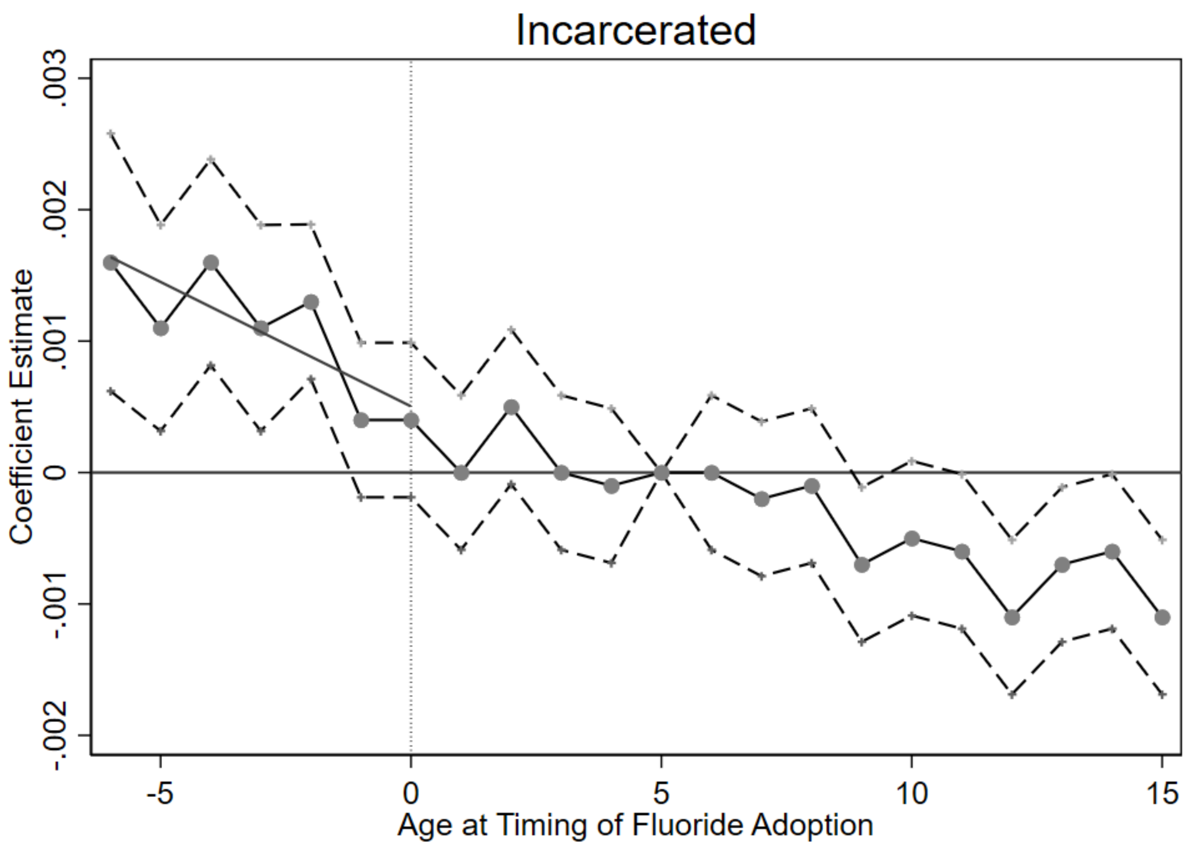
Note: This figure shows the dynamic effects of fluoride exposure on Self-sufficiency by cohort age at the time of county fluoride adoption. Cohorts left of zero were born after fluoride adoption and were potentially exposed to fluoride for their entire childhood. Cohorts to the right of zero received less childhood exposure depending on their age when fluoride was first adopted.

Figure 2.5: Event Study - Graduated High School



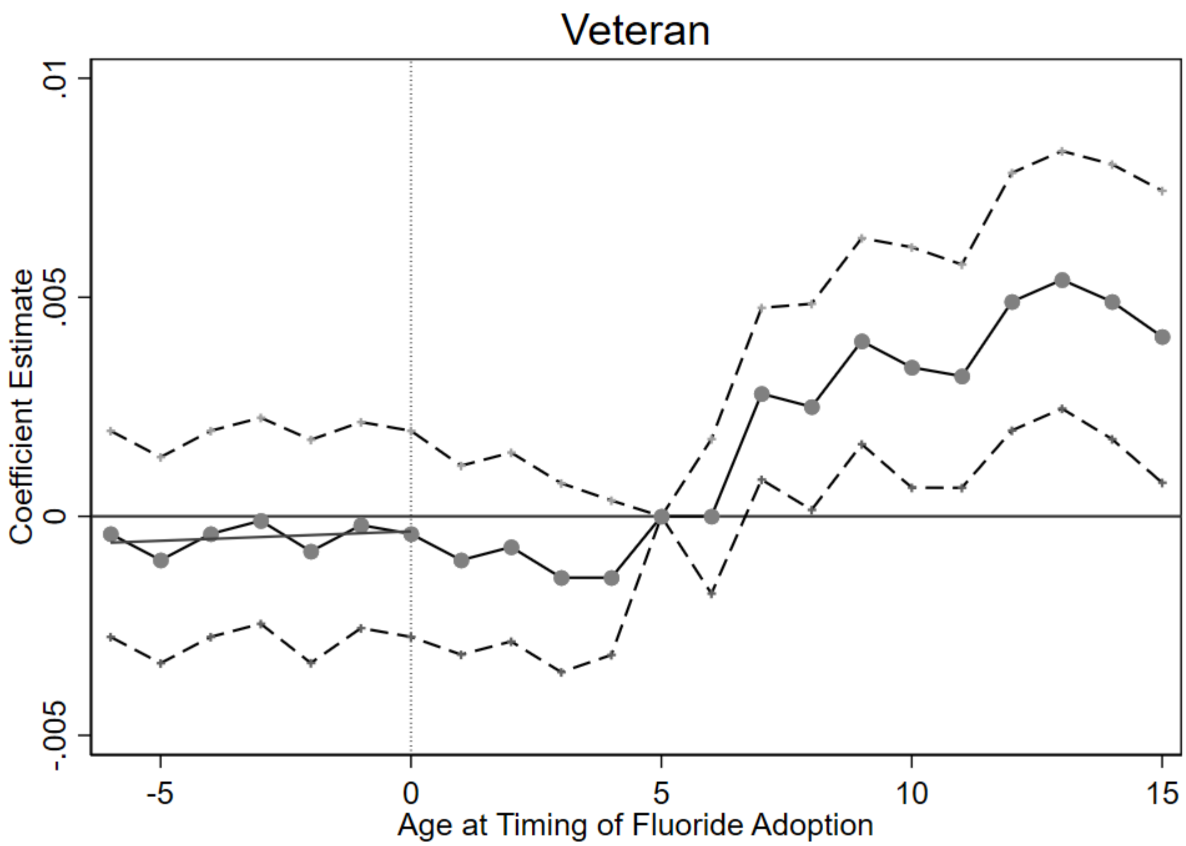
Note: This figure shows the dynamic effects of fluoride exposure on High School completion by cohort age at the time of county fluoride adoption. Cohorts left of zero were born after fluoride adoption and were potentially exposed to fluoride for their entire childhood. Cohorts to the right of zero received less childhood exposure depending on their age when fluoride was first adopted.

Figure 2.6: Event Study - Incarcerated



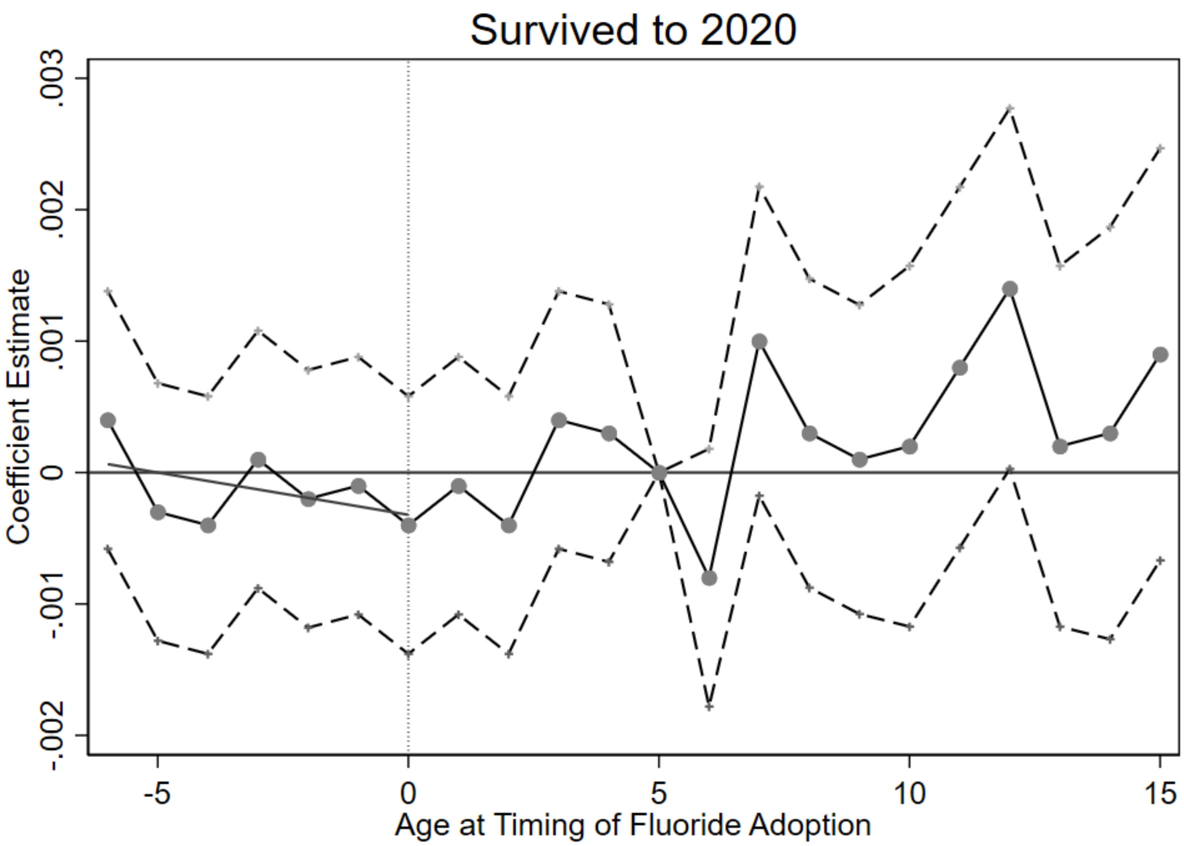
Note: This figure shows the dynamic effects of fluoride exposure on Incarceration by cohort age at the time of county fluoride adoption. Cohorts left of zero were born after fluoride adoption and were potentially exposed to fluoride for their entire childhood. Cohorts to the right of zero received less childhood exposure depending on their age when fluoride was first adopted.

Figure 2.7: Event Study - Veteran



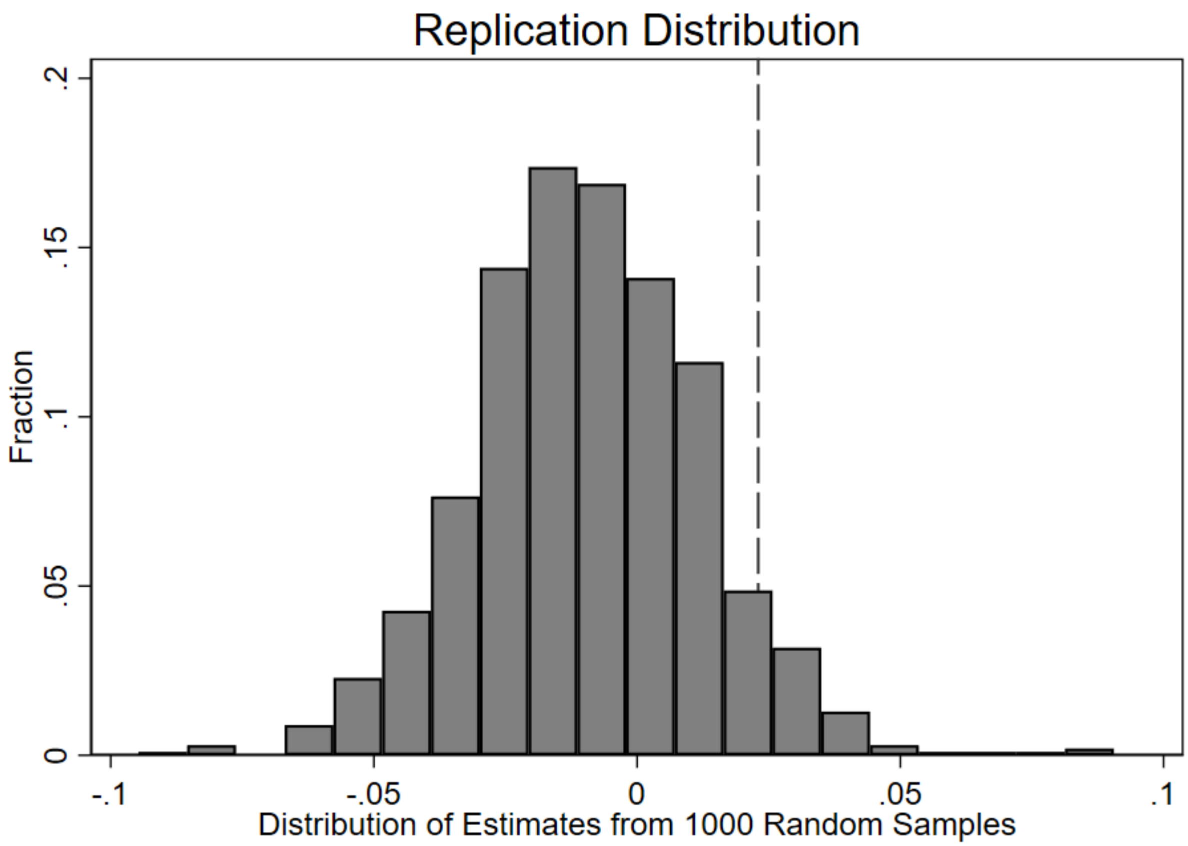
Note: This figure shows the dynamic effects of fluoride exposure on military service by cohort age at the time of county fluoride adoption. Cohorts left of zero were born after fluoride adoption and were potentially exposed to fluoride for their entire childhood. Cohorts to the right of zero received less childhood exposure depending on their age when fluoride was first adopted.

Figure 2.8: Event Study - Survived to 2020



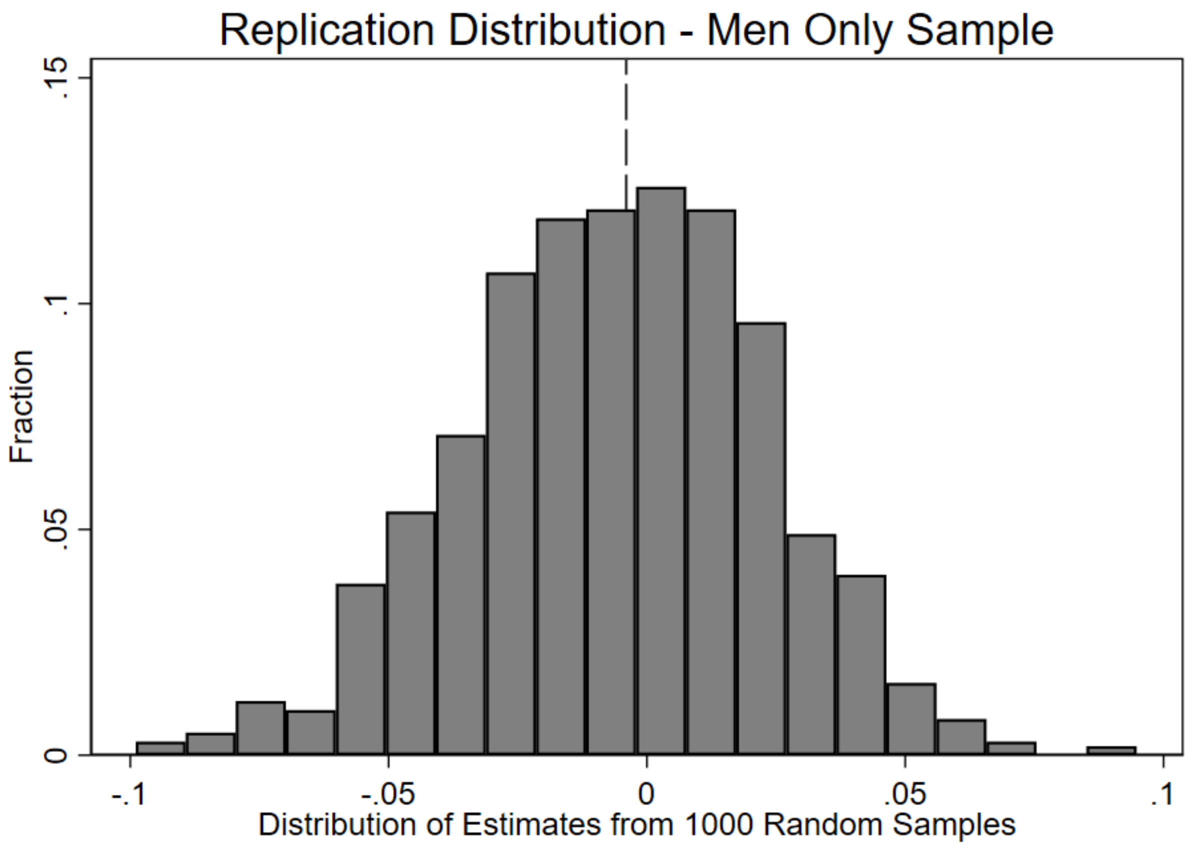
Note: This figure shows the dynamic effects of fluoride exposure on longevity by cohort age at the time of county fluoride adoption. Cohorts left of zero were born after fluoride adoption and were potentially exposed to fluoride for their entire childhood. Cohorts to the right of zero received less childhood exposure depending on their age when fluoride was first adopted.

Figure 2.9: Replication Distribution



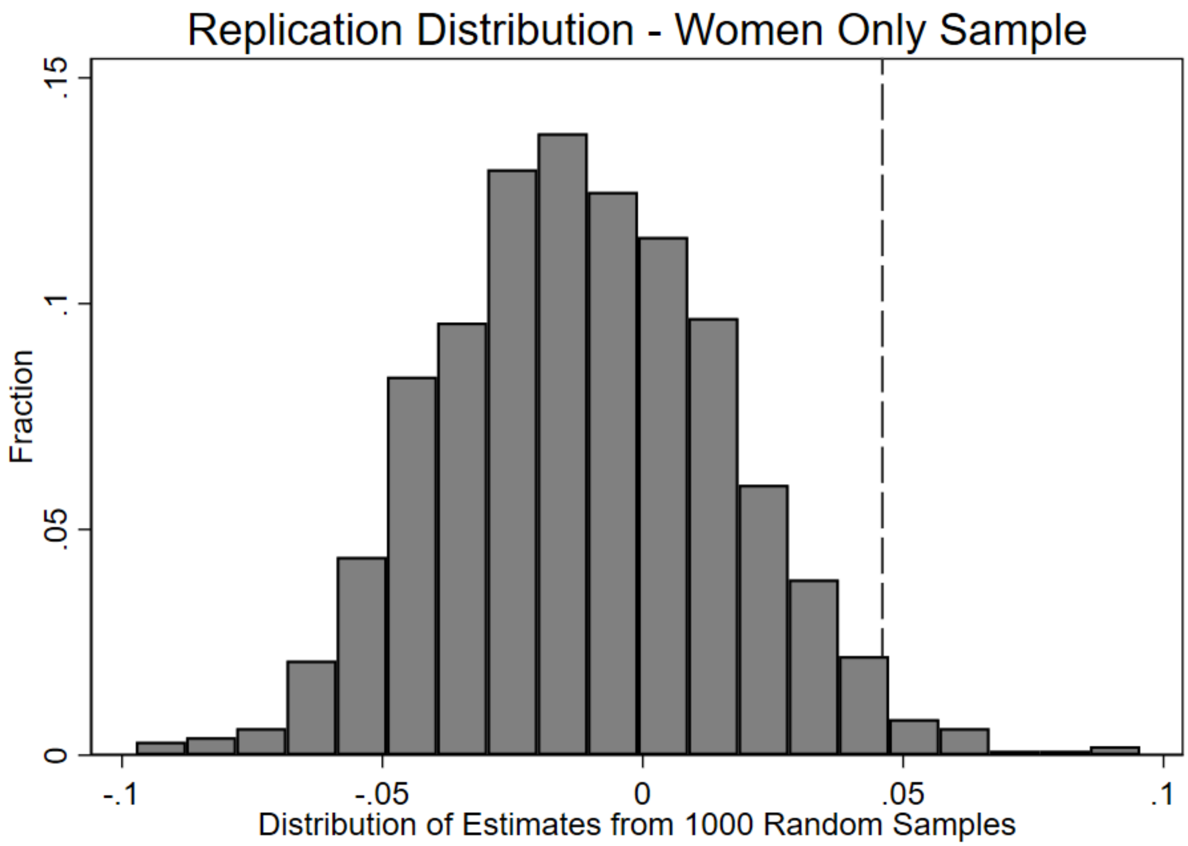
Note: This figure shows the distribution of estimates generated replicating Glied and Neidell (2010) with random 1000 random samples. The details of this replication procedure are described in Section 2.5.3. The dashed line indicates the coefficient estimated by Glied and Neidell (2010).

Figure 2.10: Replication Distribution - Men



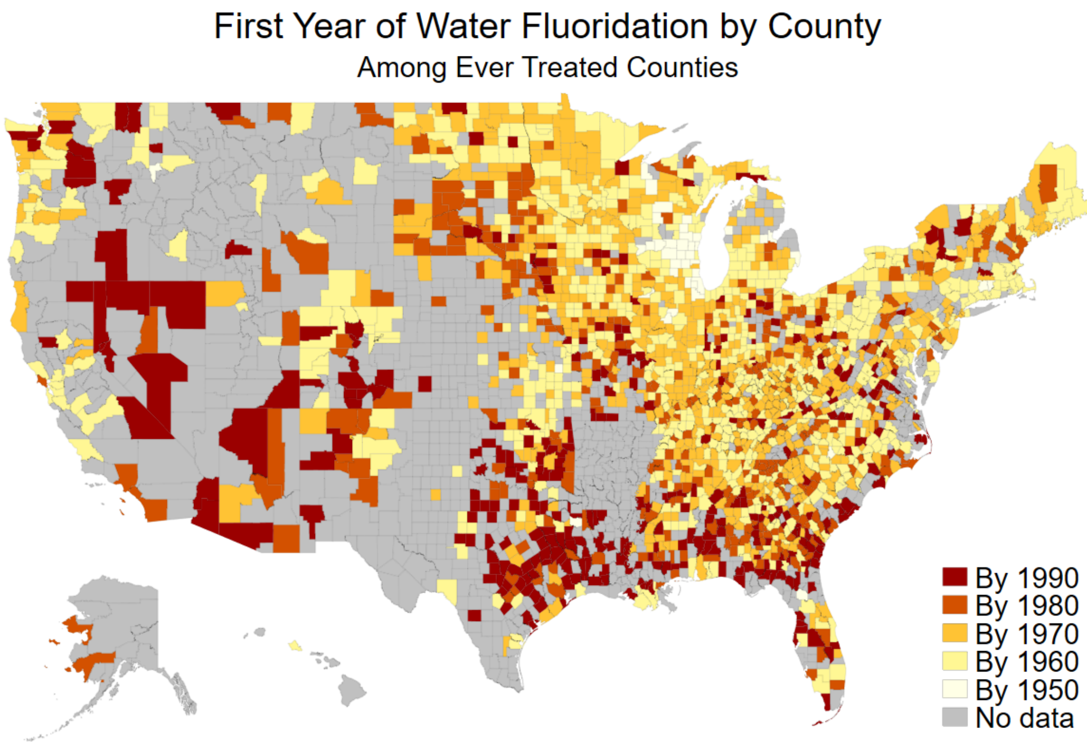
Note: This figure shows the distribution of estimates generated replicating Glied and Neidell (2010) with random 1000 random male only samples. The details of this replication procedure are described in Section 2.5.3. The dashed line indicates the coefficient estimated by Glied and Neidell (2010) in their male only sample.

Figure 2.11: Replication Distribution - Women



Note: This figure shows the distribution of estimates generated replicating Glied and Neidell (2010) with random 1000 random female only samples. The details of this replication procedure are described in Section 2.5.3. The dashed line indicates the coefficient estimated by Glied and Neidell (2010) in their female only sample.

Figure 2.12: First Year of Water Fluoridation by County



Note: This figure shows the timing of county level adoption of water fluoridation. Counties with missing water fluoridation data and never treated counties are included in the "no data" group.

3. IN-KIND BENEFITS BENEFITS: EVIDENCE FROM PUERTO RICO

This paper is joint work with Andrew Barr (Texas A&M University) and Alex Smith (West Point Academy).

3.1 Introduction

The recent surge in interest regarding universal basic income programs has revived the long-standing debate over how best to design the social safety net. A central element of this debate is whether benefits should primarily be provided in cash, where the spending decision is left to the recipient, or in kind, where much of the spending decision is made by the government. In an economic framework, the choice between these two options depends on whether household spending decisions fail to maximize the welfare of all the members of the household, as well as any externalities arising from household spending decisions. Nutritional assistance is a particularly interesting setting for investigating the tradeoffs between different benefit forms because it is a large program, providing in-kind benefits to more than 42 million Americans at a cost of more than $68 billion, in which the benefits are shared among the members of each recipient household (unlike Medicaid for example).[1] In this context, policymakers may be concerned that, under a cash benefit system, those in the household that are not the "spending decision-maker" (e.g. children) may not receive the optimal level of nutrition This deficit may yield substantial externalities if it translates into immediate or longer-term health effects.

In this paper, we investigate whether the form of nutritional assistance affects household food consumption patterns and, if so, whether these effects are large enough to generate improvements in health outcomes. We take advantage of a previously unexplored change in the Puerto Rican Nutritional Assistance Program (NAP) in 2001 that converted a cash benefit to one in which recipients were required to spend 75% of their benefits on approved food items. As the total value of

[1] USDA beneficiary and cost estimates for 2017.

the benefit remained fixed, this shift provides a unique opportunity to isolate the impact of providing food relative to an equivalent amount of cash in the context of a large-scale program. NAP's high participation rate, with more than one third of residents receiving assistance, make it an ideal program for observing the impact of this type of shift.

While a significant body of work has demonstrated the short- and long-term effects of nutritional assistance on consumption patterns, nutrition, crime, and health, this work has focused on the *availability* of nutritional assistance and not the form of the assistance.[2] Research on the effects of the form of the benefit is significantly more limited. Evaluations of randomized "cash out" experiments conducted by the U.S. Department of Agriculture in the 1990s provide mixed evidence, with three out of four sites implying that a change from food stamps to cash benefits led to a reduction in food expenditures of between 18 and 28 cents for each dollar of food stamps cashed out (Fraker, Martini and Ohls, 1995; Ohls and Bernson, 1992; Whitmore, 2002).[3] More recent quasi-experimental evidence from Hastings and Shapiro (2018) indicates that even inframarginal individuals do not treat nutritional assistance benefits as fungible with cash, but instead tend to earmark benefits for food spending. In a very poor and rural context in Mexico, Cunha (2014) evaluates a randomized control trial of the government's Food Assistance Program ('PAL') in which villages were randomly assigned baskets of goods or their cash equivalents. While they find no significant difference in the overall effect of the form of the benefit on total consumption or food consumption, the point estimates suggest that in-kind provision resulted in food consumption that was 40% higher overall, with expenditures on basket items more than 200% higher.

Leveraging earlier variation in Puerto Rico in 1982, Moffitt (1989) uses a structural model

[2]See, for example, Gertler (2004); Fox, Hamilton and Lin (2004); Lee and Mackey-Bilaver (2006); Whitake, Philips and Orzol (2006); Baum (2007); Hoynes and Schanzenbach (2009); Almond, Hoynes and Schanzenbach (2011); Hoynes and Schanzenbach. (2012); Hoynes, Schanzenbach and Almond (2016); Hastings, Kessler and Shapiro (2018); Barr and Smith (2018).

[3]Fraker, Martini and Ohls (1995) argue that the cash out in Alabama had no effect on food expenditures due to the significantly shorter nature of the cash out (8 months versus 4 years) and the provision of food stamp benefits separate from other benefits (in the other sites the "cash" food stamp benefits were combined with other cash transfers such as AFDC).

to find that replacing food stamps with paper checks (equivalent to cash) had no impact of food expenditures. However, this variation differed substantially from the later policy change that we examine. First, as with the Mexican and USDA experiments mentioned above, the generosity of the benefits was low and the transfer was inframarginal for most (92 percent of) households, suggesting that large differences in food expenditures should not be expected, at least under neoclassical assumptions. Second, there was substantial anecdotal evidence that food stamps were already operating as a second currency before the policy change, potentially negating any effect in practice.[4] Third, there is a strong possibility that changing selection into benefit receipt that coincided with the policy change explains the lack of changes in food consumption. During this period, there were substantial reductions in funding for the food stamp program, as well as changes in resource thresholds for eligibility and benefit generosity (Moffitt (1989)).

We make three contributions to this literature. First, we distinguish between effects of the availability of nutritional assistance and the form of the benefit in a context where the benefit is not inframarginal for the majority of recipients.[5] Second, we do so in a large-scale context with a permanent change in benefit structure; 30-40% of Puerto Ricans receive food assistance and the benefit structure changed permanently and island-wide in 2001. The effects of a system-wide and permanent shift may differ from those observed in small scale or short-term experiments if there are general equilibrium effects or the effort required to adjust household spending patterns is sufficiently costly. Finally, we provide the only estimates of which we are aware that the form of nutritional assistance, independent of the amount, can influence the effect of assistance on short- and long-term health.

We use a difference-in-differences strategy to exploit a natural experiment where Puerto Rico

[4] Both Fox, Hamilton and Lin (2004) and Moffitt (1989) observe this point. Moffitt (1989) notes that interviews conducted at the time revealed that merchants accepted food stamps for items other than food, as well as discrepancies between reported food consumption and reported food stamp usage that suggest stamps and cash may have been interchangeable.

[5] This is particularly relevant in thinking about the potential effects of increasing the generosity of in-kind nutritional assistance outside existing ranges in the United States.

implemented requirements that 75% of nutritional assistance benefits be spent on approved food items. We find that this change produced dramatic nutritional improvements, with fruit and vegetable consumption increasing by 0.5 servings per day (20 percent). These nutritional improvements, in turn, yielded some contemporaneous improvements in health; maternal anemia fell by 0.3 to 0.7 pp (13 to 23 percent), but there was no effect on the incidence of low birth weight, a marker of extreme malnourishment. Additionally, we explore whether these contemporaneous effects from the policy change translated to longer-term health improvements by leveraging variation in individuals' childhood exposure to the in-kind benefits. Those who grew up after the shift, and therefore more likely to have had better nutrition in early childhood, are taller and more likely to be normal weight as adolescents.

While our results provide compelling evidence that the form of a nutritional assistance benefit can have considerable effects, the strength of the evidence is limited by the nature of the variation. The estimates rely critically on the assumption that relevant outcomes would have trended similarly in Puerto Rico and comparison states if not for the policy change in Puerto Rico. We address this concern with event studies that show similar trends in most outcomes prior to the policy change. Additionally, we find no effect on maternal diabetes, a slow developing and long-term illness that is unlikely to be affected by short-term improvements in nutrition but does reflect the a priori health of the sample population.

It is important to note that our results do not imply that in-kind benefits are necessarily welfare improving over cash. Rather, they suggest that unconstrained household spending decisions may not prioritize health, particularly of young children. To the extent that recipients of nutritional assistance are also likely to receive government-provided or subsidized healthcare (i.e. Medicaid or individual marketplace subsidies), constraining household decisions with in-kind benefits may yield a substantial fiscal externality for the government in the form of reductions in healthcare spending.

3.2 In-Kind Transition in Puerto Rico's Nutrition Assistance Program

In September 2001, in an effort to align more closely with federal regulations, the Puerto Rican Nutrition Assistance Program (NAP) transitioned from a 100 percent cash redeemable EBT card, to one for which 75 percent of the benefits was required to be spent on approved food items (Trippe et al., 2015). While 25 percent of the benefit remained redeemable in cash, the government for the first time made it clear that 100 percent of the benefit was intended for food.[6] This shift came about as a result of pressure from the federal government to "align NAP more closely with the regulations that governed SNAP; specifically, regulations on what and where benefits could be used" (ADSEF, 2000). This pressure was motivated by a desire to "encourage and enforce the utilization of the benefit as established in Federal law and regulations — only for the acquisition of food". Program administrators expected this change to "dramatically reduce improper use of targeted funds" and "promote better nutrition for... participants" (ADSEF 2001b, p.7).

The newly mandated in-kind component of NAP benefits was not inframarginal for most Puerto Ricans.[7] It accounted for the majority of resources in 72 percent of recipient households, while between 35 and 40 percent of recipient households had *no other source* of income.the average Puerto Rican household spends only 18 percent (and the average welfare recipient spends only 21 percent) of their income on food, this suggests that a large portion of NAP recipients were constrained to spend more on food as a result of the 2001 benefit shift.[8],[9] The potential for trafficking of benefits to circumvent this constraint also was limited by the prior adoption of EBT cards.[10] This differed dramatically from an earlier shift from in-kind to cash benefits in Puerto Rico, where food stamps operated as a "second currency" prior to the shift (Moffitt, 1989).

[6] 25 percent was kept in cash due to concerns that some Puerto Ricans would not be able to access an approved retailer for all of their purchases.

[7] This was not the case prior to the 1980s shift from in-kind to cash benefits analyzed by Moffitt (1989).

[8] Authors' calculations using data from (Trippe et al., 2015).

[9] Average monthly benefits per person in Puerto Rico were $103 in 2003 (compared to $84 in the 50 states) according to the Government Accountability Office. In Puerto Rico, the average benefits per recipient relative to median income is more than 3 times what it is in the 50 states.

[10] Consistent with this, part of the motivation for the 2001 benefit shift was to "dramatically reduce improper use of targeted funds."

Anecdotal evidence also suggests that recipients expected the in-kind constraint to bind, with many claiming that the benefit shift would force them to consume additional food. As one young mother put it "I'm going to have the cupboard full, but I will not have the light to cook what I bought... That's logical, they know it, I do not have any more money... " Indeed, there is ample evidence from a number of sources that her case was not unique.[11] Consistent with recipient expectations, food industry revenue estimates at the time imply a roughly 40 to 60 percent increase in food expenditures for NAP recipients as a result of the benefit shift.[12]

Reports from after the benefit shift also indicate that the food assistance was not inframarginal. Surveys of NAP recipients indicate that despite the requirement to spend 100 percent of their benefit on food (including the 25 percent cash component), only 32 percent report using any of the cash benefit on food, suggesting at least 68 percent were constrained.[13] Similarly, EBT expenditure data from after the benefit shift indicate that while over 60 percent of NAP recipients spend roughly 75 percent of their benefit on food (the in-kind share of the benefit), only 6 percent spend something in between 80 and 99 percent of their benefit on food.[14] This level of bunching suggests that many households were constrained in their expenditures on food.

[11] Gotay, Benjamin Torres. "Empieza hoy el nuevo sistema de uso del PAN." *El Nuevo Dia* 1 Sep. 2001 (translated to English).

[12] The president of the Association of Marketing, Industry and Food Distribution (MIDA) indicated that "the economic boom experienced by the food industry is the result of a fundamental factor: the modifications to the Nutritional Assistance Program (PAN) under the formula 75-25 through the Family Card." A MIDA report suggests that the benefit shift increased food sector revenue by $300 to $400 million, 22.5 to 30 percent of the total PAN budget. Given that roughly 80 percent of benefits were spent on food after the shift, we obtain a rough estimate of the increase in food expenditure among recipients by dividing the $300 to $400 million by the implied amount spent on food prior to the shift, calculated as the amount of total PAN expenditures spent on food (0.8*$1335 million), minus the estimated increase in food sector revenue ($300 to $400 million). Román, Miguel Díaz. "Próspera la industria de alimentos." 22 Oct. 2001 (translated to English).

[13] Rosado-González, R., Puerto Rico Department of the Family, Administration for Socioeconomic Development of the Family (ADSEF). (2008). PANECO pregunta. Trujillo Alto, PR: RRG Universe and Assoc.

[14] Trippe et al. (2015) "Examination of Cash Nutrition Assistance Program Benefits in Puerto Rico." Prepared by Insight Policy Research under Contract No. AG-3198-C-14-0006. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

3.3 Data

We explore the contemporaneous nutritional effects of the shift from cash to in-kind NAP benefits using the Behavioral Risk Factor Surveillance System (BRFSS). Following Almond, Hoynes, and Schanzenbach (2011), we use Vital Statistics Natality data to explore whether any observed nutritional improvements translate into health improvements during the critical window surrounding childbirth. Finally, we use the Youth Risk Behavior Survey (YRBS) to look for long run health effects on teenagers who experienced early childhood exposure to in-kind rather than cash NAP benefits.

3.3.1 Behavioral Risk Factor Surveillance System

The BRFSS is a health survey collected by the CDC through phone interviews. It is the largest continuously collected health survey in the world, and collects information on health-related behaviors, conditions, and services. The survey asks a limited set of nutrition-focused food consumption questions, which we use as proxies for food consumption as a whole. Our key food consumption outcomes are (1) daily servings of fruits and vegetables, and (2) daily servings of fruit juice.[15],[16],[17] Our sample includes the BRFSS data collected from all 50 states and Puerto Rico in 1998, 2000, 2002, 2003, and 2005; these are the years in which the key questions about fruit and vegetable consumption were included in the core survey. There are an average of over 230,000 observations per year, with roughly 4,000 per year in Puerto Rico.

[15]The "daily servings of fruits and vegetables" variable is created by the CDC and based on individuals combined reported consumption of fruit (excluding juice), potatoes, carrots, green salad, and all other vegetables.

[16]Although fruit juice consumption is sometimes reported in longer time horizons in the survey, we have adjusted all consumption variables to the level of daily intake.

[17]To account for a small number of unrealistic outliers in reported consumption, we top code each consumption response to be equal to the 99th percentile response for that variable. Respondents who refused to answer consumption questions or indicated that they did not know how much they consumed were coded as missing; these missing observations account for 3% of the data.

3.3.2 Vital Statistics Natality Data

We use Vital Statistics Natality data for the same period.[18] The Vital Statistics Natality data are collected from birth certificates of children born in every state (and Puerto Rico) by the CDC. Our main maternal and infant health outcomes are binary indicators for infant low birth weight, mother's diabetes, and mother's anemia. Anemia is ex-ante the most likely outcome to reveal a positive health impact from nutrition changes produced by the switch to in-kind benefits. Iron deficiency is the most common cause of anemia worldwide and dietary changes, specifically increases in consumption of iron rich foods or use of iron supplements, are the most common form of iron-deficiency anemia prevention and treatment (Habershon, 1863; Camaschella, 2015). The incidence of low birth weights could also be affected by the changes in food consumption, but is less likely in this setting because of the relatively low levels of extreme malnourishment in Puerto Rico in recent decades.[19] Finally, we examine diabetes as a placebo test. Diabetes is generally understood to be a slow developing, progressive, and long-term illness that is closely linked with long-term over-consumption of certain food. It is therefore unlikely that constrained increases in food consumption would have any effect on diabetes in the short run.[20]

The data also include month and year of birth, state or territory of birth, mother's race, mother's education, plurality of birth, and order of birth, which we use as control variables. There are an average of 2 million observations per year, with an average of roughly 55,000 births per year occurring in Puerto Rico.

[18]Vital Statistics Natality data doesn't include any state level geographic identifiers after 2004.

[19]Risk of low birth weight has been shown to increase when mothers suffer from extremely poor nutrition when they become pregnant or from caloric deprivation during the third trimester of pregnancy. But, the same studies found no effect of moderately poor nutrition or caloric deprivation in the first two trimesters (de Bernabé et al., 2004).

[20]While a balanced diet can decrease the risk of developing diabetes, the largest benefits come from decreasing consumption of saturated fat and cholesterol (Schoenaker et al., 2016). Decreases in total energy intake has even had some success at reversing diabetes (Lim et al., 2011).

3.3.3 Youth Risk Behavior Survey

The Youth Risk Behavior Survey (YRBS) is a biannual CDC survey of high school students across the United States. It collects information on the health and behaviors of adolescents and contains an average of 89,000 observations per survey year during our sample period, with roughly 2,300 observations occurring in each survey year in Puerto Rico.[21] We use the 2009-2017 surveys and focus our analyses on individuals born between 1994 and 2003. Our key outcomes of interest are student height and a binary variable indicating if students are normal weight for their age and gender.[22]

3.3.4 Summary Statistics and Data Limitations

Table 4.1 displays summary statistics from each of the datasets mentioned above. It shows some substantial differences between Puerto Rico and comparison states. While infant and adolescent characteristics are similar, levels of income and education, as well as fruit and vegetable consumption are much lower in Puerto Rico, even relative to the poorest states. These level differences do not affect the internal validity of our difference-in-differences approach so long as the trends in these outcome variables do not differ substantially. However, the extent to which the marginal benefit of additional nutrition is greater in Puerto Rico than in other states, due to these level differences, will affect the generalizability of our results to other contexts.

An important limitation of our data sources is that none of them allow us to directly identify those eligible for or enrolled in Puerto Rico's NAP program. Therefore, our main specifications use the full sample of data available in each year, which means that many Puerto Rican individuals are included who did not participate in NAP. As a result, we estimate an average treatment effect for all Puerto Ricans that incorporates both direct effects on the treated as well as potential spillover effects on non-participants. Spillover effects are likely given the high participation rate

[21] Not every state is included in each survey year, but at least 44 states are included in each year of our sample. Puerto Rico is included in the 2009, 2015, and 2017 surveys.

[22] These calculations are based off BMI and are equivalent to measuring the probability that a student is not underweight, overweight, or obese.

(30-40%) of Puerto Ricans in NAP.

3.4 Empirical Strategy

We explore how exposure to in-kind benefits relative to equivalent cash benefits affects food consumption and health. We use a difference-in-differences design, estimating the following equation:

$$F_{ist} = \alpha_s + \lambda_t + \beta(PR_s * Post_t) + \gamma X_{ist} + \epsilon_{ist}, \tag{3.1}$$

where F_{ist} is a measures of the food consumption or health of individual i in state (or territory) s in year t. The terms α_s and λ_t are state and year fixed effects. X_{ist} are individual covariates, including age indicators interacted with gender for regressions with BRFSS and YRBS data, and mother's race, plurality of birth, and birth order for regressions with natality data. $PR_s * Post_t$ is an interaction term that equals one for observations in Puerto Rico after Puerto Rico's nutrition assistance program began requiring benefits to be spent on approved food items.[23]

In the analysis of adolescent outcomes, we follow Hoynes, Schanzenbach and Almond (2016), who find that food stamps has the largest long run effects for children exposed from age 0-5. We define treatment as the fraction of the first 5 years of life that the adolescent spent under the in-kind benefits policy. In other words, $PR_s * Post_t$ is replaced with $Frac5_t$, where $Frac5_t$ is the estimated fraction of an individual's life from birth to age five that occurred after the in-kind benefits policy change. This variable is zero for untreated Puerto Ricans (born before 1996), between zero and one for partially treated Puerto Ricans (born 1996-2001), and equal to one for Puerto Ricans who were exposed to the policy from birth (born after 2001). Standard errors are clustered at the state/territory level.

The coefficient of interest, β, provides an estimate of the effect of in-kind benefits, relative

[23] The official change happened in September 2001, but there was a ramp up period through the end of 2001. Given this and data constraints (2001 is not available in the BRFSS and YRBS data is limited to year of birth), we set 2002 as the first post year.

to cash benefits, on each outcome variable. These estimates can be interpreted as causal if Puerto Rico and the comparison states would have maintained similar trends if not for Puerto Rico's policy change. We assess this assumption for each outcome with event studies that show year-by-year differences between Puerto Rico and the comparison states (controlling for individual covariates and state/territory and year fixed effects). We find no evidence of differential trends prior to the policy change.

We also conduct supplementary analyses using a synthetic control approach, which constructs the comparison group from a weighted average of the outcomes of other states. Puerto Rico has lower levels of fruit and vegetable consumption than the rest of the United States, limiting our ability to conduct a standard synthetic control analysis since no linear combination of states is equivalent to Puerto Rico. To overcome this, we de-mean the data using each state's pre-period outcome levels, and then create the synthetic control based on those demeaned outcomes.[24] Synthetic control estimates of food consumption, based on each state's deviation from the mean, are similar in magnitude to our main results. We discuss this estimation strategy and the results in more detail in Appendix 3.8.1.

3.4.1 Threats to Internal Validity

The primary internal validity concern for our empirical strategy is that Puerto Rico implemented the benefit change when food consumption was increasing and Puerto Rican's were becoming healthier for some other reason. For example, Puerto Rico may have adopted the benefit shift as part of a larger initiative to improve a variety of services for mothers and young children. If this were the case, we might observe improved nutrition and health due to a comprehensive effort to help these cohorts and not because of the benefit shift. However, the cause of the shift was not internally motivated. In fact, it came about as a result of pressure from the U.S. mainland to "align NAP more closely with the regulations that governed SNAP; specifically, regulations on what and where benefits could be used" (ADSEF, 2000). This pressure was motivated by a desire to "en-

[24]We subtract each state's pre-2001 average consumption levels from each observation from that state.

courage and enforce the utilization of the benefit as established in Federal law and regulations – only for the acquisition of food." The externally motivated benefit shift is consistent with the evidence provided in Figure 3.1, which demonstrates a flat trend in fruit and vegetable consumption followed by a sharp jump between 2000 and 2002.

Still, it is possible that the benefit shift coincided with another shock that generated improvements in nutrition and health around or after the point of the benefit shift. To address this concern we conducted an extensive review of Puerto Rican policies and events during this time period.[25] We uncovered few policy shifts or events that seem likely to have generated the observed results. The greatest potential confounds are the 2001 recession and migration out of Puerto Rico.

While the 2001 recession coincides with the timing of the benefit shift, it seems unlikely to have generated our results as it had similar effects on Puerto Rico and our sets of comparison states. If anything, the recession was somewhat more pronounced in Puerto Rico, which we would expect to negatively affect nutrition and birth outcomes. Regardless, the recession was rather short lived and thus seems unlikely to account for the persistent effects we observe.

Another potential source of confounding variation is differential migration out of Puerto Rico. Migration out of Puerto Rico could yield spurious estimates of improvements in nutrition and health if impoverished families or mothers began leaving the island in greater numbers around the time of the benefit shift. While out-migration did increase somewhat in the early 2000s, the numbers can account for only a tiny fraction of our observed effects. Further, there was no "jump" in out migration that occurred around 2001 that could account for the observed improvements in food consumption and birth outcomes.

3.5 Results

We employ the difference-in-difference strategy discussed above to estimate the impact of the form of nutritional benefits on food consumption, mother and infant health, and adolescent health

[25]This task was undertaken with the assistance of a Puerto Rican legal researcher with extensive experience searching Spanish-language periodicals and reports.

outcomes using a variety of data sources. We find that the transition to in-kind benefits in Puerto Rico improved measures of nutrition and mother's health, and yielded long-run effects on the adolescent health of those who experienced early childhood after the shift in benefits.

3.5.1 Effects on Food Consumption

We find evidence of substantial improvements in nutrition from in-kind relative to cash benefits. While we observe limited measures of food consumption, these results are consistent with economic theory and anecdotal evidence. In columns 1 and 2 of Table 3.2, we present estimates of Equation 3.1 using data from BRFSS. We find that fruit and vegetable consumption increased by 0.53-0.56 servings per day (21%) as a result of the transition to in-kind benefits, while fruit juice consumption increased by 0.23 servings per day (28%). These results are robust to changes in the construction of the comparison group, as well as the use of the synthetic control method (method (Appendix Figures 3.11-3.14). The internal validity of our difference-in-differences strategy is supported by event studies that demonstrate a flat (or slightly downward) trend in fruit and vegetable and fruit juice consumption prior the benefit transition, and a large increase just afterward (Figure 3.1).

The large magnitudes of these nutrition effects is consistent with the large increases in revenue reported by food industry in Puerto Rico after the transition to in-kind benefits. We can inflate our estimates by the fraction of the population enrolled in NAP to estimate the effect on consumption among benefit recipients. Given that 30-40 percent of Puerto Ricans were receiving benefits during this time period, our estimates imply an increase of 1.5 servings of fruit and vegetables per day, which represents a consumption increase of over 50 percent, within the range of the 40 to 60 percent increase in food expenditures expected based on the industry's revenue changes.[26]

[26]We return to a discussion of magnitude below.

3.5.2 Contemporaneous Effects on Health

We use natality data to explore whether the nutritional improvements from the transition to in-kind benefits were substantial enough to yield observable improvements in the contemporaneous health of mothers and newborn children. Specifically, we estimate the effect of the transition to in-kind benefits on the incidence of diabetes and anemia among mothers and the incidence of low birth weight among newborn children. As discussed above, anemia is ex-ante the most likely outcome to reveal a positive health impact from the switch to in-kind benefits, given the prevalence of iron deficiency and the short-term responsiveness to dietary changes. The incidence of low birth weights could also be affected by changes in food consumption, but this is less likely to be observed, given the relatively low levels of extreme malnourishment in Puerto Rico during the 1990s and 2000s.

In columns 3 and 4 of Table 3.2, we present estimates of Equation 3.1 for low birth weight and maternal anemia. We find no robust evidence of significant effects on low birth weight, but a substantial reduction in maternal anemia of 0.35-0.67 pp (13-23 percent). Table 3.3 shows that these results are robust to the inclusion of birth and mother covariates. As with the nutrition results, the validity of the maternal anemia estimates is supported by event studies that demonstrate a flat trend in maternal anemia prior to the benefit transition and a relative decrease in Puerto Rican maternal anemia afterward (Figures 3.17, and 3.18).[27] The event studies for low birthweight suggest caution in interpreting the difference-in-difference estimates for this outcome, given the lack of parallel trends in the pre-period (Figures 3.17, and 3.18).

In columns 5 of Table 3.2, we present estimates of Equation 3.1 for maternal diabetes primarily as a placebo test. While diabetes reflects underlying nutrition and health, it is generally understood to be a slow developing, progressive, and long term illness and is unlikely to be affected by short-

[27] Due to data restrictions, we are not able to explore heterogeneous treatment effects by education level in the natality data. The natality data includes information on mothers' education level, but a number of states changed the coding of these variables in 2003 and 2004. Because the timing of the change coincides with the post treatment period in this study, we don't conduct any health effect analysis by education subgroups.

term changes in food consumption. An estimated effect on maternal diabetes would therefore likely reflect differences in the composition of mothers rather than a change in nutrition from the transition to in-kind benefits. We find no effect on maternal diabetes, providing additional evidence that the transition to in-kind benefits did, in fact, produce the estimated effects on mother's anemia.

3.5.3 Long-Run Health Effects on Children

Given the observed contemporaneous effects of the transition to in-kind benefits on food consumption and maternal health, a natural question is whether these effects persist. If so, this could suggest an important role for nutrition in explaining the long-run effects on young children of the rollout of food stamps observed by Hoynes, Schanzenbach and Almond (2016). The YRBS data on high school students age 14-18 enable us to explore the long-run health effects that the transition to in-kind benefits had on young children. Following Hoynes, Schanzenbach and Almond (2016), we define our treatment variable as the proportion of time an individual was exposed to the in-kind benefits policy from age 0-5. Columns 6 and 7 of Table 3.2 show our results using this exposure measure. We find that additional exposure to in-kind benefits in early childhood results in increased height and the likelihood of being classified as normal weight in adolescence. For each year of in-kind benefit exposure from age 0 to 5, height increases by 0.04 to 0.06 inches and the likelihood of being normal weight in high school increases 1 pp. Relative to those who never received in-kind benefits during this critical period, exposure for the first five years of life results in an increase in adolescent height of $\frac{1}{6}$ to $\frac{1}{3}$ of an inch and a 5pp (7%) increase in normal weight.

Figures 3.7 and 3.10 provide graphical evidence of the effects, demonstrating the relationship between the age at exposure to in-kind benefits and height or normal weight in adolescence. Given the nature of treatment, the presentation is somewhat non-standard, following Hoynes, Schanzenbach and Almond (2016). The horizontal axis presents the number of years between the transition to in-kind benefits (2001) and an individual's year of birth. Those individuals with a value of 0 or less are "fully treated" in that the in-kind benefits were available from the year of their birth onward. Moving to the right, the age at transition to in-kind benefits increases, and therefore childhood in-kind benefit exposure decreases. As observed in the figure, the earlier in an individual's life that the transition to in-kind benefits occurred (and nutrition improved), the larger the increase in height or likelihood of being of normal weight in adolescence. The effects on height and weight in adolescence are largest at or prior to conception and decrease between conception and age 5. Consistent with our estimates representing a causal effect of the transition to in-kind benefit provision, the timing of the transition prior to conception has no effect on the size of the reduction (i.e., the effect of the availability of in-kind benefits is the same for those born one or two years after the transition).[28]

When combined with previous results, these results suggest that increases in food consumption during early childhood translate into improved health through adolescence.

3.5.4 Magnitude of Effects

In the absence of spillover effects on non-recipients from the transition to in-kind benefits, the effect of the transition on recipients could be obtained by inflating our results by the fraction of the population enrolled in NAP to estimate average effects for recipients. Given that 30-40 percent of Puerto Ricans were receiving benefits during this time period, this implies an average treatment effect for recipients of 1.5 servings of fruits and vegetables per day, which represents an over 50 percent increase in consumption. Similar calculations suggest reductions in maternal anemia of 1-1.7 percentage points (30-50 percent) as well as substantial increases in height (7-12 percent) and the likelihood of being normal weight (14-18 percent) among recipients.

However, it is unlikely that those receiving NAP were the only ones affected by the transition to in-kind benefits. Puerto Rico is more densely populated than any state, and has a culture that puts a strong focus on family and community. Since the shift in NAP policy constrained roughly 1 out of every 3 people to purchase more food, we expect that a non-trivial portion of that food would be

[28]Unfortunately, the timing of the policy change and the availability of data limits our ability to explore effects on cohorts born more than one or two years after the transition.

shared with family and neighbors, either through direct gifts of food or by sharing prepared meals with individuals outside the household.

It is also possible that the policy change had substantial general equilibrium effects, such as impacting food prices, the types of foods supplied/consumed, and the location of sellers. Indeed, anecdotal evidence suggests that the policy change had meaningful implications for food distributors across Puerto Rico. One newspaper wrote that,"the commercial food sector, which lobbied intensely [in favor of the in-kind restrictions], has registered significant increases in sales, product of the captive market that provide... 75% of the $1.8 billion that the PAN distributes annually in Puerto Rico. A specific estimate of how much sales have risen was not available, but the head of the Socioeconomic Development Administration (Adsef) of the Family, Gretchen Coll, says it is 'very much.' "The large-scale shift may have led to broader changes in the availability of different foods and/or consumption patterns.

Heterogeneity in the effect of the shift in NAP policy by education level provides further support for spillovers onto non-recipients. In Table 3.7, we find meaningful increases in food consumption at every education level, including individuals with college degrees who have low rates of NAP eligibility (though the magnitudes of these effects are smaller than for lower levels of education).[29] Even after accounting for higher average consumption of fruits and vegetables among individuals college graduates, these estimates imply a TOT that is twice as large for college graduates as would be expected if NAP recipients in this group were affected similarly as NAP recipients with only a high school diploma. This suggests a substantial spillover of the NAP policy change onto non-recipients with higher levels of education.

The presence of spillover effects has important implications for how we interpret the reduced form and scaled effects of the shift to in-kind benefits. While we discuss above the average treatment effects for recipients implied by NAP participation rates, these inflated estimates are only

[29]Estimates using the 2000 Puerto Rico census suggest that about 13% of those with bachelor's degrees had incomes below the federal poverty line (FPL), compared to 48% of those with only a high-school degree and 63% of high school drop outs.

accurate under the assumption of no spillover effects and are biased upward when spillovers are present. Because spillovers are likely, our main specifications and results focus on the reduced form effect of the policy shift for the population, accounting for direct effects for NAP recipients as well as any spillover and general equilibrium impacts.

3.6 Discussion and Conclusion

While prior evidence suggests that the FSP increased food consumption and improved short- and long-term health outcomes, it is not clear whether these effects were driven, at least partially, by constraining households' consumption decisions, or whether an equivalent increase in income would have generated the same effects. To shed light on this question, we leverage a natural experiment where Puerto Rico converted a cash benefit to one in which recipients were required to spend 75% of their benefits on approved food items. This allows us to examine the impact of constraining household consumption decisions in the absence of any shock to overall income.

Using a difference-in-differences strategy, we find that providing the benefits in-kind increases fruit and vegetable consumption by 0.5 servings per day (20 percent). These improvements in nutrition led to decreases in maternal anemia of 0.3-0.7 pp (13-23 percent) and resulted in those who grew up after the shift being taller and more likely to be normal weight as adolescents.

Our results conflict somewhat with a body of work that suggest more modest effects of the form of nutritional assistance on food expenditures or health. Most of these results can be reconciled when one considers the degree to which program participants' food consumption levels are constrained by the generosity of the in-kind benefit. Given the lack of a binding constraint and the subsequent minimal effects on food consumption and nutrition it is perhaps not surprising that prior studies were unable to detect effects on health. Unlike in most prior cash out evaluations, the majority of benefit recipients in Puerto Rico were constrained by the shift to in-kind benefits. This resulted in large increases in food consumption and subsequent improvements in health.

While the results provide compelling evidence that the form of the benefit can matter, the analyses are not without limitations. First, the strength of the evidence is limited by the nature of the variation. The estimates rely critically on a comparison of outcomes across time in Puerto Rico and how those outcomes evolved relative to outcomes in sets of comparison states in the U.S. While we are unaware of other policy changes or events that could have generated these effects, we present the results with this caution in mind. Second, it is important to emphasize that the results do not imply that in-kind benefits are welfare improving over cash, but rather that individual spending may not prioritize health, particularly of young children, under a cash-based system. This may have additional implications for the long run costs of cash vs in-kind benefits, particularly when low income individuals have medical costs that are covered or subsidized by the government. A more holistic understanding of the costs and benefits of provision of benefits in-kind or in cash is outside the scope of this work.

3.7 Figures and Tables

Table 3.1: Summary Statistics

	(1) All Comparison States	(2) 10 Poorest States	(3) Puerto Rico
Panel A: BRFSS			
Female	0.60	0.63	0.64
Age	48.68	48.77	49.15
Daily Servings of Fruit and Vegetables	3.83	3.71	2.97
Daily Servings of Fruit Juice	0.65	0.61	0.90
Income Below 10k	0.05	0.07	0.36
Income Above 50k	0.35	0.29	0.05
Graduated High School	0.96	0.93	0.78
Graduated College	0.31	0.26	0.22
Observations	1,829,113	308,772	29,924
Panel B: Natality			
Mother's Diabetes	0.03	0.03	0.02
Mother's Anemia	0.02	0.03	0.03
Low Birth Weight	0.08	0.09	0.11
Plural Birth	0.03	0.03	0.02
Observations	27,591,343	3,574,871	389,729
Panel C: YRBS			
Female	0.51	0.52	0.50
Age	15.77	15.79	15.74
Hispanic	0.19	0.18	0.20
Weight (lbs)	146.91	150.45	147.04
Normal Weight	0.67	0.62	0.65
Height (in)	66.57	66.60	66.33
Observations	449,867	57,916	7,412

Note: Table presents descriptive statistics for each data set used in analysis. Statistics are shown separately for all states, the 10 poorest states, and Puerto Rico as indicated by column titles.

Table 3.2: Main Outcomes

	BRFSS		Natality			YRBS	
(lr)2-3(lr)4-6(lr)7-8	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Fruit and Vegetables	Fruit Juice	Low Birth Weight	Mother's Anemia	Mother's Diabetes	Normal Weight	Height in Inches
Panel A: All States							
PR*Post	0.5306***	0.2290***	0.0005*	-0.0035***	-0.0006	0.0385***	0.1904***
	(0.0167)	(0.0046)	(0.0003)	(0.0008)	(0.0006)	(0.0034)	(0.0464)
Observations	1,236,218	1,219,453	27,555,353	26,350,409	27,294,011	454,397	454,397
Panel B: 10 Poorest States							
PR*Post	0.5576***	0.2253***	-0.0008	-0.0067***	-0.0001	0.0514***	0.3216**
	(0.0546)	(0.0120)	(0.0007)	(0.0016)	(0.0007)	(0.0054)	(0.1021)
Observations	222,938	217,407	3,924,711	3,769,091	3,899,538	57,501	57,501
PR Pre-Treatment Average	2.67	0.81	.11	.03	.02	0.68	65.46

Note: Each coefficient is the result of a separate regression. Panel A displays the results when all states are included in the comparison group, while panel B restricts the comparison group to include only the 10 poorest states. Outcomes are indicated by column titles, with the data source for each outcome indicated by its multi-column header. All standard errors are clustered at the state level. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 3.3: Natality - With or without controls

	(1) No Controls	(2) All Controls
Panel A: Low Birthweight		
Low Birthweight: Poor States	-0.0003	-0.0008
	(0.0009)	(0.0007)
Low Birthweight: All States	0.0004*	0.0005*
	(0.0002)	(0.0003)
Panel B: Anemia		
Anemia: Poor States	-0.0068***	-0.0067***
	(0.0016)	(0.0016)
Anemia: All States	-0.0037***	-0.0035***
	(0.0008)	(0.0008)
Panel B: Diabetes		
Diabetes: Poor States	-0.0018	-0.0001
	(0.0012)	(0.0007)
Diabetes: All States	-0.0016**	-0.0006
	(0.0008)	(0.0006)

Note: Each coefficient is the result of a unique regression, where the outcome variable and comparison set is indicated by the row label. Both columns include state and year fixed effects. The second column also includes all controls used in our main analysis (education level, birth month, race, plural birth, and birth order). Standard errors are clustered at the state level. Significance levels indicated by: * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

Table 3.4: Natality by Marital Status

	(1) Low Birth Weight	(2) Anemia	(3) Diabetes
Panel A: All States			
PR*Post*Married	-0.001***	-0.007***	0.001
	(0.000)	(0.001)	(0.001)
PR*Post*(Married or Cohabiting)	-0.001***	-0.004***	-0.001
	(0.000)	(0.001)	(0.001)
PR*Post*(Not Married)	0.002***	-0.000	-0.001*
	(0.001)	(0.001)	(0.000)
PR*Post*(Not Married or Cohabiting)	0.005***	-0.000	-0.000
	(0.001)	(0.001)	(0.000)
Panel B: 10 Poorest States			
PR*Post*Married	-0.002**	-0.009***	0.001
	(0.001)	(0.001)	(0.001)
PR*Post*(Married or Cohabiting)	-0.001*	-0.007***	-0.000
	(0.001)	(0.001)	(0.001)
PR*Post*(Not Married)	0.000	-0.004	-0.001
	(0.001)	(0.003)	(0.001)
PR*Post*(Not Married or Cohabiting)	0.002*	-0.004	0.000
	(0.001)	(0.003)	(0.001)
PR Pre-Treatment Average	.11	.03	.02

Note: This table shows Natality results by marital status. Each coefficient is the result of a separate regression. Outcomes, comparison groups, and marital status are indicated by column titles, panel labels, and row labels respectively. All standard errors are clustered at the state level. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 3.5: BRFSS by Marital Status

	(1) All	(2) Married	(3) Married or Cohabiting	(4) Not Married	(5) Not Married or Cohabiting
Panel A: All States					
Daily Servings of Fruits & Vegetables	0.531***	0.445***	0.458***	0.637***	0.480***
	(0.017)	(0.017)	(0.016)	(0.019)	(0.017)
Daily Servings of Fruit Juice	0.229***	0.219***	0.225***	0.243***	0.221***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Observations	1,219,453	666,219	696,818	549,617	1,017,215
Panel B: 10 Poorest States					
Daily Servings of Fruits & Vegetables	0.558***	0.472***	0.480***	0.662***	0.504***
	(0.055)	(0.045)	(0.046)	(0.066)	(0.052)
Daily Servings of Fruit Juice	0.225***	0.207***	0.215***	0.250***	0.215***
	(0.012)	(0.011)	(0.011)	(0.016)	(0.012)
Observations	217,407	114,749	118,974	102,113	180,736

Note: Significance levels indicated by: * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

3.7.1 BRFSS Figures

Figure 3.1: Fruit Vegetables - All Education Levels - All States

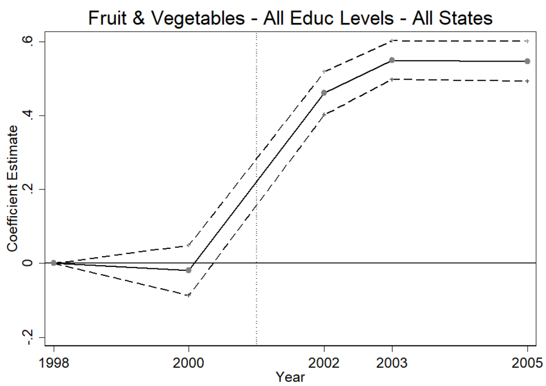


Figure 3.2: Fruit Juice - All Education Levels - All States

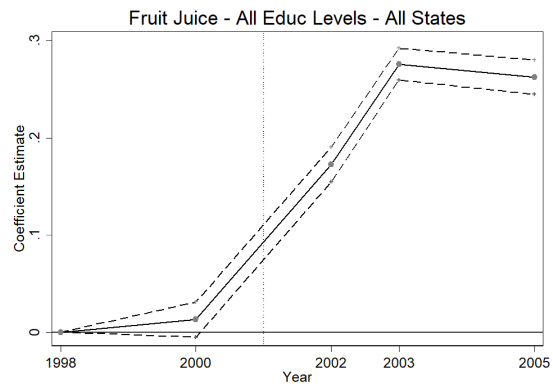


Figure 3.3: Fruit Vegetables - All Education Levels - Poorest States

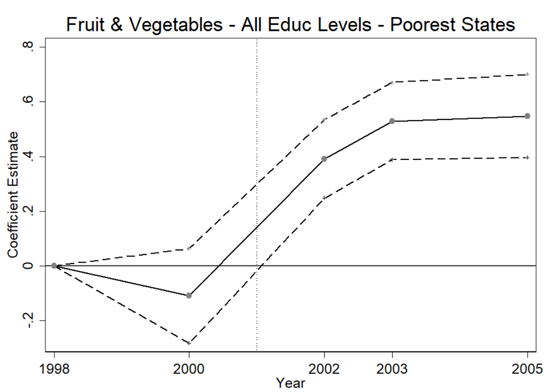
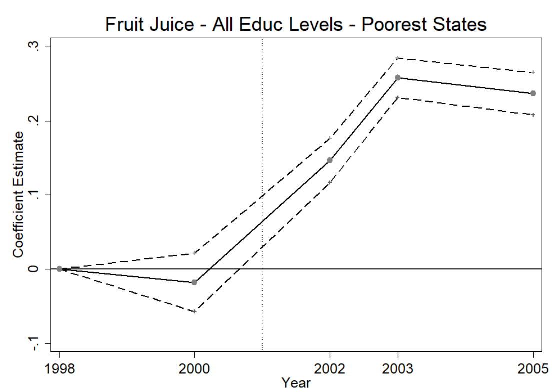


Figure 3.4: Fruit Juice - All Education Levels - Poorest States



3.7.2 Natality Figures

Figure 3.5: Natality Outcomes: All States

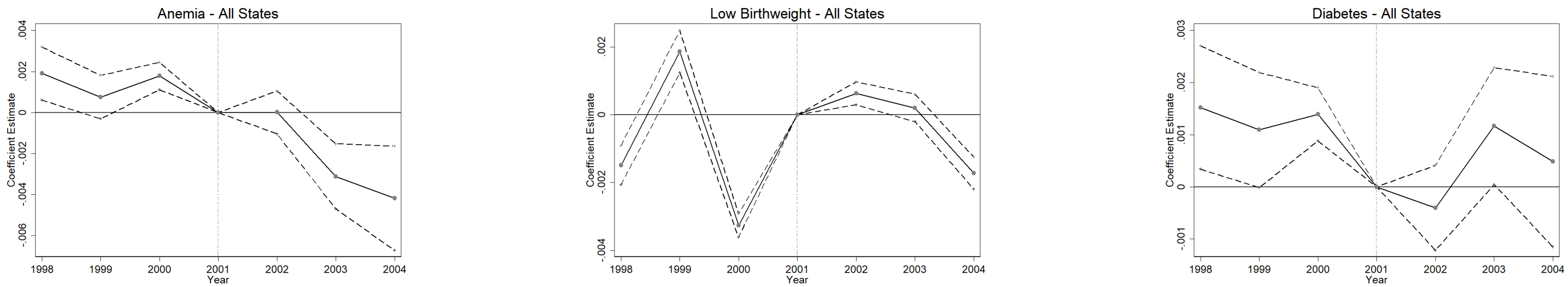
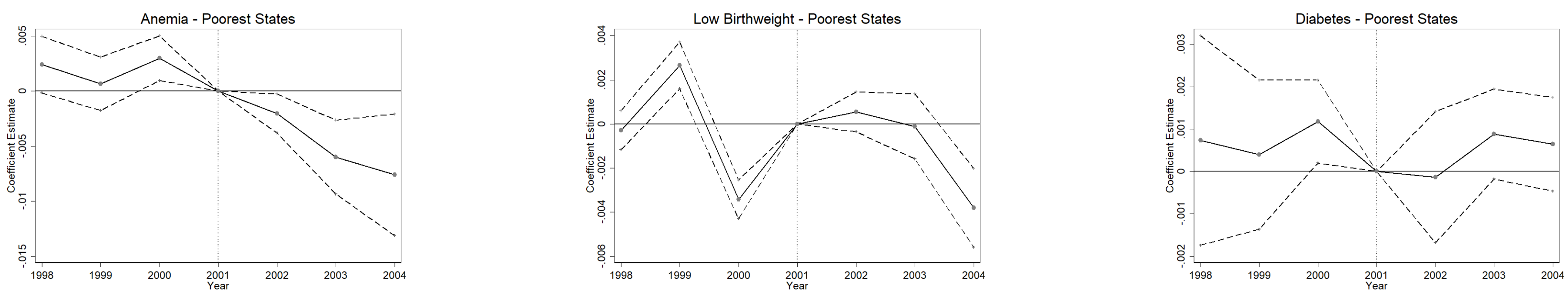


Figure 3.6: Natality Outcomes: 10 Poorest States



3.7.3 YRBS Figures

Figure 3.7: Height - All States

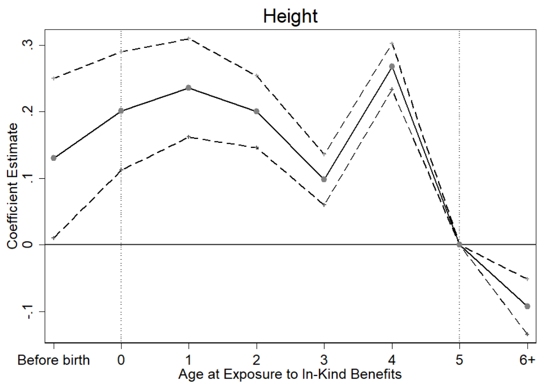


Figure 3.8: Normal Weight - All States

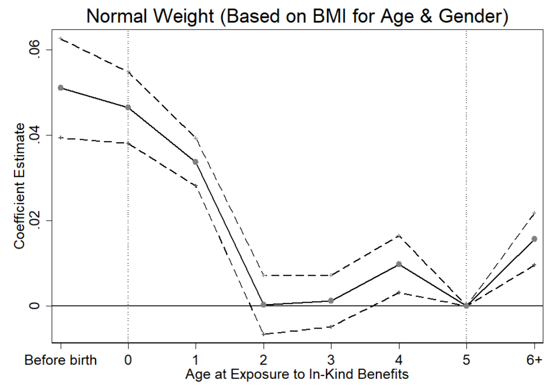
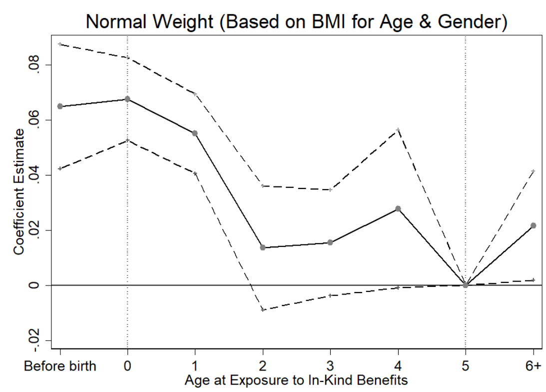


Figure 3.9: Height - 10 Poorest States

Figure 3.10: Normal Weight - 10 Poorest States



3.8 Additional Tables and Figures

Table 3.6: BRFSS Results by Education Level

	(1) All	(2) Dropout	(3) High School Diploma	(4) Some College	(5) Bachelor's Degree
Panel A: All States					
Daily Servings of Fruits & Vegetables	0.531***	0.587***	0.623***	0.559***	0.415***
	(0.017)	(0.019)	(0.017)	(0.018)	(0.020)
Percent Change	19.9	24.3	24.3	20.7	13.3
Daily Servings of Fruit Juice	0.229***	0.183***	0.269***	0.236***	0.216***
	(0.005)	(0.008)	(0.005)	(0.006)	(0.006)
Percent Change	28.3	25.4	33.3	27.8	24.5
Observations	1,219,453	134,941	380,692	329,336	371,905
Panel B: 10 Poorest States					
Daily Servings of Fruits & Vegetables	0.558***	0.633***	0.638***	0.567***	0.452***
	(0.055)	(0.030)	(0.051)	(0.051)	(0.080)
Percent Change	20.9	26.2	24.9	21.0	14.5
Daily Servings of Fruit Juice	0.225***	0.202***	0.267***	0.226***	0.205***
	(0.012)	(0.012)	(0.011)	(0.016)	(0.018)
Percent Change	27.8	28.1	33.0	26.6	23.3
Observations	217,407	37,256	71,225	52,943	55,463
PR Pre-Treatment Average	0.81	0.72	0.81	0.85	0.88

Note: Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 3.7: Natality - Composition of Births

	Education				Race		Other	
(lr)2-5(lr)6-7(lr)8-9	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	No HS Diploma	HS Diploma	Some College	Earned BA	White	Black	First Child	Number of Births
Panel A: All States								
PR*Post	-0.032***	0.024***	0.010***	-0.003	-0.009***	0.016***	0.013***	-8780.438***
	(0.003)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(625.515)
Observations	21,461,549	21,754,521	21,461,549	21,461,549	27,981,072	27,981,072	27,836,044	364
Panel B: 10 Poorest States								
PR*Post	-0.034***	0.028***	0.004**	0.002	-0.019***	0.024***	0.019***	-7723.250***
	(0.005)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(874.788)
Observations	3,027,519	3,062,206	3,027,519	3,027,519	3,062,206	3,062,206	3,054,231	77
PR Pre-Treatment Average	.27	.29	.23	.21	.92	.08	.43	58790

Note: Each coefficient is the result of a unique regression, where the outcome variable is indicated by the column title. Regressions include state and year fixed effects and no other controls. Regressions for education outcomes exclude 9 comparison states that adjusted their coding of education categories in the post period. Standard errors are clustered at the state level. Significance levels indicated by: * ($p<0.10$) **($p<0.05$), ***($p<0.01$).

3.8.1 Synthetic Control

Table 3.8: Synthetic Control - Main Outcomes

	BRFSS			Natality	
(lr)2-3(lr)4-6	(1) Fruit and Vegetables	(2) Fruit Juice	(3) Low Birth Weight	(4) Mother's Anemia	(5) Mother's Diabetes
Demeaned Outcomes					
Puerto*Post	0.480	0.232***	0.003	-0.002	-0.001
	(0.115)	(0.038)	(0.885)	(0.256)	(0.269)
	[0.019]	[0.019]	[0.192]	[0.605]	[.904]
Not Demeaned					
Puerto*Post			0.005	-0.004*	-0.003
			(0.788)	(0.093)	(0.192)
			[0.038]	[0.349]	[0.596]
PR Pre-Treatment Average	2.7	0.8	.11	.03	.02

Note: Natality post period is defined as starting 2002. P-Values (Calculated from Post-treatment RMSE divided by Pre-treatment RMSE) are in parentheses and significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01). Standard Random Inference P-Values are in brackets.

3.8.2 BRFSS Results

Figure 3.11: Synthetic Control - Fruit and Vegetable Servings

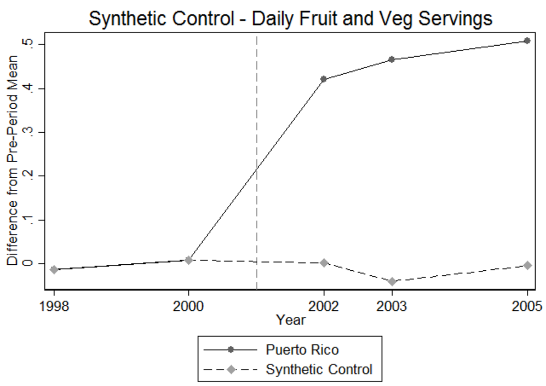


Figure 3.12: Synthetic Control - Fruit Juice

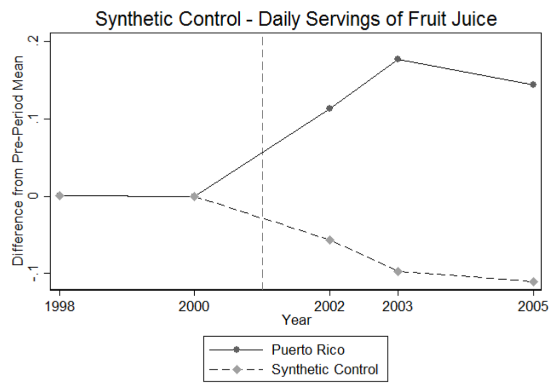


Figure 3.13: Randomization Inference - Fruit and Vegetable Servings

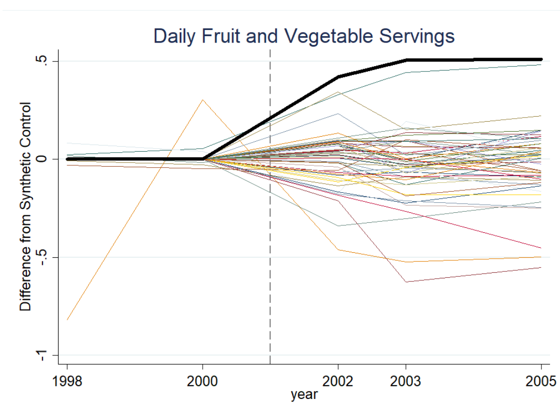
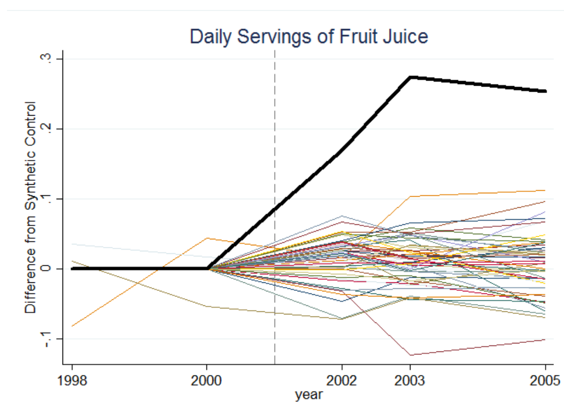


Figure 3.14: Randomization Inference - Fruit Juice



3.8.3 Natality Results

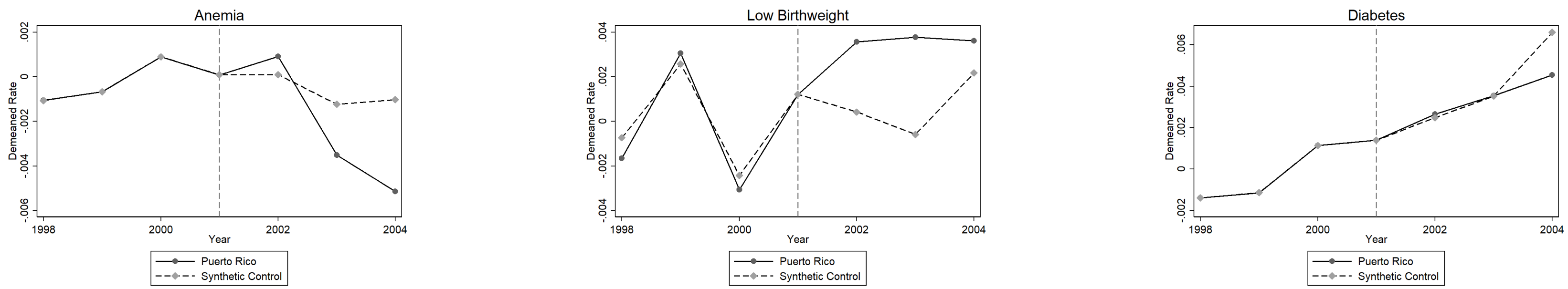


Figure 3.15: Demeaned Outcomes

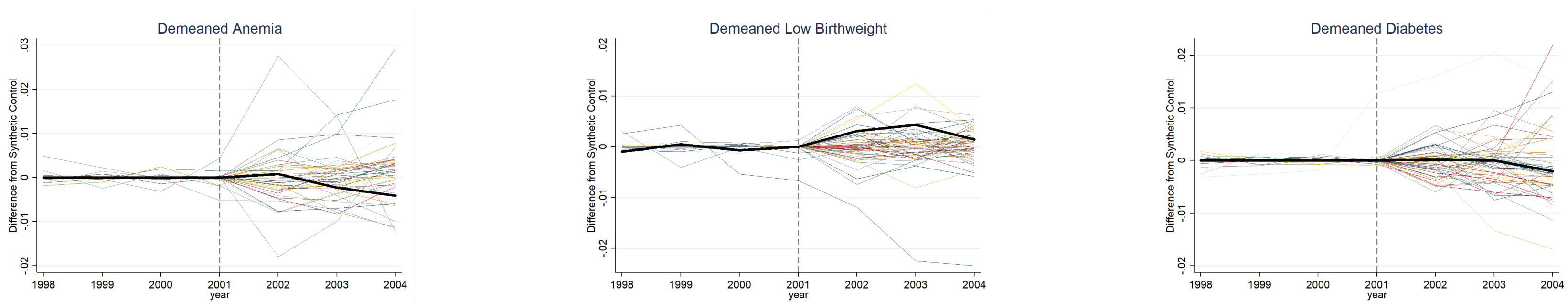


Figure 3.16: Randomization Inference - Demeaned Outcomes

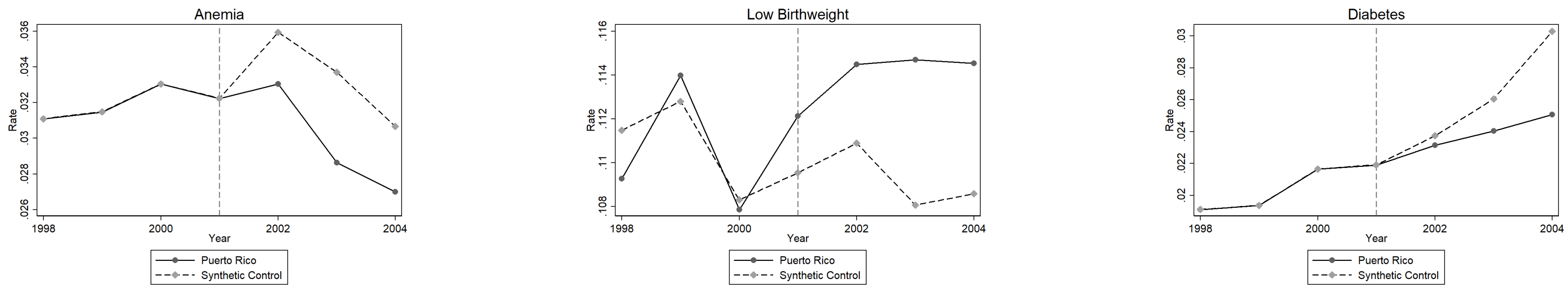


Figure 3.17: Unadjusted Outcomes

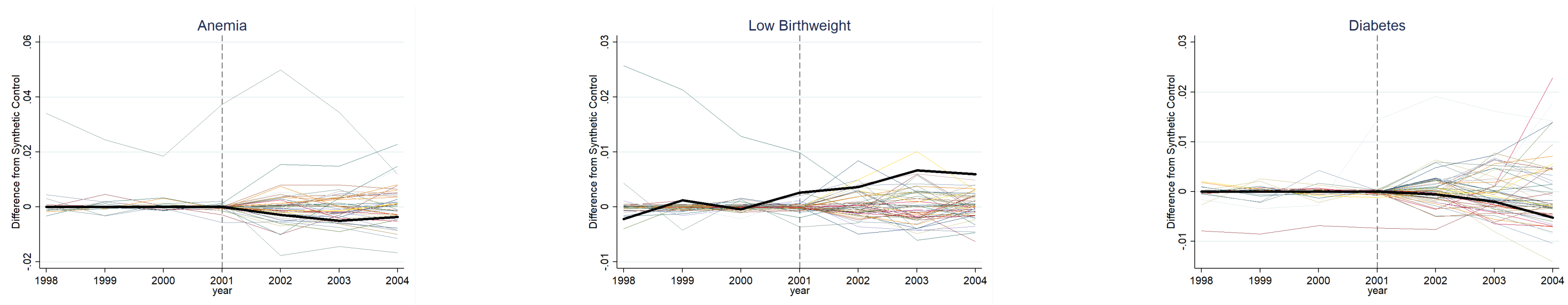


Figure 3.18: Randomization Inference - Unadjusted Outcomes

4. THE EFFECTS OF (FREE) COLLEGE ON EARNINGS AND HEALTH ACROSS THE LIFE CYCLE

This paper is joint work with Andrew Barr (Texas A&M University), Jonathan Eggleston (Census Bureau) and Alex Smith (West Point Academy).

4.1 Introduction

Understanding the returns to higher education has relevant implications both for policymakers seeking to optimize resource allocation for education as well as individuals considering large investments in either their own or their children's education. While there has been significant inquiry into the extent to which higher education enhances labor market outcomes, much of the evidence is for very short-term outcomes, relies on questionable identifying assumptions, or is limited by the context. Even less understood are the effects of higher education on health and health behaviors.[1] While those who attend or complete college exhibit healthier behaviors and have better health, we have little understanding as to whether these relationships are causal. The answer to this question has important implications for optimal investments in and the subsidization of higher education. For example, those with a college degree are significantly less likely to have a chronic disease than those with a high school degree (Choi et al., 2011).[2] If this relationship is causal, it suggests enormous social benefits of higher education through reduced healthcare costs and increased resiliency to disease. Measuring the extent of health effects is critical to thinking about how investments in education may translate to downstream improvements in the health and functioning of society as well as understanding the optimal level of public funding for higher education. Unfortunately, we

[1]These outcomes are often associated with the positive externalities generated by education. To the extent that these externalities are causally related to investments in higher education, the market will lead to sub-optimal investments because individuals do not consider the effects of these externalities in their investment decisions. For example, an 18-year-old considering college enrollment may not consider the spillover effects of worsened health (e.g., others getting sicker, higher insurance costs) that might result from not obtaining a college degree.

[2]For example, those with a college degree are roughly 40 percent less likely to have cardiovascular disease than those with a high-school degree.

have little understanding of the causal relationship between higher education and these types of non-pecuniary outcomes.

The existing evidence linking higher education and improved health relies almost entirely on correlations between college attainment and non-pecuniary outcomes or twin approaches.[3] While these studies suggest a positive relationship between college attainment and health, they rely on strong assumptions about selection into schooling to reach these conclusions. It remains unclear whether the positive association between college attainment and health outcomes is a result of college or merely reflects the selection of healthier individuals, even within twin sets, into college. The few studies that attempt to address this selection issue are limited by their contexts, relying heavily on variation in attendance generated by the Vietnam War draft and associated draft avoidance behaviors (MacInnis, 2006a; Buckles et al., 2016; Grimard and Parent, 2007; de Walque, 2007).

In contrast, we leverage large changes in college attainment generated by the elimination of a large subsidy to higher education, the Student Benefit Program.[4] Under the Student Benefit Program, students age 18-21 who were the children of retired, disabled, or deceased Social Security beneficiaries were eligible to receive monthly payments if they enrolled as full-time college students. At the program's peak in the late-1970s, 12 percent of full-time college students were receiving benefits that averaged roughly $7,500 (2019 dollars). For context, this was roughly equivalent to the student-weighted average total tuition, fees, room, and board charged across all public four-year institutions at the time. It was three times the average tuition and fees charged to in-state students at universities at that time.[5] In 1981, Congress voted to eliminate the program. As demonstrated by Dynarski (2003), the elimination of benefits resulted in a large reduction in

[3]See Oreopoulos and Salvanes (2011) for one recent overview.

[4]At the time of its elimination, program expenditures were around $7 billion (2019 dollars), roughly equivalent to annual Pell grant expenditures at the time and roughly three-quarters average annual Pell expenditures throughout the 1980s and 1990s

[5]See Table 306 of the National Center for Education Statistics 1995 Digest of Education Statistics (https://nces.ed.gov/programs/digest/d95/dtab306.asp).

college attainment among those who were previously eligible. We leverage the decrease in college attainment resulting from the elimination, combined with a unique administrative dataset containing Social Security benefit records linked with administrative and survey data, to examine the impacts of changes in college attainment. As a result of this new and large administrative dataset, we are be able to precisely identify those who were impacted by the elimination of the program and estimate the subsequent impacts of reduced college attainment on earnings and health.

Our unique and large panel of linked administrative and survey data provides two advantages in exploring the relationship between higher education and earnings. First, we have self-reported measures of income and earnings from a national sample. This overcomes concerns related to out of state migration, lack of coverage, or attrition that are often present with Unemployment Insurance wage data or small panel surveys. Second, we are able to estimate impacts on earnings during prime earnings years when earning trajectories are more settled.

Effects on earnings provide an incomplete picture of the returns to education. We bring new evidence to the extremely limited literature on the causal effects of higher education on health. First, our study overcomes concerns related to the endogeneity of college investments that limit prior correlational and twin studies. We provide some of the only evidence on this question that takes advantage of exogenous variation in college attainment. We are able to do so because the elimination of student benefits was a broad policy change that affected a large number of individuals while also being targeted directly at college enrollment. This approach has advantages over prior studies in (1) the extent to which we our results can be interpreted as causal (i.e., internal validity), and (2) the extent to which we believe the estimated effects are generalizable to other contexts (i.e., external validity). Our policy variation provides compelling identifying variation in a recent and demographically varied cohort of individuals. Second, we take advantage of a unique panel of linked administrative and survey data that supports a deeper examination of our central questions. These data allow us to explore effects of college attainment on a broad set of primary outcomes (including measures of earnings, health, and health expenditures).[6] We combine our estimated effects on health with estimates of the dollar values of the associated costs to provide a more complete quantification of the effects associated with higher education.

4.2 Brief Overview of the Wage and Health Returns to Higher Education

While we know a great deal about the returns – both pecuniary and non-pecuniary – to compulsory middle and high school education[7], our understanding of the returns to higher education (or financial aid) is more limited.

4.2.1 Pecuniary Returns

There is an extensive literature estimating the relationship between wages (or log wages) and measures of education (see Oreopoulos and Petronijevic (2013) for an overview). However, much of this evidence relies on questionable identifying assumptions. Only a small subset of the literature makes an attempt to leverage plausibly exogenous variation in college going or attainment to estimate the causal effect on wages. Even among the set of pioneering papers that attempt to leverage exogenous shifters of college going, there are questions about the reliability of the identifying assumptions or the relevance of the estimates to more general or recent populations (Oreopoulos and Petronijevic, 2013). For example, a number of studies take advantage of geographic variation in college proximity to estimate the return to schooling (Card, 1995). However, there is a concern that proximity to a college is endogenous, even after controlling for certain individual characteristics (Card, 2001). Other estimates of the pecuniary return to higher education take advantage of differences in attainment generated by the Vietnam and World War II Era GI Bills (Bound and Turner, 2002; Stanley, 2003; Angrist and Chen, 2011). While interesting and important given the

[6] As a result of data limitations, the few earlier studies to address this broad question focus on a handful of outcomes at specific points in time. We can explore effects on both pecuniary and non-pecuniary outcomes a broad set of ages.

[7] For example, a number of studies take advantage of changes to compulsory schooling laws in middle and high school to examine effects on health (Kemptner, Jürges and Reinhold, 2011; Mazumder, 2008; Lleras-Muney, 2005; Oreopoulos, 2007; Fletcher, 2015) civic participation (Milligan, Moretti and Oreopoulos, 2004), occupational prestige (Oreopoulos and Salvanes, 2011), religious views (Hungerman, 2014; Arias-Vasquez, 2012), and fertility (Tequame and Tirivavi, 2015; McCrary and Royer, 2011; Dincer, Kaushal and Grossman, 2014; James and Vujić, 2019)

scale of these policies, the extent to which the estimates inform our general understanding of the returns to higher education is severely limited by the context of the policies. Furthermore, all of these studies are focused on individuals who attended college in the 1960s or before.

More recently there have been a number of papers to estimate the labor market effects of higher education (or financial aid programs) using compelling identification strategies and more recent cohorts of individuals. One of the most convincing studies compares high school seniors from Florida who barely qualified to attend one of the state's public colleges with seniors who barely missed the academic cutoff (Zimmerman, 2014). Using students from the late 1990s through early 2000s, Zimmerman finds a return to marginal students of 8.7 percent per year at a four-year college. Other recent and well-identified studies estimate returns to additional attainment induced by financial aid (Denning, 2019; Bettinger et al., 2019). Denning (2019) estimate positive effects of the Pell grant on attainment and short-term earnings outcomes in Texas, whereas Bettinger et al. (2019) find evidence of modest increases in earnings resulting from Cal Grant eligibility, but the estimates are imprecise and vary across groups. Interestingly, aid has no effect on enrollment in either of these contexts. Both studies provide compelling strategies to estimate the short-term returns to additional financial aid.

We take advantage of the elimination of a significantly more generous subsidy to study the effects of shifts into higher education in a broad population. This provides several advantages over previous studies. First, our unique and large panel of linked administrative and survey data provides better and more comprehensive measures of income and earnings. We do not have to worry about out of state migration or lack of coverage that is often a concern with Unemployment Insurance wage data. Second, we estimate impacts of college attainment on earnings, including the prime working years when earnings trajectories are more stable. Finally, earnings and income are incomplete measures of the return to additional education. Higher education may have effects on other important measures such as health. These effects may be driven by increases in earnings and income, but there may also be direct effects of higher education on health behaviors and preferences.

4.2.2 Health Returns

Our understanding of the non-pecuniary returns to higher education is even more limited. Most studies rely on the strong positive correlation between college attainment and adult outcomes to suggest that going to or completing college results in better health.[8] To interpret these relationships as causal relies on strong assumptions about selection into schooling. For example, it is likely that children from wealthier families are more likely to attend and complete college. If wealth contributes to health, health behaviors, and longevity as an adult, the positive relationship between college attainment and these outcomes may not be driven by college attendance or completion itself.

A number of studies address this particular type of concern by comparing the health and health behaviors of siblings or twins who complete different amounts of education (Lundborg, Nordin and Rooth, 2018; Lundborg, 2013; Fujiwara and Kawachi, 2009; V.Amin, Behrman and Spector, 2013). This approach "controls" for differences across families, addressing the selection associated with differences in family wealth and upbringing. However, there are still selection concerns related to why one sibling completed more schooling than another. For example, perhaps the twin who was naturally healthier was able to go to college because of her better health. This type of reverse causality could easily explain the positive correlations between college attainment and health, even within families.[9] Putting aside whether sibling education differences are essentially randomly assigned, we might still worry that the effects of education on health outcomes might spillover between siblings, leading us to underestimate the true effect of education.[10] In addition to these fundamental concerns, the small sample sizes underlying these studies severely limit the

[8]See Oreopoulos and Salvanes (2011) for a review.

[9]Recent evidence suggests that even if these estimates are internally valid, a more nuanced type of selection into identification may result in misleading parameter estimates that are substantially biased relative to the population average treatment effect (Miller, Shenhav and Grosz, 2019).

[10]For example, one twin going to college might result in improved health behaviors and health for that individual, but it might also improve the health behaviors and health for the other twin via shared information or peer effects (e.g., eating healthier, going to the doctor more, etc.).

extent to which researchers can draw strong conclusions about even the sign of the correlations.

A handful of innovative studies have attempted to overcome the limitations of the correlational and siblings studies by taking advantage of natural experiments.[11] For example, a number of studies use Vietnam War draft and GI Bill-induced increases in college attainment to examine impacts on mortality, health, and smoking (Buckles et al., 2016; MacInnis, 2006b; Grimard and Parent, 2007; de Walque, 2007). While these studies provide evidence of the effect of education for a large and policy-relevant sample, it is not clear the extent to which they would generalize to other populations, including women and/or those attending college under different and perhaps more typical circumstances. Furthermore, the difficulties associated with disentangling the effect of Vietnam War service, avoidance, and GI-bill induced college attainment limit confidence in the interpretation of the resulting effects as causal.

4.3 Social Security Administration's Student Benefits Program

We contribute to this extremely limited literature by exploiting changes in higher education attainment stemming from the elimination of a large federal student aid program in the United States. The Student Benefit program operated under the umbrella of the Social Security Administration (SSA). In general, SSA benefits were designed to partially offset the loss of income a family experiences when a worker retires, dies, or becomes disabled. This includes increased benefits to support dependent children who have an SSA eligible parent who is retired, disabled, or deceased. Through age 17, dependent children are eligible to receive 75% of the benefits of their parent, although that amount may be reduced to keep the total household benefit level below the household cap.[12] Beginning in 1965, the SSA recognized that children enrolled in school are typically dependent on their parents for support even after the age of 17. To support these students, the SSA

[11]We focus here on studies that estimate quasi-experimental effects on health and health-related outcomes and behaviors, but there is also a small literature that estimates quasi-experimental effects on other non-pecuniary outcomes such as civic participation (for example, Dee (2003)) or estimates effects on health outcomes under structural models (for example, Heckman, Humphries and Veramendi (2018))

[12]Household benefit caps limit the total benefits that can be received by a beneficiary, their spouse and children to 150-180% of the recipient's base benefit eligibility.

expanded the definition of “child” to include individuals age 18-21 who were enrolled in school full time. These extended child benefits are referred to as the SSA’s “Student Benefits Program.”

As with the child benefits, individuals were only eligible for extended student benefits if they had a social security eligible parent who was retired, disabled, or deceased. Additionally, they were required to be enrolled in school full-time, unmarried, and under the age of 22.[13] When a child beneficiary was nearing the age of 18 (and each year after) they would receive a form letter from the SSA verifying their continued student status. Additional verification of enrollment was provided annually by schools (Dynarski, 2003). These students would then receive a separate SSA check each month for as long as they stayed enrolled in the student benefits program. The size of the child benefits did not change after the age of 17 but continued to be calculated based on their parent’s base eligibility and the total family benefits cap. At the program’s peak in the mid-1970s, close to 17 percent of full-time college students were receiving benefits that averaged over $7,500 (2019 dollars) per year (SSA, 1982).[14]

In 1981, as a part of the Omnibus Budget Reconciliation and motivated by a desire to reduce government spending, congress voted to end the SSA’s student benefits program. The new rules required that:

1. Benefits to secondary school students older than 18 ended in August 1982.[15]
2. Students who first enrolled in post-secondary education during or after May 1982 were not eligible for any student benefits.
3. Students who were enrolled in post-secondary education prior to May 1982, continued receiving reduced benefits during the phase-out period. During the phase-out:
    - Students did not receive cost of living adjustment (COLA) increases

[13]If an undergraduate, benefits ended at the end of the semester/quarter that an individual turned 22 (SSA, 1982).

[14]See Table 174 of `https://nces.ed.gov/pubs99/1999036.pdf`, which estimates that there were 4,601,000 18-21 year old full time students in 1975. Combined with number of student beneficiaries listed on the SSA website for 1975 (`https://www.ssa.gov/history/studentbenefit.html`), we estimate that 16.8% of full-time college students in 1975 were receiving benefits (DeWitt, n.d.).

[15]Notably, this preserved benefits for 18-year-old students who had not yet finished high school, which are still available today.

- Student benefits were not payable from May through August
- Student benefits were reduced by 25% (of their August 1981 benefit amount) each year starting in September 1982
- No post-secondary student benefits were paid after April 1985

Effectively, these rules meant that, for children receiving SSA child benefits, those graduating high school before 1982 were eligible for the student benefits program while those graduating in 1982 or later were not. This change in available benefits, including the reduction in benefits for those graduating between 1979 and 1981, is shown in Figure 4.1 and Table 4.3. Eligible students went from having (on average) total cost of four-year public attendance taken care of to not having any assistance over the course of a few years.

4.4 Data

To answer our research questions, we make use of a unique panel of linked administrative and survey data. The Annual Social Economic Supplement (ASEC) sample of the Current Population Survey (CPS) for the 1991, 1994, and 1996-2019 survey years serve as our base sample and contain roughly 185,000 individuals per year (Flood et al., 2020). The ASEC contains a number of questions on health, employment, and income, which makes it an ideal dataset to explore long-run effects of education on various outcomes.

Critically, these data can be confidentially linked to individual-level Social Security benefit records, which include the type and amount of benefits received, allowing us to precisely identify individuals who were eligible for and receiving benefits due to the retirement, death, or disability of a parent.[16,17] We use the receipt of these benefits at age 17 (which was unaffected by the elim-

[16]The SSA data also allow us to observe administrative records on mortality and long-run earnings streams.

[17]The CPS and SSA data are confidentially linked together using the Census Bureau's Person Identification Validation System (PVS) described in Wagner and Lane (2014). This process links both survey and administrative data to a master reference file. All individuals who are matched are then assigned a Protected Identification Key (PIK), which is an anonymized identification number that we use to link the CPS surveys with the administrative SSA dataset.

ination of the Student Benefit program) to assign treatment status. Individuals observed receiving any Social Security Administration benefits at age 17 are in the treatment group, while all other individuals are in the control group.

Our identifying variation takes advantage of whether an individual was expected to graduate high school before or after the elimination of the benefits program in 1982. We use exact date of birth information from the SSA records combined with researcher compiled information on age cutoff requirements for kindergarten in each state and year to code an individual's expected year of high school graduation.[18]

Individuals who were receiving social security benefits as a minor (i.e., those in our treatment group) were eligible to continue receiving benefits during college through age 21 as long as they enrolled in college before May 1982. From September 1982 to April 1985 existing beneficiaries were eligible for reduced benefits, as shown in Figure 4.1 and Table 4.3. We limit our sample to include 5 years of graduates with expected access to full benefits (1974-1978) and 5 years without any expected benefit access (1982-1986). As a result, our analysis sample includes anyone with an expected year of high school graduation between 1974 and 1986. We also exclude individuals with less than an 11th-grade education level since these individuals are unlikely to be affected by treatment.[19] Individuals surveyed in the CPS samples listed above with an expected year of high-school graduation outside of this period are excluded from all analyses.

The merged dataset contains one observation for each year that an individual appears in the March CPS. Everyone who completes their 4-8-4 rotation in the CPS has two observations in the resulting dataset in consecutive years.[20]

Table 4.1 presents summary statistics from publicly available CPS data. Just over half the sample is female. Just under 11 percent of the sample is black, and 5 percent is of Hispanic Origin.

[18]The substantial variation in kindergarten start age cutoffs is shown in Appendix Figure 4.10.

[19]There are interesting questions as to how the promise of financial aid may affect earlier investments but we find it unlikely that this type of investment effect will meaningfully shift the sample of students that reach 11th grade.

[20]As discussed further below, outcomes are clustered at the individual level to account for the correlation between observations of the same individual.

These characteristics vary little between those in the treatment and control groups, with exception of the fact that individuals in the treatment group 6 percentage points more likely to be black. The average years of education is 13.68, with slightly higher education levels in the control group. The average age of observation is 44 years old. Given the available years of linked SSA-linked CPS data, we observe most of the pivotal cohorts between 15 and 35 years after the period of expected high school graduation and potential benefit receipt. As shown in Figure 4.2, this allows us to observe individuals in both the treatment and control group from their late 20s into their 50s, providing a broad window for the observation of income and health.

4.4.1 Key Outcomes

To address multiple inference concerns, reduce measurement error, and simplify the presentation of information, we focus our inquiry on a set of key outcome variables. This includes:

- Earnings Measure
    - ∗ Self-Reported Personal Income
- Health Measures
    - – Primary measures from self-reported scale:
        - ∗ Poor health (as measured by 4 or 5 on the scale)
        - ∗ Estimated health expenditures
    - – Other measures:
        - ∗ Summary index of health (components below)
            - · SSA disability receipt
            - · Physical or cognitive difficulty
            - · Self-reported health status (1-5)
            - · Retired or quit due to health
            - · Disability that limits or prevents work

* Deceased by age 50

The self-reported health measure comes from a question in the CPS ASEC. We use the Likert measure, which asks individuals if their health is excellent, very good, good, fair, or poor, to construct a binary indicator for poor health (defined as "fair" or "poor" health) and to monetize effects on health by mapping the individual Likert scale answers to cost measures from the Medical Expenditures Panel Survey (MEPS). We also construct a health index to combine information from a larger set of variables in a useful way as well as estimating effects on mortality by age 50.

4.4.1.1 Health Index

We construct several summary indices of our outcome measures, grouping components by theme. The index of health outcomes, which is a primary outcome of interest, combines measures of self-reported health status (on a scale of 1-5), having a disability that limits or prevents work, any SSA disability receipt, and having ever left a job for health reasons (Kling, Liebman and Katz, 2007).

We generate z-scores by subtracting the mean and dividing by the standard deviation of each variable. The index is the average across the standardized z-score measures of each component.[21] For each component of these index, individual variables are adjusted so that a higher index score reflects a "better" outcome. Due to the rotating panel nature of the CPS and the varied timings of supplemental surveys, some individuals are missing some components of an index. For the construction of the index, missing values are imputed using the mean value of that variable for individuals of the same cohort (expected HS graduation year), treatment group, and gender.

4.4.1.2 Monetizing and Quantifying Health Effects

To monetize and quantify the health effects of additional college attainment, we also generate variables estimating the direct effect on individual health expenditures. These measures may be used as a quantification of the effect on health to an individual or as a measure of the burden they

[21] The health index itself is also standardized using the same procedure.

place on the healthcare system. The role that private and public health insurance play in paying for medical expenses, as well as the obligation of physicians to provide lifesaving care even to patients who are unable to pay, means only a small portion of medical expenses are paid for directly by individuals receiving care. In 2015, only 11% of total health spending was paid for out-of-pocket (HHS, 2016).

While we are unable to directly observe health expenditures of our sample, we are able to estimate expected health expenditures based on individuals reported health level. From 1996 onward, respondents to the ASEC survey have been asked to rate their health as excellent, very good, good, fair, or poor. This type of self-reported health has been extensively used in health surveys and there is a wide body of research evaluating its reliability.[22] The evidence suggests that perceived health levels primarily reflect underlying disease burden (Kaplan et al., 1996).

In fact, DeSalvo et al. (2009) find that a simple model of age and self-reported health predicts future health expenditures just as well as a set of more complex health expenditure prediction models using significantly more detailed information. Inflation adjusted to 2015, annual averages range from $2,069 for individuals in excellent health to $15,946 for those in poor health. Given that only 11% of medical expenses were paid out of pocket, one might view the remaining 89% of the expenses (whether paid by private insurance, Medicare or Medicaid) to be a cost borne by society more generally. We use MEPS data to estimate average yearly medical expenses by self-reported health status, age, and gender. These estimated expenditures are then mapped into our CPS sample to give an estimate of yearly medical expenditures for each individual.

4.4.2 Education Gradients

As described in Section 4.1, prior work has illustrated a positive correlation between postsecondary educational attainment and a number of long run outcomes. However, determining whether these relationships are causal has proved challenging. Because we expect individuals of higher socioeconomic status, health, and ability to select into college, we also expect those same individuals

[22] See Idler and Benyamini (1997) for a review.

to report higher levels of health, longer lifetimes, and improved employment outcomes relative to their peers even in the absence of any causal impacts of higher education. Despite this, correlational results and trends by education level are often misrepresented as, or implied to be, causal effects both within and outside of the academic literature. In light of this, we aim to not only provide accurate estimates of the causal effects of higher education, but also establish what proportion of observed education gradients can be attributed to education itself rather than endogenous selection. In this section, we will outline the relationship between each of our main outcomes and educational attainment, and how those gradients were constructed.

It is important to note that our policy variation is variation in the availability of financial aid. While we think that the associated effects on college going and attainment are the most likely drivers of any downstream effects, it is possible that the aid itself has direct effects on student outcomes. For example, for many inframarginal enrollees the Student Benefit Program essentially provides an additional $7,500 in income per year. If this income has direct effects on student choices and behaviors outside of the educational investment decision, this might influence downstream outcomes. Some recent evidence suggests that the effects of this type of windfall income (and even much larger amounts) on health and health behaviors are essentially zero (Cesarini et al., 2016). If this is true, our "instrumental variables" estimates may be unbiased. Of course, we can also think more directly about the reduced form estimated effects of the program as the earnings and health returns to "free college" or generous financial aid.

For the purposes of our first stage estimation and creation of education gradients, we define educational attainment as the years of completed education up to completion of a bachelor's degree.[23] We top code years of education at 16 due to the linear relationship between years of education and our main outcomes of interest. Appendix Figures 4.11 through 4.13 show a relatively linear trend between years of education and our main health outcomes, which levels off for advanced degrees.

[23]Degree completion is mapped into years of education such that a high school diploma, associate's degree, or bachelor's degree represent 12, 14, or 16 years of education respectively.

Our choice to focus on years of college attainment is further motivated by the availability of student benefits, which were only available through age 21.

4.5 Empirical Strategy

We implement a difference-in-differences design, comparing the outcomes of SSA benefit recipient children who likely graduated high school before and after the elimination of the Student Benefit Program. The treatment group is defined as all individuals who received any Social Security Administration benefits at age 17. All other individuals are in the control group. The "before" variable indicates all individuals who were expected, based on their exact date of birth and state school age cutoffs, to graduate high school before the end of the student benefits program in 1982. The analysis does not use any sample weights. Standard errors are clustered at the individual level to account for correlation between multiple observations from the same individual. The basic reduced form specification is as follows:

$$Y_{it} = \lambda_t + \beta_1 * SSA_i + \beta_2(SSA_i * Before_t) + \gamma X_{it} + \epsilon_{it}, \tag{4.1}$$

where Y_{it} is an outcome measure for individual i in cohort t. Cohort fixed effects are represented by λ_t while X_{it} are individual covariates including state of birth, gender, age, race, and Hispanic origin. SSA_i indicates whether the individual received SSA benefits at age 17 and $Before_t$ indicates whether an individual's expected year of high school graduation was before the program ended in 1982. The parameter of interest is the coefficient on the interaction term $SSA_i * Before_t$, which equals one for individuals who received student SSA benefits at age 17 and whose cohort graduated high school before the program ended in 1982.

All specifications include birth state and birth cohort (defined by the expected year of HS graduation) fixed effects. Our preferred specification (#5 below, shown in **bold**) includes binary controls for individual gender, age, age-by-gender, race, and Hispanic origin.[24] To illustrate robustness, we

[24]To maintain consistency across survey years, only two binary race variable were created; White and Black.

estimate and show results using each of the following sets of controls in Table 4.2:

1. No Additional Controls
2. Gender
3. Gender, Age (indicators)
4. Gender, Age (indicators), Race, Hispanic origin
5. **Gender, Age (indicators), Gender by Age (indicators), Race, Hispanic origin**

4.6 Results

We find preliminary evidence that access to the Social Security Administration's student benefits program resulted in over $7,000 (2015) dollars of additional benefit income during college going years (age 18-21) and a corresponding increase of 0.185 years of education which are both significant at the at the 1% level. On the other hand, these preliminary results show no statistically significant increase in income. Despite estimating statistically significant increases in medical expenses and decreases in overall health for eligible beneficiaries, these results do not appear to causal (and will be discussed in more detail later in this section). There is also no statistically significant impact on rates of poor health or mortality by age 50. These results are estimated using the reduced form specification described in Section 4.5 and are shown, along with their robustness to alternative sets of control variables, in Table 4.2.

In prior work leveraging the end of the student benefits program Dynarski (2003) found that access to benefits increased educational attainment by 0.754 years. These results were estimated using five cohorts of graduating seniors in the National Longitudinal Survey of Youth (NLSY), focusing on students with deceased fathers. Using our larger administrative data sources, our preliminary estimates show a much smaller effect of 0.185 education years. Our results may differ from Dynarski's findings for a variety of reasons, including the fact that we are using a different sample, more cohorts, or that we also include children of disabled parents. Given the relatively small sample size of the NLSY, those findings may simply be an outlier that is higher than the population average.

We also estimate dynamic effects relative to each cohort's expected year of high school graduation. These event studies are shown in Figures 4.3 through 4.9. For most outcomes, these figures show level effect sizes across cohorts graduating high school prior to 1982, which suggests that estimated treatment effects are driven by differential access to student benefits rather than some other factor. In the case of the Health Index and Poor Health these event studies show a clear trend in outcomes prior to the end of the student benefits program. This suggests an increased likelihood that the parallel trends assumption is violated in these cases, we are therefore cautious about interpreting the resulting estimates as causal. Considering the pre-existing differences in trends, we interpret these results as showing no evidence of a long run impact of on health outcomes. These figures are discussed further in the Section 4.6.1.

Somewhat surprising, we also find no statistically significant impact of financial aid on income. Taking the point estimates at face value, these preliminary findings show that access to student benefits resulted in a $354 (2015) dollar increase in income, which, if driven by the education increases of 0.185 years, suggests a return of $1,910 dollars per year of additional education. Given the average income of individuals in the treatment group of almost $45,880 (showing in Table 4.1), this suggests a 4% per year return to education, which is much smaller than has been estimated in other contexts (Zimmerman, 2014). Unfortunately, the confidence intervals are large, and we are unable to reject even large increases or decreases in personal income.

Overall, our findings confirm that the SSA student benefit program was providing substantial amounts of financial aid, but this financial support resulted in relatively small increases in educational attainment—only one fourth of previous estimates. As a result, we find no evidence in these preliminary estimates that the receipt of additional financial aid or the resulting increase in educational attainment led to substantial improvements in health or labor market outcomes.

4.6.1 Testing Identifying Assumption

The key identifying assumption is that, conditional on cohort and group fixed effects, the likelihood of going to or persisting in college or obtaining better later-life outcomes is orthogonal to one's membership in the group eligible for Student Benefits. In other words, conditional on cohort and group fixed effects, any difference in outcomes among the group eligible for Student Benefits is a result of the benefit availability and not some other factor.

In our context, the primary assumption is that the shift in attainment patterns of non-beneficiaries from before to after effectively proxies for the shift in attainments patterns that would have occurred among beneficiaries absent the elimination of student benefits, the standard difference-in-differences parallel trends assumption. While it is impossible to observe the counterfactual behavior of the treated group, we conduct several tests to explore whether or not that assumption is reasonable in this setting.

First, we estimate dynamic treatment effects, allowing the impact of being in the treatment to be different for each cohort. This method allows for treatment effects to change over time and provides a natural test of the parallel trends assumption. We present the results of these tests in event study Figures 4.3 through 4.9. If the treatment and control groups are trending differently prior to the elimination of benefits, we would see a positive or negative slope in the figure. A flat trend of effect sizes prior to treatment indicates that there were no pre-existing differences in trends between the two groups and provides suggestive evidence that the identifying assumption holds.

Given parallel trends prior to the elimination of benefits, our identifying assumption may still fail if the group of people in our treatment and control groups are not consistent across time. For example, one might be concerned that the end of the student benefits program could itself cause a shift in the unobservable characteristics of the treatment group, if some parents were claiming retirement benefits specifically so that their children were eligible for student benefits. These types of parents would stop claiming retirement after the Student Benefit program ended. In that scenario, those individuals would be assigned to the treatment group in the pre-period and the control group in the post-period, biasing our estimates. It is possible that behavioral responses could change the unobservable characteristics of the treated group without necessarily changing the demographic makeup of included individuals. In order to avoid this issue, Dynarski (2003) excludes individuals who are receiving benefits due to parental disability or retirement from her analysis, focusing instead on children with a deceased father, which is unlikely to be manipulated in response to the availability of the program. We limit our sample to exclude children of retired beneficiaries, as this group is most likely to be able to manipulate the timing of their retirement relative to children's college attendance.

In addition to testing the validity of our identifying assumptions, we also provide additional evidence on the sensitivity of our results to minor specification changes and whether the pattern of effects (or lack thereof) matches theoretical predictions. One way that we demonstrate robustness to specification changes is by estimating our results using the various sets of controls described in Section 4.5.

4.6.2 Future work

The results presented in this paper are preliminary and ongoing. Future versions of this paper may contain

- Additional outcomes made available through linked CPS supplements and administrative SSA records.
- Further exploration of secondary outcomes and discussion of mechanisms.
- Heterogeneity analysis, focused on any subgroups that are more strongly affected by financial aid.
- Alternative analysis approaches including variations of treatment group definitions and sample restrictions as well as an instrumental variables approach.
- Exploration of similarities and differences between the education gradient in treatment and

control groups.

- Comparison of estimated effect sizes relative to education gradients for those variables.
- Additional robustness checks, including tests of potential strategic responses in the timing of parental retirement or disability receipt.

4.7 Conclusion

Understanding the returns to higher education has relevant implications both for policymakers seeking to optimize resource allocation for education as well as individuals considering large investments in either their own or their children's education. While there has been significant inquiry into the extent to which higher education enhances labor market outcomes, much of the evidence is for very short-term outcomes, relies on questionable identifying assumptions, or is limited by the context. Even less understood are the effects of higher education on health and health behaviors. While those who attend or complete college exhibit healthier behaviors and have better health, we have little understanding as to whether these relationships are causal. The answer to this question has important implications for optimal investments in and the subsidization of higher education. Measuring the extent of health effects is critical to thinking about how investments in education may translate to downstream improvements in the health and functioning of society as well as understanding the optimal level of public funding for higher education. Unfortunately, we have little understanding of the causal relationship between higher education and these types of non-pecuniary outcomes.

In this paper, we leverage large changes in college attainment generated by the elimination of a large subsidy to higher education, the Student Benefit Program. In 1981, Congress voted to eliminate the program. Using a unique administrative dataset containing Social Security benefit records linked with administrative and survey data, we find that the elimination of the program resulted in a relatively small decrease in educational attainment of 0.185 years, about one fourth the size of previous estimates. Although we can verify that average financial aid fell by over $7,000 in the treatment group, these preliminary results show no statistically significant causal relationship between the end of the student benefit program and long run income, health, or mortality.

These findings have important policy implications for financial aid policies. Strong correlations between educational attainment and labor market and health outcomes are frequently used as support for public and private investments in educational attainment. If the causal effect of education reflects a relatively small portion of these correlations, it suggests a need for careful reevaluation of optimal higher education policy. The results presented in this study are preliminary and future work will include a variety of additional robustness checks and heterogeneity analysis and may focus on more targeted subsamples.

4.8 Figures and Tables

Table 4.1: Summary Statistics

	(1) Control	(2) Treated	(3) Full Sample
Demographics			
Black	0.1024	0.1657	0.1089
Age	44.48	44.33	44.47
Male	0.4795	0.4835	0.4799
Hispanic	0.0526	0.0596	0.0533
Outcomes			
SSA Benefits from Age 18–21	12	6724	698
Years of Education	13.71	13.41	13.68
Health Index	0.0443	-0.1402	0.0254
Died by age 50	0.0123	0.0168	0.0128
Total Personal Income	53210	45880	52460
Observations	490,000	60,000	540,000
Poor Health	.0975	.1337	.1012
Health Expenditure	3435	3743	3466
Observations	420,000	50,000	460,000

Note: Table presents descriptive statistics for demographic and main outcome variables. Statistics are shown separately for treated individuals (those who received any SSA benefits at age 17) and the control group. Health questions were not asked during the 1991 CPS survey and have a slightly smaller sample size as a result. All output is rounded in accordance with Census Bureau and Social Security Administration policy.

Table 4.2: Results

	(1)	(2)	(3)	(4)	(5)
Outcomes					
Benefits from 18–21	7099***	7099***	7099***	7099***	7099***
	(88)	(88)	(88)	(88)	(88)
Years of Education	0.1827***	0.1847***	0.185***	0.1852***	0.1853***
	(.0193)	(.0193)	(.0193)	(.0192)	(.0192)
Total Personal Income	918	281	291	310	354
	(621)	(602)	(600)	(598)	(598)
Died by age 50	-0.0017	-0.0018	-0.0018	-0.0019	-0.0018
	(.0015)	(.0015)	(.0015)	(.0015)	(.0015)
Health Index	-0.0272**	-0.028**	-0.0275**	-0.0269**	-0.0268**
	(.012)	(.012)	(.0119)	(.0119)	(.0119)
Observations	540,000	540,000	540,000	540,000	540,000
Health Expenditure	83**	102***	87***	86***	84***
	(33)	(32)	(29)	(29)	(29)
Poor Health	0.0049	0.0052	0.0048	0.0047	0.0046
	(.0037)	(.0037)	(.0037)	(.0037)	(.0037)
Observations	460,000	460,000	460,000	460,000	460,000
Gender FE	N	Y	Y	Y	Y
Age FE	N	N	Y	Y	Y
Race FE	N	N	N	Y	Y
Hispanic Origin FE	N	N	N	Y	Y
Age by Gender FE	N	N	N	N	Y

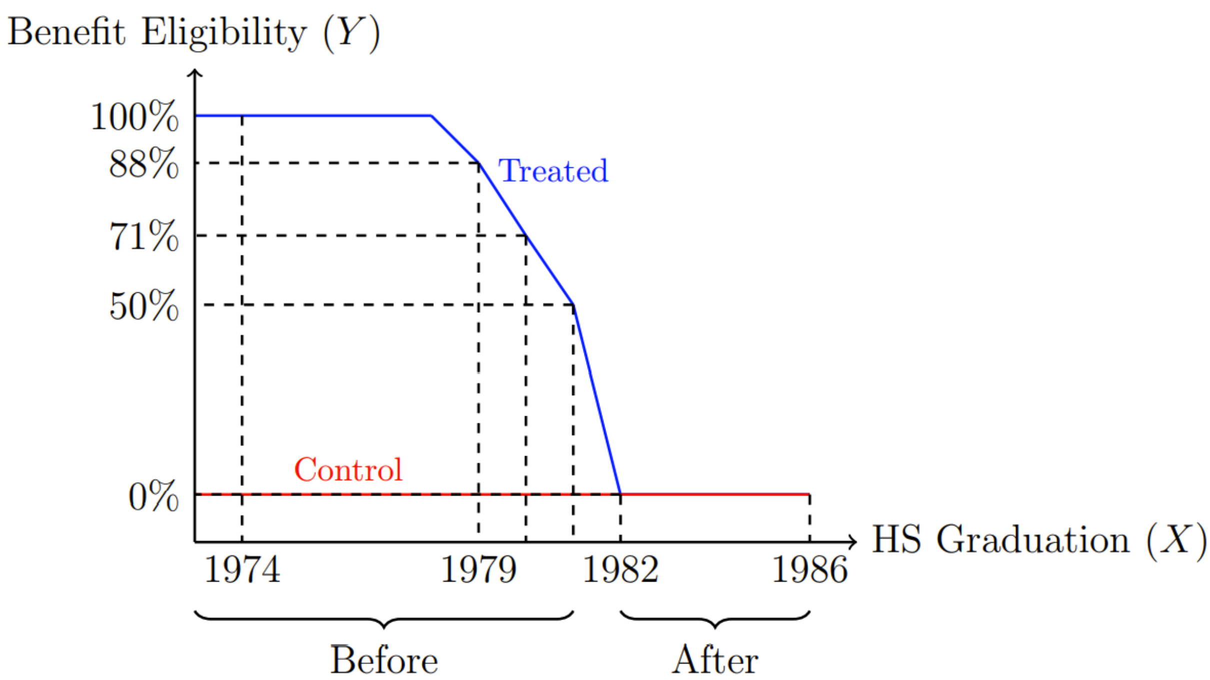
Note: This table displays the primary outcomes as additional controls are added - ending with the preferred specification in column (5). Health questions were not asked during the 1991 CPS survey and have a slightly smaller sample size as a result. All output is rounded in accordance with Census Bureau and Social Security Administration policy. Significance levels indicated by: * (p<0.10) **(p<0.05), ***(p<0.01).

Table 4.3: Benefit Eligibility

	Freshman (1)	Sophomore (2)	Junior (3)	Senior (4)	Total (5)
High School Graduation Year					
1977	100%	100%	100%	100%	100%
1978	100%	100%	100%	100%	100%
1979	100%	100%	100%	50%	88%
1980	100%	100%	50%	33%	71%
1981	100%	50%	33%	17%	50%
1982	0%	0%	0%	0%	0%
1983	0%	0%	0%	0%	0%

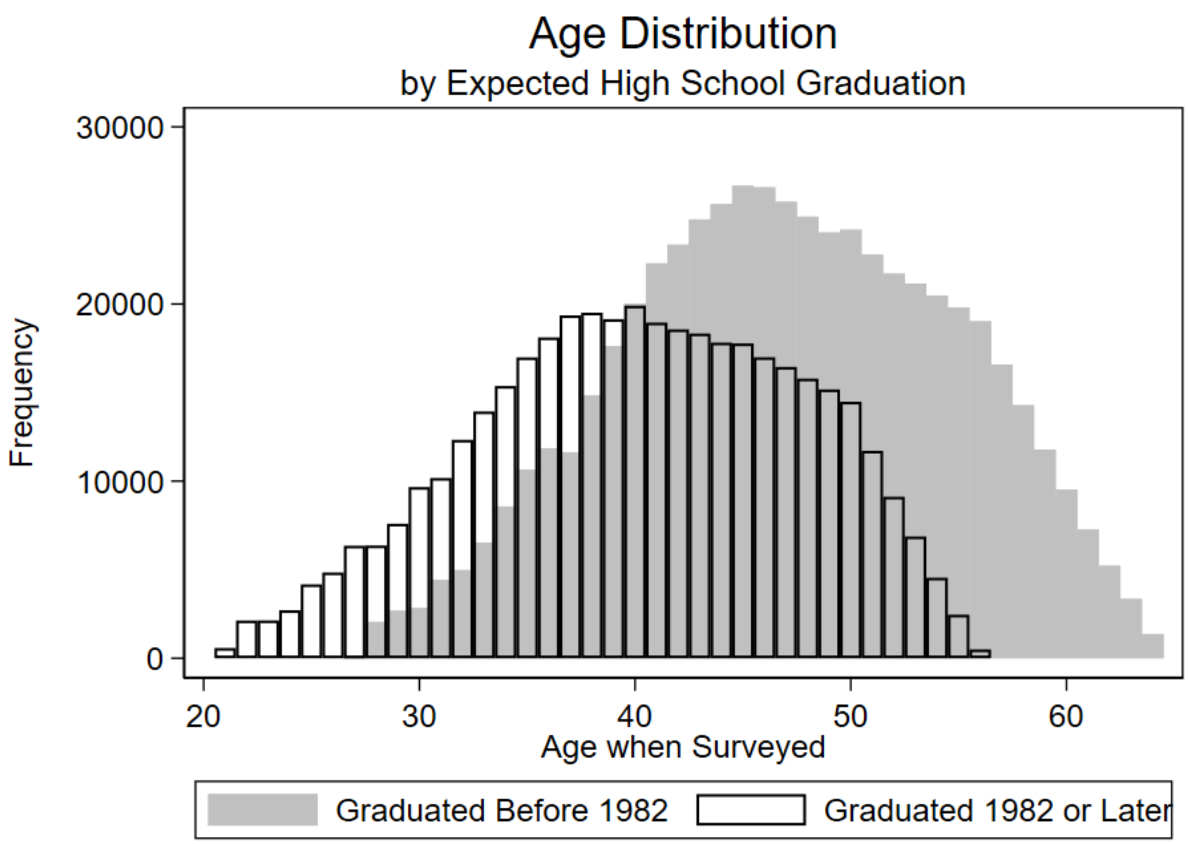
Note: Individuals who were receiving student benefits before the elimination of the program continued to receive reduced benefits after 1981. These students had their benefits reduced by 25% each year, cost of living adjustments were halted, and they were not eligible for any benefits during summer months (May-August). Because students were being reduced by 25% each year and only being paid 8 months a year, continuing students received 50% benefits (.75*8/12) for the 1982 academic year, 33% benefits (.5*8/12) for 1983 and 17% (.25*8/12) for 1984. No post-secondary students were eligible to receive benefits after April 1985. The cumulative effect of these reductions by cohort are displayed in column (5) above. These calculations do not account for the halt in cost-of-living adjustment after August 1981, which further reduced the benefits in real terms for cohorts graduating high school from 1979-1981.

Figure 4.1: Benefit Eligibility



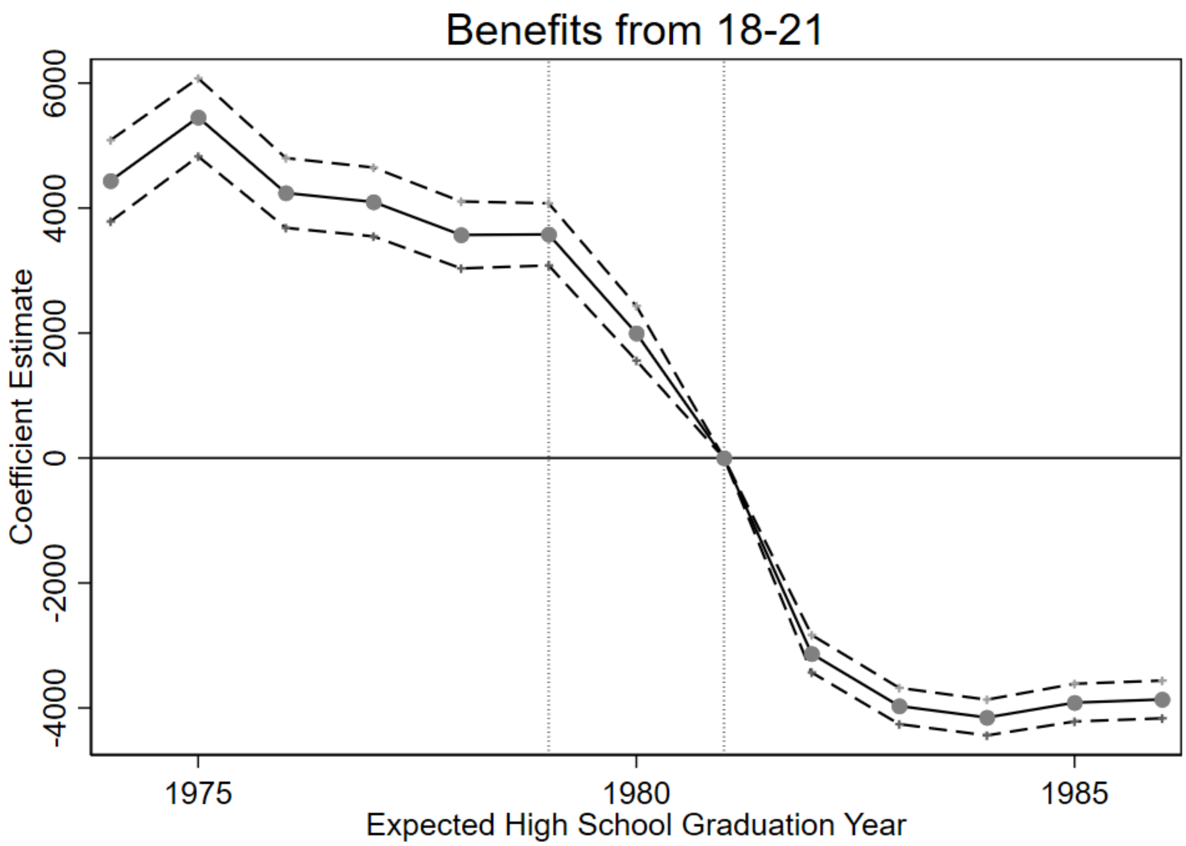
Note: This figure displays the total 4-year SSA student benefits eligibility by high school graduation cohort from Table 4.3 Column 5. While the control group was never eligible for benefits, access for the treated group dropped off sharply for cohorts graduating high school between 1979 and 1982. These calculations do not account for the halt in cost-of-living adjustment after August 1981, which further reduced the benefits in real terms for cohorts graduating high school from 1979-1981.

Figure 4.2: Age Distribution



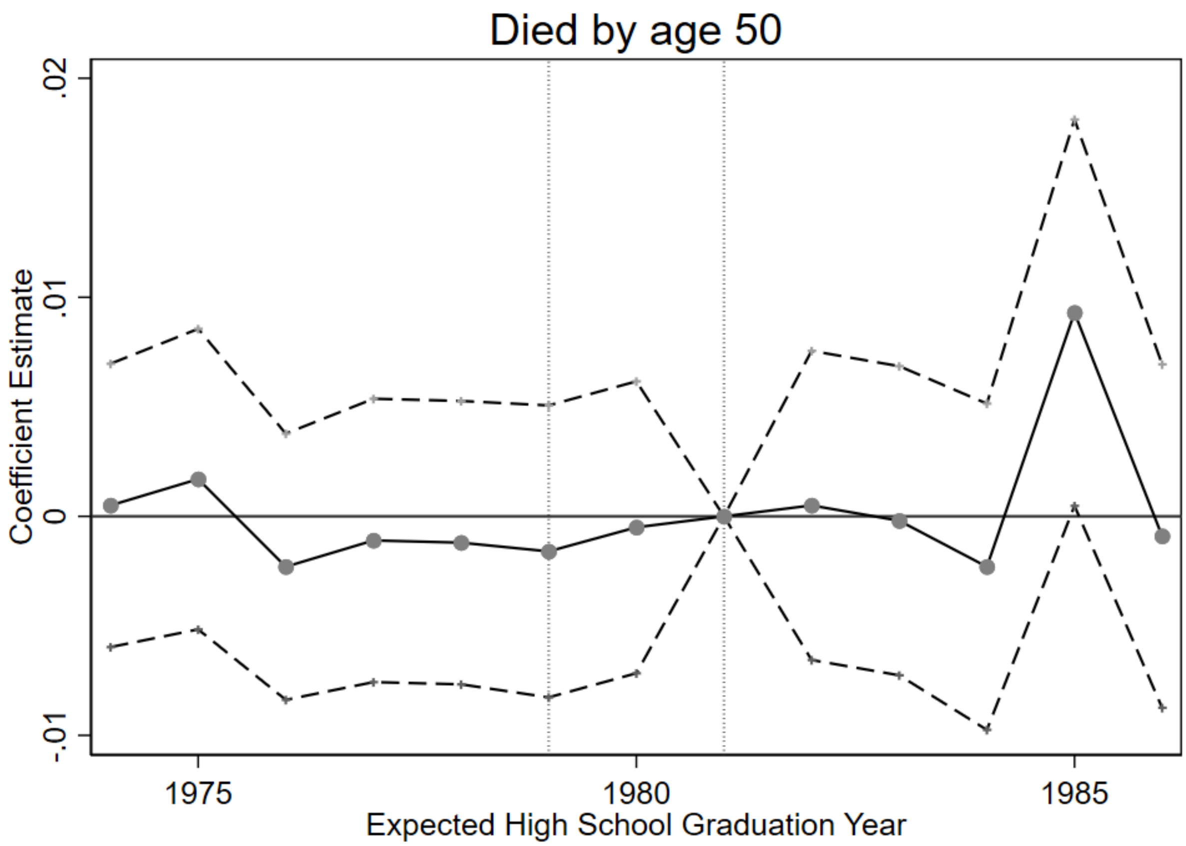
Note: This figure shows the distribution of individuals' age at the time of survey for those who graduated before 1982 (when the SSA Student Benefit program was still in existence) versus those who graduated in 1982 and later. Statistics presented here are calculated from publicly available data.

Figure 4.3: Event Study - Total Benefits from 18-21



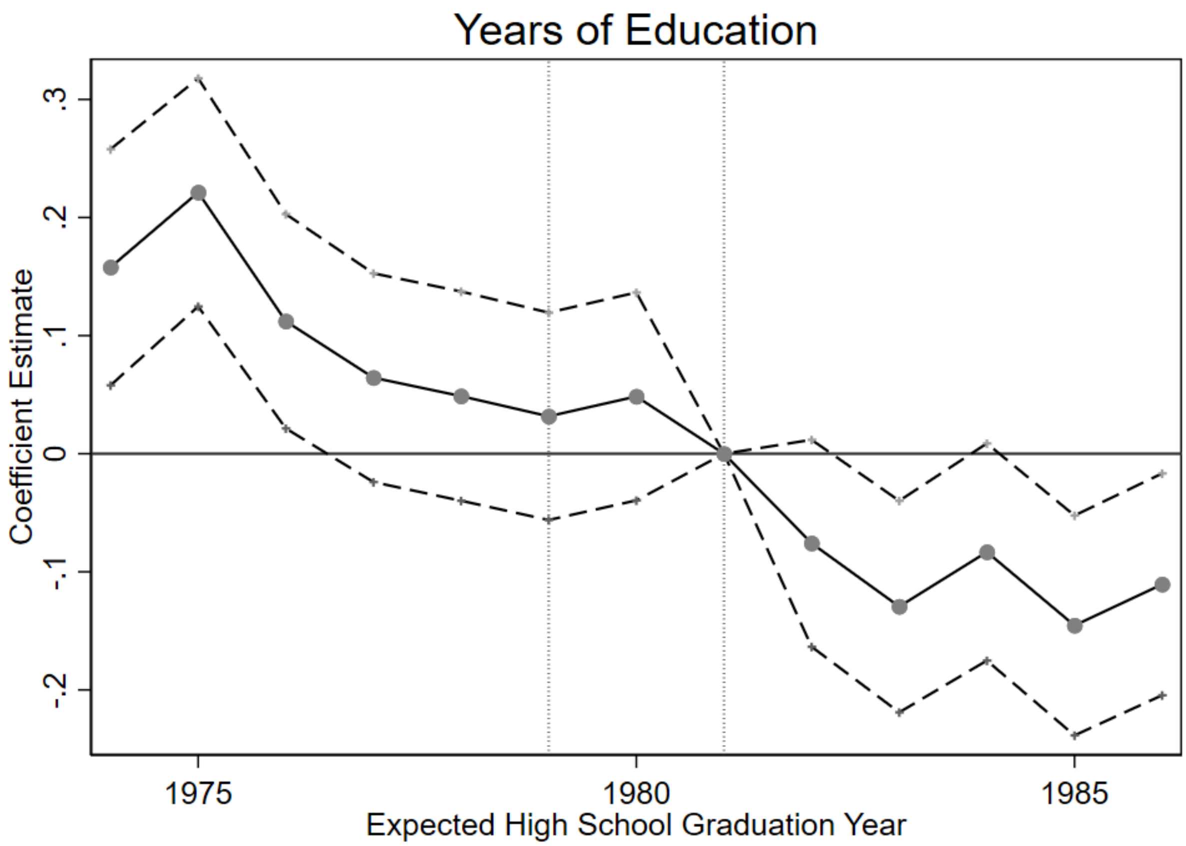
Note: This figure shows the dynamic effects of access to SSA student benefits on total SSA benefits received from age 18-21 by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

Figure 4.4: Event Study - Died by Age 50



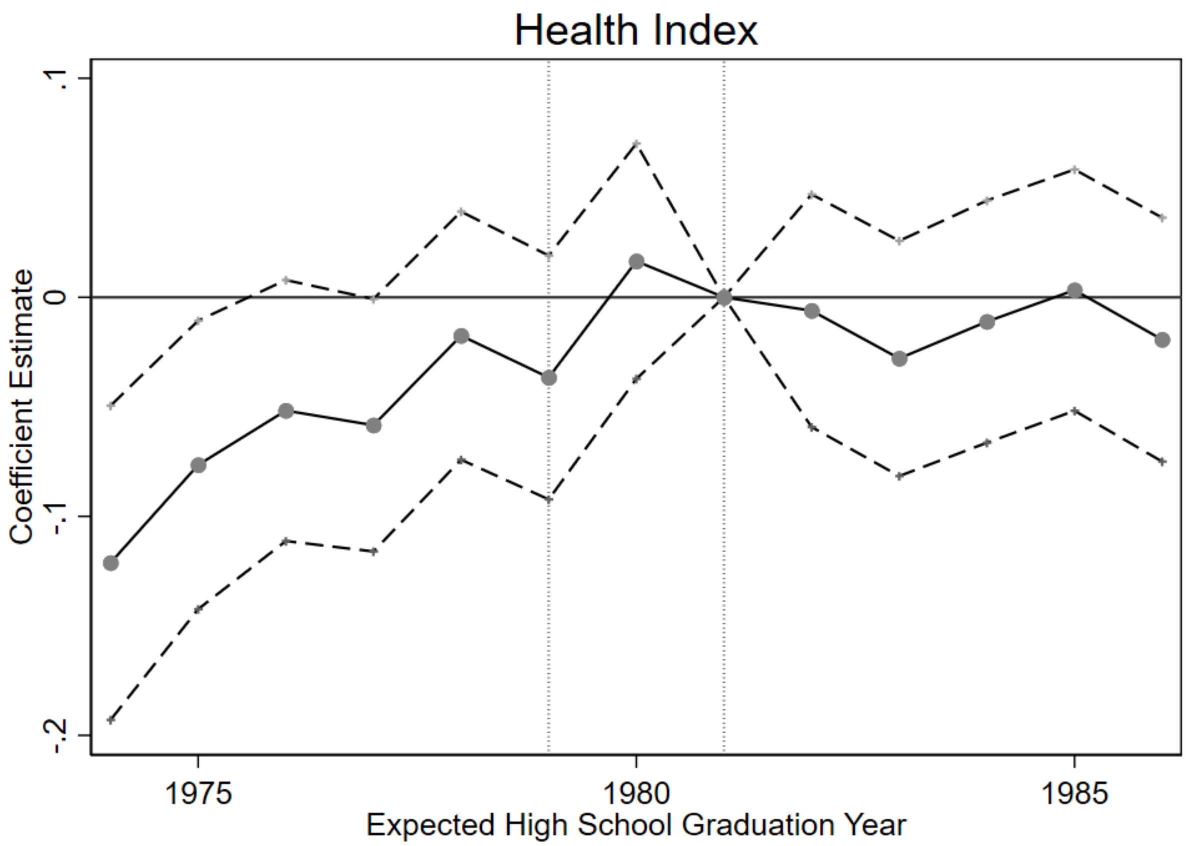
Note: This figure shows the dynamic effects of access to SSA student benefits on mortality by age 50 by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

Figure 4.5: Event Study - Years of Education



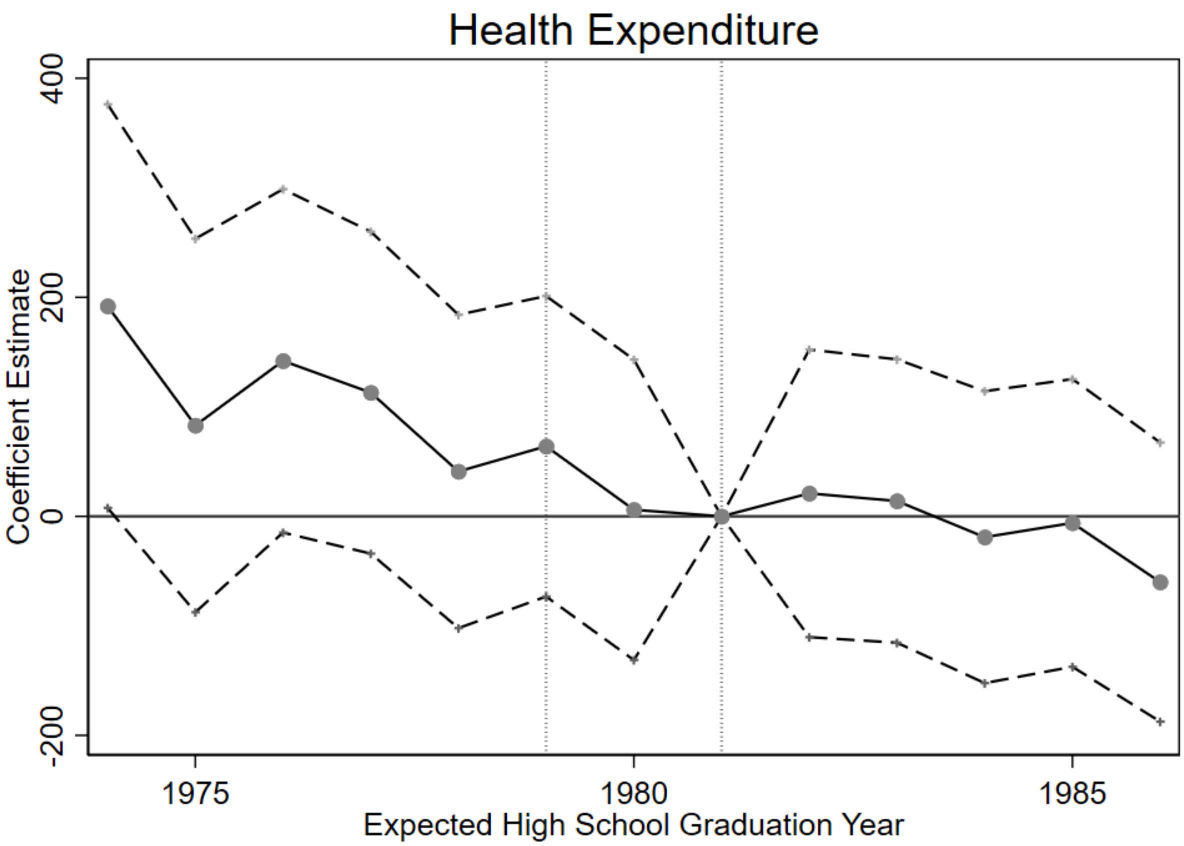
Note: This figure shows the dynamic effects of access to SSA student benefits on years of education by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

Figure 4.6: Event Study - Health Index



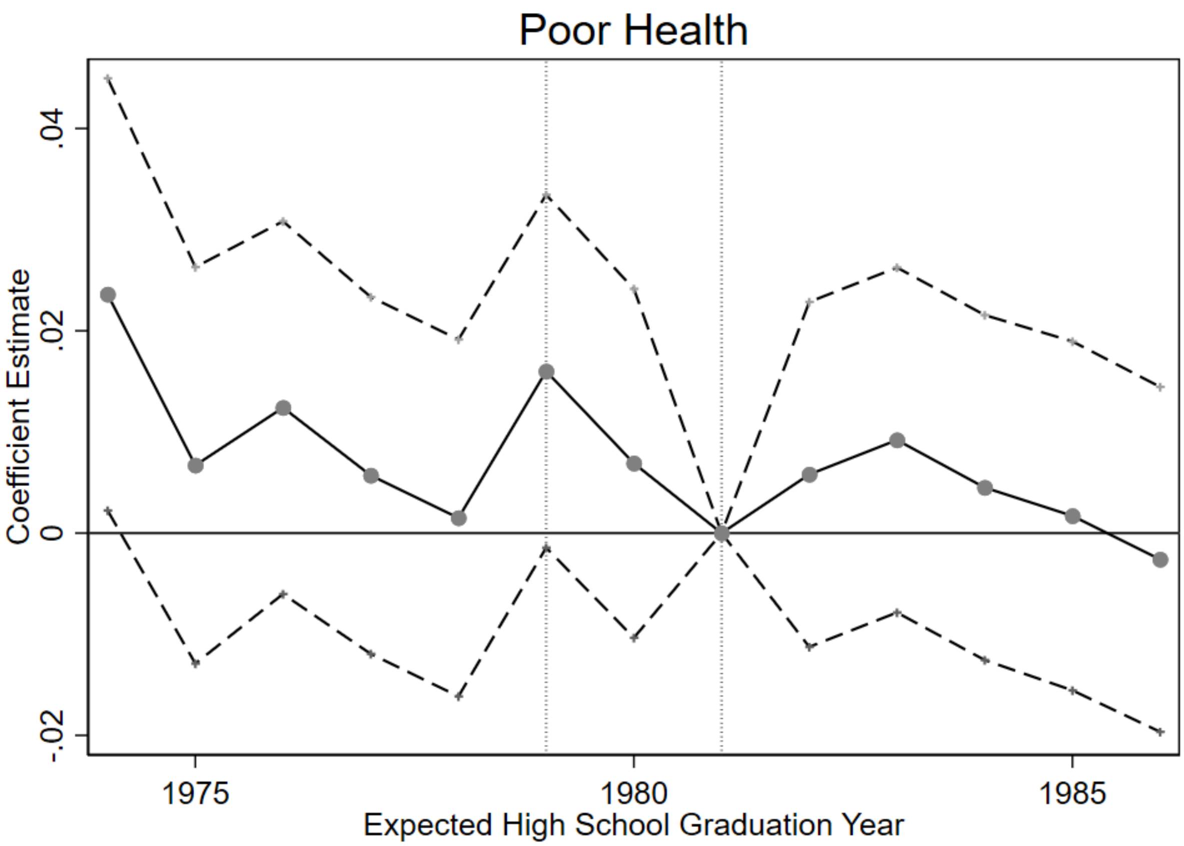
Note: This figure shows the dynamic effects of access to SSA student benefits on the health index by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

Figure 4.7: Event Study - Health Expenditure



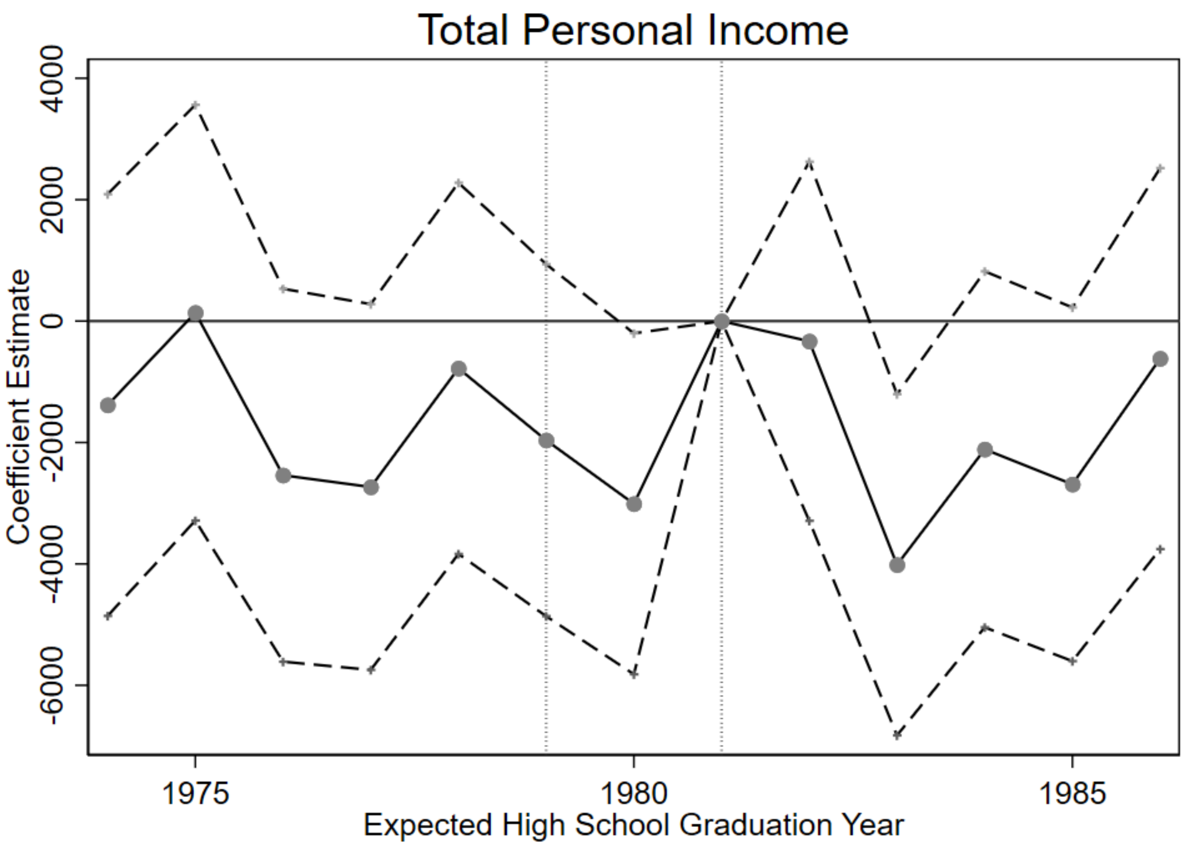
Note: This figure shows the dynamic effects of access to SSA student benefits on health expenditure by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

Figure 4.8: Event Study - Poor Health



Note: This figure shows the dynamic effects of access to SSA student benefits on poor health by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

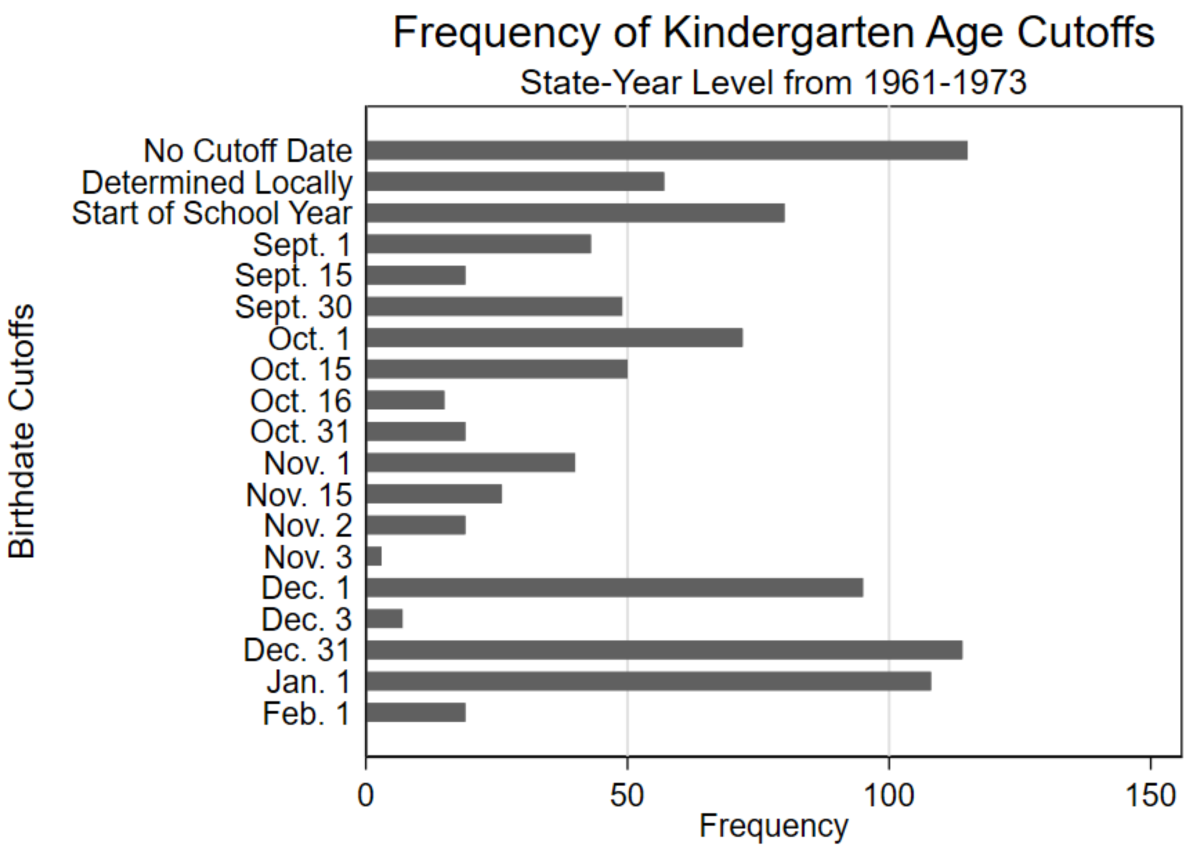
Figure 4.9: Event Study - Total Personal Income



Note: This figure shows the dynamic effects of access to SSA student benefits on total personal income by expected high school graduation year. Cohorts left of 1982 were eligible for student benefits (those to the left of 1979 received full benefits) while cohorts in or after 1982 had no access to the student benefits program after graduating high school. All coefficients are rounded in accordance with Census bureau policy.

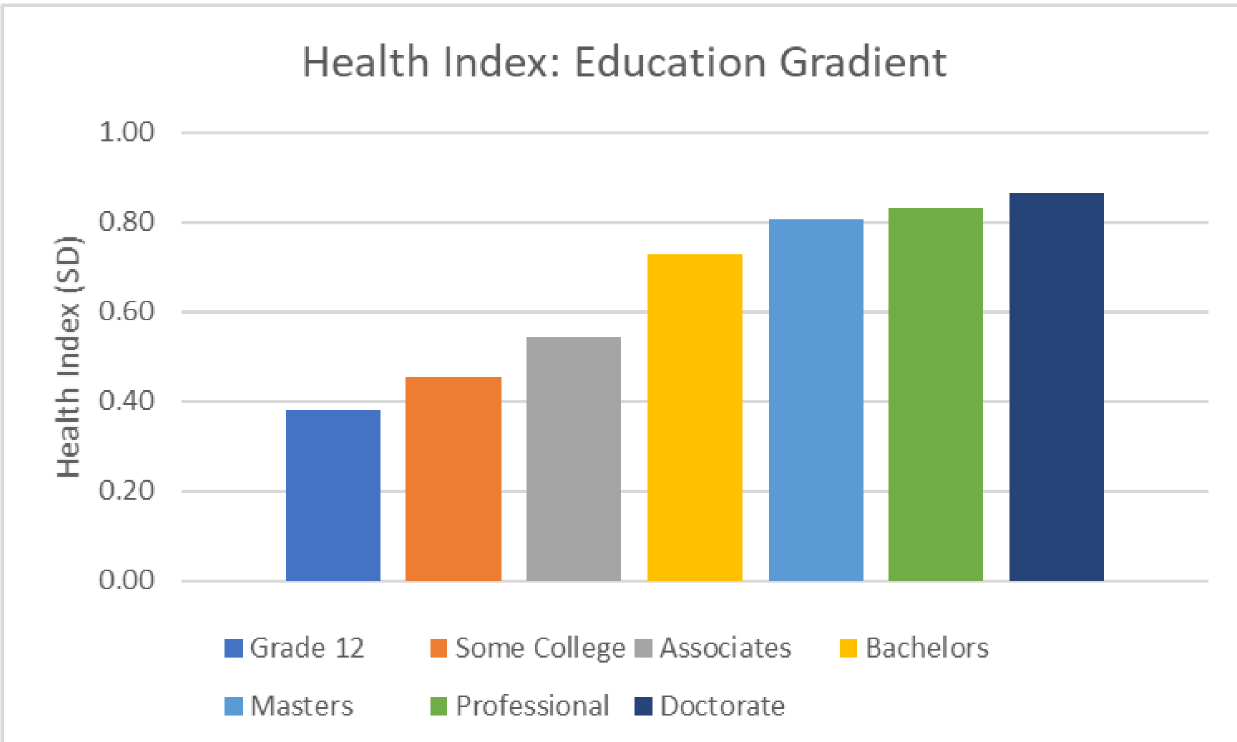
4.9 Additional Tables and Figures

Figure 4.10: Kindergarten Age Cutoffs



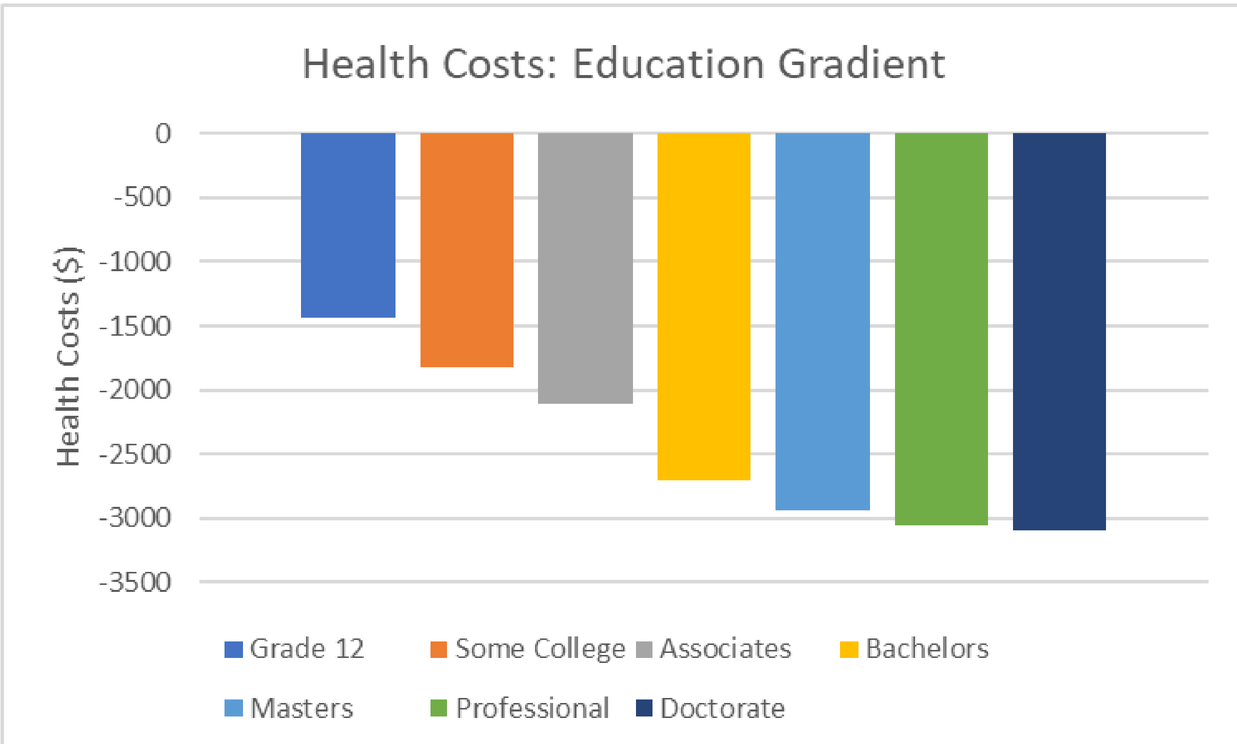
Note: This figure displays the distribution of school start cutoff dates for children entering kindergarten. Observations are at the state-year level. These kindergarten age cutoffs are used in conjunction with exact date of birth to define each CPS survey respondent's expected year of high school graduation. We approximate the age cutoff date using common school start date(s) for state-years with a "start of school year" cutoff system. For state-years where the cutoff was determined locally, missing, or where there were no age cutoff dates, we approximate age cutoffs using the most common nationwide cutoffs. These data were compiled by the researchers.

Figure 4.11: Health Index - Education Gradient



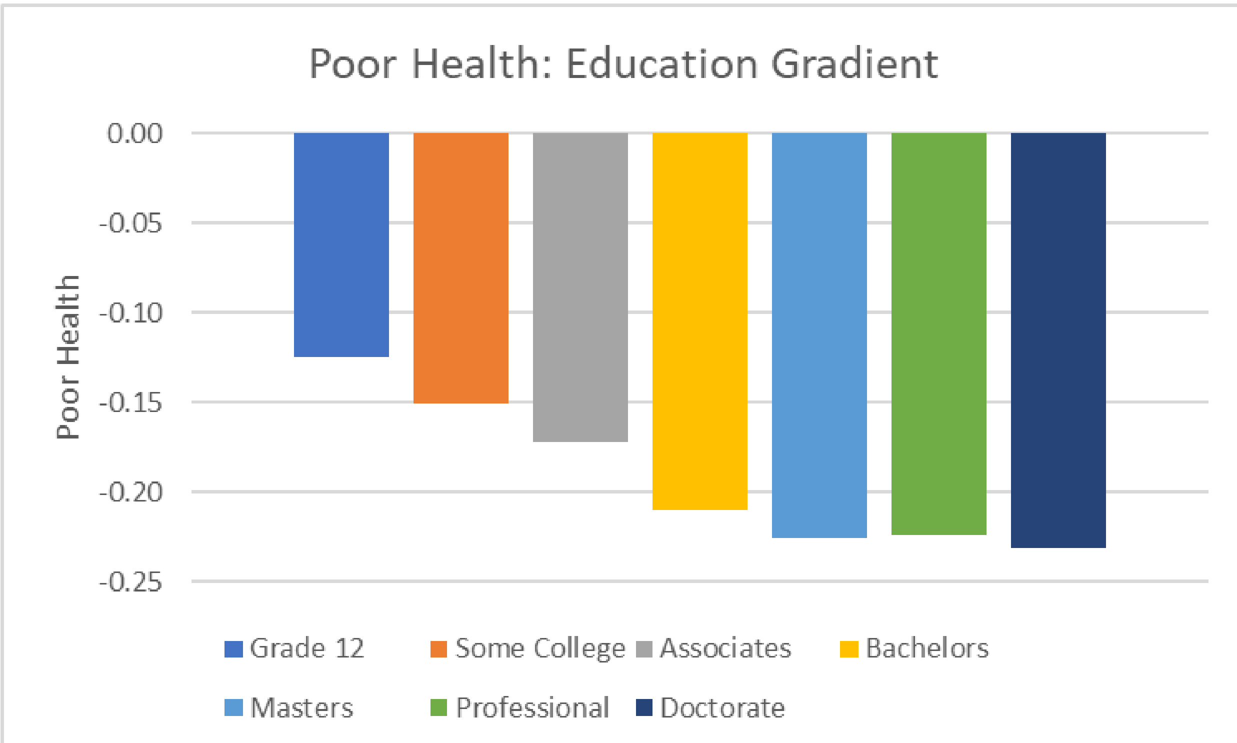
Note: This figure shows the gradient between educational attainment and the health index, a key outcome. The health index was regressed on education indicator variables, one for each bar on the X-axis. Each bar's height shows the estimated relationship between the associated level of educational attainment and the health index, measured in standard deviations. The sample was restricted to individuals with at least an 11th grade education. All estimates are relative to the omitted group, individuals whose highest level of education was 11th grade. Statistics presented here are calculated from publicly available data.

Figure 4.12: Health Costs - Education Gradient



Note: This figure shows the gradient between educational attainment and health costs, a key outcome. Health costs (measured in dollars) were regressed on education indicator variables, one for each bar on the X-axis. Each bar's height shows the estimated relationship between the associated level of educational attainment and health costs, measured in dollars. The sample was restricted to individuals with at least an 11th grade education. All estimates are relative to the omitted group, individuals whose highest level of education was 11th grade. Statistics presented here are calculated from publicly available data.

Figure 4.13: Poor Health - Education Gradient



Note: This figure shows the gradient between educational attainment and the poor health, a key outcome. The poor health indicator was regressed on education indicator variables, one for each bar on the X-axis. Each bar's height shows the estimated relationship between the associated level of educational attainment and poor health, which is a binary variable indicating that the individual reported their health as either "fair" or "poor". The sample was restricted to individuals with at least an 11th grade education. All estimates are relative to the omitted group, individuals whose highest level of education was 11th grade. Statistics presented here are calculated from publicly available data.

5. CONCLUSION

Many factors combine throughout early childhood development and early life years to impact the trajectory of an individual's life. While we may never be able to quantify all of these mechanisms, each channel that we understand provides a lever, both for personal and public policy decisions, that can be used to improve individuals' long run outcomes. Focused specifically on long run health and labor market outcomes, early childhood access to clean drinking water and good nutrition has clear long run benefits for affected children. Because these impacts are focused during early years, children are not the decision makers who determine their access to these key resources—that responsibility is left to parents and policy makers. At older ages, investments in education become driven by individual choice in addition to outside factors like access to financial aid and other resources. While post-secondary educational attainment undoubtedly has the ability to shift long term trajectories in income and health, self-selection into educational attainment may mean that they impact of education itself is smaller than correlational evidence would suggest. Our preliminary results also suggest that, while increased access to financial aid increases educational investments, the effects are limited and may be as small as one fourth the size of previous estimates.

In Chapter I, I showed that water fluoridation, which has been promoted since 1945 as a simple, cost effective, and egalitarian approach to improving dental health, has negative long-term consequences that out weight the dental health benefits. Using restricted U.S. Census data linked to childhood fluoride exposure, I find that children exposed to community water fluoridation from age zero to five experience a 1.9 percent of a standard deviation decrease in their adult economic self-sufficiency, 1.2 percent of a standard deviation increase in rates of adult disability, as well as a 1.5 percentage point decrease in high school graduation.

These results show that the net effect of fluoride is negative even at relatively low levels of exposure. These findings have important policy implications for water fluoridation policies. Fluoride is still being added to a majority of public water supplies in the U.S. and regulations for regions with naturally high levels of fluoride allow water to carry up to 4 mg/L, four times the level of water fluoridation level evaluated in this study. Many regions around the globe have groundwater that is naturally high in fluoride. The results of this study demonstrate the need for a re-evaluation of water fluoridation policies. The observed negative impacts of fluoridated drinking water combined with widespread access to the enamel strengthening benefits of fluoride through toothpaste and dental treatments provides a strong argument for ending the practice of water fluoridation and lowering the maximum levels of fluoride allowed by safe drinking water standards. If water fluoridation practices continue, more research is needed to determine the optimal level of fluoride such that the marginal benefits to dental health are not overwhelmed by negative health and labor market costs. Further study is needed to determine the exact biological mechanisms that are driving these negative effects and discover solutions that mitigate them.

As a direct counterpart to early childhood access to drinking water, Chapter II identifies the importance of early childhood access to nutritional benefits, specifically the impact of increased consumption driven by in-kind food benefits. We leverage a natural experiment where Puerto Rico converted a cash benefit to one in which recipients were required to spend 75% of their benefits on approved food items. This allows us to examine the impact of constraining household consumption decisions in the absence of any shock to overall income.

Using a difference-in-differences strategy, we find that providing the benefits in-kind increases fruit and vegetable consumption by 0.5 servings per day (20 percent). These improvements in nutrition led to decreases in maternal anemia of 0.3-0.7 pp (13-23 percent) and resulted in those who grew up under in-kind benefits being taller and more likely to be normal weight as adolescents.

While the results provide compelling evidence that the form of the benefit can matter, the analyses are not without limitations. It is important to emphasize that the results do not imply that in-kind benefits are welfare improving over cash, but rather that individual spending may not prioritize health, particularly of young children, under a cash-based system. This may have additional implications for the long run costs of cash vs in-kind benefits, particularly when low-income individuals have medical costs that are covered or subsidized by the government.

Turning our attention to the role that education plays as a determinant of long run health and labor market outcomes, Chapter III focuses on the role of increased financial aid in increasing educational attainment, and the subsequent impact of increased post-secondary education on long run outcomes. Understanding the returns to higher education has relevant implications both for policymakers seeking to optimize resource allocation for education as well as individuals considering large investments in either their own or their children's education. Those who attend or complete college have better health and higher incomes but determining the underlying causality of that relationship is difficult.

To answer these questions, we leverage large changes in college attainment generated by the elimination of a large subsidy to higher education, the Student Benefit Program. Using a unique administrative dataset containing Social Security benefit records linked with administrative and survey data, we find that the elimination of the program resulted in a relatively small decrease in educational attainment of 0.185 years, about one fourth the size of previous estimates. Although we can verify that average financial aid fell by over $7,000 in the treatment group, these preliminary results show no statistically significant causal relationship between the end of the student benefit program and long run income, health, or mortality.

These findings have important policy implications for financial aid policies. Strong correlations between educational attainment and labor market and health outcomes are frequently used as support for public and private investments in educational attainment. If the causal effect of education reflects a relatively small portion of these correlations, it suggests a need for careful reevaluation of optimal higher education policy. The results presented in this study are preliminary and future work will include a variety of additional robustness checks and heterogeneity analysis and may focus on more targeted subsamples.

Each chapter of this dissertation contributes to our understanding of the determinants of health and labor market outcomes in adult life. These results broadly support growing evidence that an individual's experiences, setting, and resource access during early childhood and educational years have a profound impact on the long run productivity, income, and health in adulthood. Despite being consistent with that overarching message, the use of large restricted and administrative data sources has resulted in several conclusions that update or in some cases directly contradict previous work. As our understanding of additional factors health and labor market outcomes improve, so does our ability to make informed individual decisions and shape beneficial public policy.

REFERENCES

ADA, American Dental Association. 2019. "ADA Statement on Study in JAMA Pediatrics."

Aggeborn, Linuz, and Mattias Öhman. 2017. "The effects of fluoride in the drinking water." Working Paper.

Almond, Douglas, Hilary W. Hoynes, and Diane Whitmore Schanzenbach. 2011. "Inside the war on poverty: The impact of food stamps on birth outcomes." The Review of Economics and Statistics, 93: 387–403.

Anders, John, Andrew Barr, and Alex Smith. 2019. "The Effect of Early Childhood Education on Adult Criminality: Evidence from the 1960s through the 1990s." Working Paper.

Angrist, Joshua D., and Stacey H. Chen. 2011. "Schooling and the Vietnam-era GI Bill: Evidence from the draft lottery." American Economic Journal: Applied Economics, 3: 96–118.

Arias-Vasquez, F. Javier. 2012. "A note on the effect of education on religiosity." Economics Letters, 117: 895–897.

Bailey, Martha, Hilary W. Hoynes, Maya Rossin-Slater, and Reed Walker. 2020. "Is the social safety net a long-term investment? Large-scale evidence from the food stamps program." Working Paper: National Bureau of Economic Research.

Barr, Andrew, and Alexander A. Smith. 2018. "Fighting Crime in the Cradle: The Effects of Early Childhood Access to Nutritional Assistance." Working Paper.

Barr, Andrew, and Alex Smith. 2021. "Fighting crime in the cradle: The effects of early childhood access to nutritional assistance." Journal of Human Resources.

Bashash, M., D. Thomas, H. Hu, E. Angeles Martinez-Mier, B.N. Sanchez, N. Basu, K.E. Peterson, A.S. Ettinger, R. Wright, Z. Zhang, and Y. Liu. 2017. "Prenatal fluoride exposure and cognitive outcomes in children at 4 and 6–12 years of age in Mexico." Environmental health perspectives, 125.

Baum, Charles. 2007. "The Effects of Food Stamps on Obesity." Southern Economic Journal.

Beltrán-Aguilar, ED, L Barker, and BA Dye. 2010. "Prevalence and Severity of Enamel Fluorosis in the United States, 1986-2004."

Bettinger, Eric, Oded Gurantz, Laura Kawano, Bruce Sacerdote, and Michael Stevens. 2019. "The Long-Run Impacts of Financial Aid: Evidence from California's Cal Grant." American Economic Journal: Economic Policy, 11: 64–94.

Bound, John, and Sarah Turner. 2002. "Going to war and going to college: Did World War II and the GI Bill increase educational attainment for returning veterans?" Journal of labor economics, 20: 784–815.

Buckles, K., A. Hagemann, O. Malamud, M. Morrill, and A. Wozniak. 2016. "The effect of college education on mortality." Journal of Health Economics, 50: 99–114.

Camaschella, Clara. 2015. "Iron-deficiency anemia." New England journal of medicine, 372: 1832–1843.

Card, David. 1995. "Using geographic variation in college proximity to estimate the return to schooling." Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp, 201–222.

Card, David. 2001. "Estimating the return to schooling: Progress on some persistent econometric problems." Econometrica, 69: 1127–1160.

Cesarini, David, Erik Lindqvist, Robert Ostling, and Björn Wallace. 2016. "Wealth, health, and child development: Evidence from administrative data on Swedish lottery players." The Quarterly Journal of Economics, 131: 687–738.

Chaisemartin, Clement De, and Xavier d'Haultfoeuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." American Economic Review, 110.

Choi, Andy I., Cristin C. Weekley, Shu-Cheng Chen, Suying Li, Manjula Kurella Tamura, Keith C. Norris, and Michael G. Shlipak. 2011. "Association of educational attainment with chronic disease and mortality: the Kidney Early Evaluation Program (KEEP)." American Journal of Kidney Diseases, 58: 228–234.

Choi, Anna L, Guifan Sun Ying Zhang, and Philippe Grandjean. 2012. "Developmental fluoride neurotoxicity: A systematic review and meta-analysis." Environmental Health Perspectives, 120: 1362–1368.

Crain, Robert L., Elihu Katz, and Donald B. Rosenthal. 1969. "The politics of community conflict: The fluoridation decision." Bobbs-Merrill Company.

Cunha, Jesse M. 2014. "Testing paternalism: Cash versus in-kind transfers." American Economic Journal: Applied Economics, 6: 195–230.

de Bernabé, Javier Valero, Trinidad Soriano, Romana Albaladejo, Margarita Juarranz, Marıa Elisa Callec, David Martınze, and Vicente Domınguez-Rojas. 2004. "Risk factors for low birth weight: a review." European Journal of Obstetrics Gynecology and Reproductive Biology, 116: 3–15.

Dee, T. 2003. "Are There Civic Returns to Education?" Journal of Public Economics, 88: 1697–1720.

Denning, Jeffrey T. 2019. “Born under a lucky star financial aid, college completion, labor supply, and credit constraints.” Journal of Human Resources, 54: 760–784.

DeSalvo, Karen B., Tiffany M Jones, John Peabody, Jay MacDonald, Stephan Fihn, Vincent Fan, Jiang He, and Paul Muntner. 2009. “Health care expenditure prediction with a single item, self-rated health measure.” Medical care, 440–447.

de Walque, Damien. 2007. “Does Education Affect Smoking Behaviors? Evidence Using the Vietnam Draft as an Instrument for College Education.” Journal of Health Economics, 26: 877–895.

DeWitt, Larry. n.d.. “The History of Social Security “Student" Benefits.” SSA Historian’s Office, January 2001, `https://www.ssa.gov/history/studentbenefit.html`.

DHHS. 2015. “Public Health Service Recommendation for Fluoride Concentration in Drinking Water for the Prevention of Dental Caries.”

Dincer, Mehmet Alper, Neeraj Kaushal, and Michael Grossman. 2014. “Women’s education: Harbinger of another spring? Evidence from a natural experiment in Turkey.” World Development, 64: 243–258.

Dye, BA, X Li, and G. Thornton-Evans. 2012. “Oral health disparities as determined by selected Healthy People 2020 oral health objectives for the United States, 2009-2010.”

Dynarski, Susan M. 2003. “Does aid matter? Measuring the effect of student aid on college attendance and completion.” American Economic Review, 93.1: 279–288.

Fletcher, Jason M. 2015. “New evidence of the effects of education on health in the US: Compulsory schooling laws revistited.” Social Science and Medicine, 127: 101–107.

Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles, and J. Robert Warren. 2020. “Integrated Public Use Microdata Series, Current Population Survey: Version 7.0 [dataset].”

Fox, Mary Kay, William L. Hamilton, and Biing-Hwan Lin. 2004. "Effects of food assistance and nutrition programs on nutrition and health: Volume 4, executive summary of the literature review." No. 1481-2016-121334.

Fraker, Thomas M., Alberto P. Martini, and James C. Ohls. 1995. "The effect of food stamp cashout on food expenditures: An assessment of the findings from four demonstrations." Journal of Human Resources, 633–649.

Fujiwara, T., and I. Kawachi. 2009. "Is education causally related to better health? A twin fixed-effect study in the USA." International Journal of Epidemiology, 38(5): 1310–1322.

Gertler, Paul. 2004. "Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment." The American Economic Review, 94.

Glied, Shery, and Matthew Neidell. 2010. "The economic value of teeth." Journal of Human Resources, 45: 468–496.

Gooch, Barbara F. 2020. "Preventing Chronic Disease Dialogue: Community Water Fluoridation - One of the 10 Greatest Public Health Achievements of the 20th Century."

Goodman-Bacon, Andrew. 2018. "Difference-in-differences with variation in treatment timing." No. w25018. National Bureau of Economic Research.

Grandjean, Philippe. 2019. "Developmental fluoride neurotoxicity: an updated review." Environmental Health, 18: 1–17.

Green, Rivka, Bruce Lanphear, Richard Hornung, David Flora, E. Angeles Martinez-Mier, Raichel Neufeld, Pierre Ayotte, Gina Muckle, and Christine Till. 2019. "Association between maternal fluoride exposure during pregnancy and IQ scores in offspring in Canada." 173: 940–948.

Grimard, Franque, and Daniel Parent. 2007. "Education and Smoking: Were Vietnam War Draft Avoiders Also More Likely to Avoid Smoking?" Journal of Health Economics, 26: 896–926.

Habershon. 1863. "On Idiopathic Anemia." The Lancet, 81: 518–519.

Hastings, Justine, and Jesse M. Shapiro. 2018. "How are SNAP benefits spent? Evidence from a retail panel." American Economic Review, 108: 3493–3540.

Hastings, Justine, Ryan Kessler, and Jesse Shapiro. 2018. "The Effect of SNAP on the Composition of Purchased Foods-Evidence and Implications." Working Paper.

Heckman, James J., John Eric Humphries, and Gregory Veramendi. 2018. "Returns to education: The causal effects of education on earnings, health, and smoking." Journal of Political Economy, 126: S197–S246.

Heller, Keith E., Stephen A. Eklund, and Brian A. Burt. 1997. "Dental caries and dental fluorosis at varying water fluoride concentrations." Journal of Public Health Dentistry, 57: 136–143.

HHS. 2016. "CMS Releases 2015 National Health Expenditures."

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-run impacts of childhood access to the safety net." American Economic Review, 106: 903–934.

Hoynes, Hilary, Diane Whitmore Schanzenbach, , and Douglas Almond. 2016. "Long-run impacts of childhood access to the safety net." American Economic Review, 106: 903–934.

Hoynes, Hilary W., and Diane Whitmore Schanzenbach. 2009. "Consumption responses to in-kind transfers: Evidence from the introduction of the food stamp program." American Economic Journal: Applied Economics, 1: 109–39.

Hoynes, Hilary Williamson, and Diane Whitmore Schanzenbach. 2012. "Work incentives and the food stamp program." Journal of Public Economics.

Hungerman, Daniel M. 2014. "The effect of education on religion: Evidence from compulsory schooling laws." Journal of Economic Behavior and Organization, 104: 52–63.

Idler, Ellen L., and Pael Benyamini. 1997. "Self-rated health and mortality: a review of twenty-seven community studies." Journal of Health and Social Behavior, 21–37.

James, Jonathan, and Sunčica Vujić. 2019. "From high school to the high chair: Education and fertility timing." Economics of Education Review, 69: 1–24.

Johnston, Robert D. 2004. The politics of healing: Histories of alternative medicine in twentieth-century North America. Psychology Press.

Kaplan, George A., Debbie E Goldberg, Susan A Everson, Richard D Cohen, Riita Salonen, Jaakko Tuomilehto, and Jukka Salonen. 1996. "Perceived health status and morbidity and mortality: Evidence from the Kuopio ischaemic disease risk factor study." International journal of epidemiology, 25: 259–265.

Kemptner, Daniel, Hendrik Jürges, and Steffen Reinhold. 2011. "Changes in compulsory schooling and the causal effect of education on health: Evidence from Germany." Journal of health economics, 30: 340–354.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental analysis of neighborhood effects." Econometrica, 75.1: 83–119.

Lee, Bong Joo, and Lucy Mackey-Bilaver. 2006. "Effects of WIC and Food Stamp Program participation on child outcomes." Children and Youth Services Review.

Lim, E. L., K. G. Hollingsworth, B. S. Aribisala, M. J. Chen, J. C. Mathers, and R. Taylor. 2011. "Reversal of type 2 diabetes: normalisation of beta cell function in association with decreased pancreas and liver triacylglycerol." Diabetologia, 54: 2506–2514.

Lleras-Muney, Adriana. 2005. "The relationship between education and adult mortality in the United States." The Review of Economic Studies, 72: 189–221.

Lundborg, P. 2013. "The health returns to schooling—what can we learn from twins?" Journal of Population Economics, 673–701.

Lundborg, Petter, Martin Nordin, and Dan Olof Rooth. 2018. "The intergeneration transmission of human capital: the role of skills and health." Journal of Population Economics, 31: 1035–1065.

MacInnis, B. 2006a. "Does College Education Impact Health? Evidence from the Pre-Lottery Vietnam." Ann Arbor: University of Michigan. Unpublished.

MacInnis, B. 2006b. "The long-term effects of college education on morbidities: New Evidence from the pre-lottery Vietnam draft." Draft presented at the NBER Summer Institute.

Mazumder, Bhashkar. 2008. "Does education improve health? A reexamination of the evidence from compulsory schooling laws." Economic Perspectives, 32.

McCrary, Justin, and Heather Royer. 2011. "The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth." American Economic Review, 101: 158–195.

Miller, Douglas L., Na'ama Shenhav, and Michel Z. Grosz. 2019. "Selection into identification in fixed effects models, with application to Head Start." National Bureau of Economic Research, No. w26174.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom." Journal of Public Economics, 88: 1667–1695.

Moffitt, Robert. 1989. "Estimating the value of an in-kind transfer: The case of food stamps." Econometrica: Journal of the Econometric Society, 385–409.

Mullenix, P. J., P. K. Denbesten, A. Schunior, and W. J. Kernan. 1995. "Neurotoxicity of sodium fluoride in rats." 17: 169–177.

Newacheck, P. W., D. C. Hughes, Y. Y. Hung, S. Wong, and J. J. Stoddard. 2000. "The unmet health needs of America's children." Pediatrics, 105: 989–997.

Ohls, James C., and Laura Bernson. 1992. "The effects of cash-out on food use by food stamp program participants in San Diego."

Oreopoulos, P., and K. G. Salvanes. 2011. "Priceless: The Nonpecuniary Benefits of Schooling." Journal of Economic Perspectives, 25(1): 159–184.

Oreopoulos, Philip. 2007. "Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling." Journal of Public Economics, 91: 2213–2229.

Oreopoulos, Philip, and Uros Petronijevic. 2013. "Making college worth it: A review of research on the returns to higher education." National Bureau of Economic Research Working Paper Series, w19053.

Ran, Tao, and Sajal K. Chattopadhyay. 2016. "Economic evaluation of community water fluoridation: a community guide systematic review." American journal of preventive medicine, 50: 790–796.

Roberts, Adam. 2021. "Evaluating the Long-Run Effects of Water Fluoridation: Evidence from Community Water Fluoridation Programs - A Pre-Analysis Plan." Open Science Framework.

Sapolsky, Harvey M. 1968. "Science, Voters, and the Fluoridation Controversy: Conflict among perceived experts leads voters to act negatively on the fluoridation innovation." Science, 162: 427–433.

Schoenaker, Danielle A.J.M., Gita D. Mishra, Leonie K. Callaway, and Sabita S. Soedamah-Muthu. 2016. Diabetes Care, 39: 16–23.

Singh, K. A., A. John Spencer, and J. M. Armfield. 2003. "Relative effects of pre-and posteruption water fluoride on caries experience of permanent first molars." Journal of public health dentistry, 63: 11–19.

SSA. 1982. "Social Security bulletin annual statistical supplement." Social Security Administration, Washington, DC: U.S. Government Printing Office `https://catalog.hathitrust.org/Record/000499755`.

Stanley, Marcus. 2003. "College education and the midcentury GI Bills." The Quarterly Journal of Economics, 118: 671–708.

Sun, Liyang, and Sarah Abraham. 2020. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." Journal of Econometrics.

Taylor, Evan J, Bryan A Stuart, Martha J Bailey, et al. 2016. "Summary of Procedure to Match NUMIDENT Place of Birth County to GNIS Places." Center for Economic Studies, US Census Bureau.

Tequame, Miron, and Nyasha Tirivavi. 2015. "Higher education and fertility: Evidence from a natural experiment in Ethiopia." UNU-MERIT Working Paper Series.

Trippe, Carole, Rachel Gaddes, Alexandra Suchman, Kate Place, James Mabli, Chrystine Tadlerand Teresa DeAtley, , and Brian Estes. 2015. "Examination of Cash Nutrition Assistance Program Benefits in Puerto Rico." Pre-pared by Insight Policy Research under Contract

No. AG-3198-C-14-0006. Alexandria, VA: U.S. Department of Agriculture, Food and Nutrition Service.

V.Amin, J. R. Behrman, and T. D. Spector. 2013. "Does more schooling improve health outcomes and health related behaviors? Evidence from U.K. twins." Economics of Education Review, 35: 134–148.

Wagner, Deborah, and Mary Lane. 2014. "The person identification validation system (PVS): applying the Center for Administrative Records Research and Applications'(CARRA) record linkage software." Center for Economic Studies, US Census Bureau, 2014-01.

Whitake, Robert, Shannon Philips, and Sean Orzol. 2006. "Food Insecurity and the Risks of Depression and Anxiety in Mothers and Behavior Problems in their Preschool-Aged Children." Pediatrics.

Whitmore, Diane. 2002. "What are food stamps worth?" Working Paper.

Zimmerman, Seth D. 2014. "The returns to college admission for academically marginal students." Journal of Labor Economics, 32: 711–754.