ESSAYS IN HEALTH ECONOMICS

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Barton Jay Willage

May 2018

ESSAYS IN HEALTH ECONOMICS

Barton Jay Willage, Ph.D., M.P.A., M.A.

Cornell University 2018

This dissertation is comprised of three essays, each examining a different topic in health economics. In Chapter 1, I investigate ex ante moral hazard from health insurance, or the effect of insurance on risky behaviors. Specifically, I examine risky sex, a behavior with quick, meaningful consequences. I exploit the requirement insurance cover prescription contraception without cost-sharing. Since this policy affects a small aspect of insurance, I isolate the effect on behavior. Using pre-policy insured rates as a measure of treatment intensity, I find insurance decreases prevention (condom sales) and increases sexually transmitted infections (STIs). To determine insurance's overall effect, I exploit the requirement insurance cover young adult children. I find the protective effect of insurance on STIs more than compensates for the reduction in prevention.

In Chapter 2, I test whether weight impacts mental health. Both average body mass index (BMI) and rates of depression increased over the past several decades; however, scant research has investigated if the correlation between weight and depression reflects a causal effect. Using genetic variation in obesity predisposition as an instrument, I determine the impact of BMI on suicidality and depression. I find BMI impacts suicidal ideation but not other measures of mental health. This effect on suicidal ideation is concentrated in white women. One mechanism through which weight affects mental health is social stigma.

In Chapter 3, I investigate the effect of financial aid on health, specifically BMI and general self-report health. Aid lowers the cost of college, which increases college-going. However, there is little evidence on the long-run impact of aid on health. I exploit a 1981 shock in Social

Security benefits to test the effect of aid on health. Minor children of retired or disabled Social Security beneficiaries and children with deceased parents are eligible for their own benefits, and before 1981 these children could receiving benefits conditional on college enrollment. Using difference-in-differences, I show benefits for college students reduced women's BMI and improved general health, but had no effect on men. I also find that aid improved educational attainment for beneficiaries, which is a plausible mechanism between aid and health.

BIOGRAPHICAL SKETCH

Barton Willage grew up primarily in Southern Indiana. He graduated with a Bachelor of Arts degree in Mathematics and Political Science from Beloit College. Before starting his doctoral studies at Cornell University, he earned a Master of Arts in the Program in the Social Sciences at the University of Chicago and a Master of Public Affairs from Indiana University's School of Public and Environmental Affairs.

ACKNOWLEDGMENTS

I am indebted to my dissertation advisors John Cawley, Donald Kenkel, and Michael Lovenheim. My special committee chair, John Cawley, has been extremely helpful and supportive of my research agenda. Additionally, he has provided me with outstanding advice and professional development opportunities. I have learned a great deal from him in his many roles: classroom teacher, dissertation advisor, and research assistant supervisor. I thank Donald Kenkel for his valuable feedback on my research as well as for being a constant source of interesting history. His door is almost always open, and he was always available to chat. I also thank Michael Lovenheim for his constant encouragement and energy. I am extremely grateful for his investment of both time and energy in the success of my research; he read multiple drafts and listened to multiple presentations. My success during my doctoral studies and on the job market are in large part due to my advisors.

This research was partially funded by a Human Ecology Alumni Association Research Grant. I would also like to thank discussants and participants at American Society of Health Economists Conference, Association for Public Policy & Management Regional Student Conference, Association for Public Policy & Management Fall Conference, International Health Economics Association Congress as well as Cornell University's Policy Analysis and Management Seminar and Institute on Health Economics, Health Behaviors & Disparities Seminar. I also thank the many Cornell University graduate students who have read drafts of these papers, particularly Bridget Brew.

Lastly, I thank my parents for helping me in countless ways as I pursued many years of education on this journey to a Ph.D.

TABLE OF CONTENTS

CHAPTER 1

Ex Ante Moral Hazard in Health Insurance: Lessons from the ACA and Risky Sex

Abstract: Given the $3 trillion spent on health care in 2015 and the political contention surrounding insurance expansions, the impact of health insurance on health behaviors, medical utilization, and health outcomes continues to be of the upmost importance. How insurance influences investment in good health and risky behavior (ex ante moral hazard) has received much less attention than the effect of insurance on the out-of-pocket cost of care (ex post moral hazard). Since many risky health behaviors take decades to result in illness, these behaviors likely respond to expectations about future insurance but could be unaffected by current insurance status. I examine the effect of moral hazard in the context of risky sex, a health behavior that results in quick and economically meaningful consequences – fertility and sexually transmitted infections. I isolate the effect of ex ante moral hazard by exploiting a policy in the Affordable Care Act, the 2012 zero cost-sharing for prescription contraception mandate. Leveraging pre-policy insured rates as a measure of policy intensity, I use dose-response event studies that estimate both a time-varying treatment effect as well as a one-time jump in outcomes in the treatment year. I find evidence ex ante moral hazard from health insurance decreases prevention and increases STIs. I then exploit the 2010 dependent coverage mandate to determine the overall effect of health insurance. Based on this policy I find that the protective effect of insurance on STIs more than compensates for the reduction in prevention.

1. Introduction

Health insurance has played a prominent role in the rapid growth of medical spending over the past half century (Finkelstein, 2007; Finkelstein et al., 2012; Manning et al., 1987). While health insurance increases utility by smoothing consumption, insurance can also have unintended consequences, known as moral hazard. Ex ante moral hazard is the effect health insurance has on risky health behaviors and investment in prevention. This kind of moral hazard receives less attention than ex post moral hazard, the increased quantity of care demanded due to lower out-of-pocket cost (Zwiefel and Manning, 2000).[1] However, determining the effect of insurance on risky behavior is important because these distortions can cause increased illness and medical spending. Additionally, ex ante moral hazard causes negative externalities; the financial burden is spread across the insurance risk pool and infections can be transmitted to individuals who have not changed their behavior.

Most studies that do examine ex ante moral hazard find little or inconsistent evidence that health insurance changes smoking, diet, exercise, and drinking (Dave and Kaestner, 2009; De Preux, 2011; Barbaresco et al., 2015; Simon et al., 2017).[2] A major obstacle to observing ex ante moral hazard is that many health shocks occur

[1] There is a large literature on moral hazard in other types of insurance markets as well. See Cummins and Tennyson (1996) for a car insurance example and Chetty (2008) for an unemployment insurance example. However, health insurance has important distinguishing characteristics. For instance, health insurance covers maintenance and prevention (such as prescription contraception), while car insurance does not cover maintenance such as oil changes. Additionally, the distinction between treatment and prevention is blurry in health; for instance, statins treat high cholesterol but also prevent heart attack. Most importantly for changes in prevention, other types of insurance can provide replacements (cars, houses, income), while health insurance can often only provide access to treatment that may not completely cure the illness or repair the injury.

[2] Dave and Kaestner (2009) examine exercise, smoking, alcohol consumption. De Preux (2011) examines exercise, smoking, alcohol consumption. Barbaresco et al. (2015) examine exercise, smoking, alcohol consumption, BMI/obesity, pregnancy. Simon et al. (2017) examine exercise, smoking, alcohol consumption, BMI/obesity.

only after years or decades of poor health behaviors, for example smoking and cancer (Department of Health and Human Services, 2010). As a result, future, rather than current, health insurance covers the eventual consequences of many current risky behaviors.

This paper empirically tests for ex ante moral hazard in risky sex decisions and the ensuing health consequences. The risky sex behavior I consider is sex between a man and a woman without a condom.[3] There are two advantages of focusing on risky sex. First, many health consequences of risky sex such as unplanned pregnancy and sexually transmitted infection (STI) occur quickly, so changes to insurance status should affect risky sex. Few risky health behaviors can cause adverse health shocks as quickly as unprotected sex. One exception is drug-use can result in an overdose and mortality almost immediately, and a recent working paper suggests drug-use is another risky behavior that is responsive to ex ante moral hazard (Doleac and Mukherjee, 2018).

The second benefit of focusing on risky sex is that a recent policy change in the Patient Protection and Affordable Care Act (ACA) provides a unique opportunity to isolate an ex ante moral hazard effect. The zero cost-sharing for prescription contraception mandate of 2012 made prescription birth control free for insured women, increasing health insurance on the intensive margin or the degree to which each person is insured. The zero cost-sharing mandate allows me to isolate an ex ante moral hazard effect, because it affects only a single aspect of health insurance: the cost

[3] I acknowledge that risky sex can occur in other contexts, such as with men who have unprotect sex with men. Here I focus on heterosexuals who should respond to the zero cost-sharing mandate.

of prescription contraception. This policy was implemented in all states simultaneously, so its effect cannot be determined by comparing treated and untreated states. To overcome this obstacle to identification, I use pre-policy insurance rates as a measure of treatment intensity, similar to the approach employed by Finkelstein (2007).

I use state-year level data and dose-response event studies to determine the effect of this policy on several outcomes related to risky sex, such as condom sales, STI incidence, and fertility. The measure of treatment intensity I use is the pre-policy insured rate among 25-to-29-year-olds in each state. The insured rate represents the percent of 25-to-29-year-olds exposed to the zero cost-sharing mandate, because the mandate only applies to insured individuals.[4] Using non-parametric event study analysis, I identify any change in relative trends that occur in the treatment year by interacting the treatment intensity in each state with year dummies. I then impose a linear parameterization on the event studies. In this analysis, I model the pre-policy trend and intercept in the event studies as well as the change in trend and one-time jump that occur in the year of policy implementation. The benefits of the parametrization are increased statistical power and clearer estimates, but it requires the assumption that trends are approximately linear.

Economic theory provides clear predictions of the ex ante moral hazard effects of the zero cost-sharing mandate. Prescription contraception decreases the cost of having sex without a condom and should decrease demand for condoms. Condoms and

[4] See Courtemanche et al. (2017) and Finkelstein (2007) for two examples using similar sources of exogenous variation in other contexts.

prescription birth control are substitutes when it comes to preventing pregnancy, but condoms also prevent the transmission of STIs. Based on the dual role of condoms, substitution to prescription contraception should lead to increased incidence of STIs. While much of the ex ante moral hazard literature considers the substitution from prevention to treatment, I am examining a different pathway between health insurance and risky behaviors: substitution from broad behavior-based prevention to narrow medical-based prevention. I find that the estimated effects of the zero cost-sharing mandate are consistent with ex ante moral hazard: a reduction in condom sales and increased incidence of chlamydia.

An increase in STIs is not consequence of all health insurance expansions. As an extension, I examine the young adult dependent coverage mandate of 2010 that allowed children 25 and under to join their parents' health insurance. This mandate caused an exogenous shock to the extensive margin of health insurance, which I leverage to determine the overall effect of health insurance. I use the same empirical strategy but now the measure of treatment intensity is the uninsured rate among people in their early 20s. I find that this policy also reduced investment in prevention. However, the protective effect of health insurance more than offset the reduction in prevention; the dependent coverage mandate resulted in fewer STIs.

This study makes two main contributions to the literature. First, it empirically tests for ex ante moral hazard in risky sex. Several studies have examined the effect of health insurance on health behaviors, but largely in contexts where health behaviors are slow to result in disease and are less likely to respond to health insurance coverage (Dave and Kaestner, 2009; De Preux, 2011; Barbaresco et al., 2015; Simon et al.,

2017).

The second main contribution is an evaluation of important and contentious aspect of the ACA. For instance, almost 90 outside groups submitted briefs during *Burwell v. Hobby Lobby,* which determined that certain corporations do not have to pay for insurance plans that cover prescription birth control (Supreme Court of the United States, 2014). Furthermore, information on the effect of this mandate provides insight on recent and current policy proposals. The executive branch recently drafted a rule limiting the zero cost-sharing mandate (Wolf, 2017), greatly expanding the ability of employers to obtain an exemption from the requirement to offer insurance that covers birth control at no out-of-pocket cost. Additionally, Congress recently considered repealing the ACA, which may have ended this mandate (Kaplan and Pear, 2017). Evidence from the implementation of this policy offers suggestive evidence about the effect of proposals that would eliminate or reduce it.

There is also great policy interest in the outcomes I examine: fertility and STIs. Fertility, especially unintended pregnancy, is a very expensive consequence of unprotected sex and is often paid for by public insurance. The government spends an estimated $12 billion on unwanted pregnancies each year (Thomas and Monea, 2011). Additionally, some STIs are becoming increasingly resistant to treatment. Bacterial STIs were previously easy to treat, but complications from STIs, such as pelvic inflammatory disease, are an increasing concern (Hersher, 2016). In fact, the World Health Organization (2017) prioritized gonorrhea as one of the eleven most important antibiotic resistant bacteria. This study also produces insight on an important outcome missing from many analyses of risky sex: condoms. By analyzing condom sales, I

provide evidence on one mechanism through which insurance affects STI incidence and fertility. Finally, in contrast to the many studies on Medicaid expansion and Medicare that focus on low income or older populations, I focus on a largely understudied group; the marginal individual in this context is a middle class young adult with private insurance or with privately insured parents.

The zero cost-sharing mandate lowered the cost of prescription contraception but not of condoms, which resulted in a reduction in condom sales. One way to counteract the increase of risky sex would be to also subsidize condoms. Another policy concern is that while ex ante moral hazard causes increased health care utilization, spending on medical care does not reflect the full economic loss of reduced prevention. Ex ante moral hazard results in more health shocks, so people experience utility loss directly from illness and injury.

The remainder of this paper proceeds as follows: Section 2 provides background including a review of the existing literature and information on the zero cost-sharing mandate. Section 3 describes the economic framework. Section 4 details the data sources and empirical research method. Empirical results are presented in Section 5. In Section 6, I examine the extensive margin of health insurance to determine the overall effect of health insurance using another policy in the ACA, the mandate that adult children under the age of 26 be allowed on their parents' insurance. Robustness and falsification tests are discussed in Section 7. Finally, Section 8 concludes with a discussion.

2. Background

2.1. Literature Review

This study sits at the intersection of two literatures: (1) responses to health insurance, particularly ex ante moral hazard effects on risky behaviors and prevention, and (2) the economics of sexual activity. While not the focus of this study, the theoretical work on ex post moral hazard in health insurance starts with Pauly (1968). Both the RAND and Oregon health insurance experiments showed strong empirical evidence of ex post moral hazard (Manning et al., 1987; Finkelstein et al., 2012). Finkelstein (2015) provides an overview of this literature. The theoretical work on ex ante moral hazard in health insurance starts with Ehrlich and Becker (1972). However, the empirical evidence of ex ante moral hazard is much less consistent than the evidence of ex post moral hazard.

Generally, researchers find weak or mixed empirical evidence of ex ante moral hazard; many studies find effects on a small subset of examined health behaviors or find effects only among certain demographic groups. A common strategy to examine the causal effect of insurance on health behaviors is to leverage the exogenous change in insurance status caused by aging into Medicare eligibility (Dave and Kaestner, 2009; De Preux, 2011). While most people entering Medicare are 65 years old, other studies examine policies that affect younger populations. For example, Barbaresco et al. (2015) study the population targeted by the 2010 requirement that insurers cover adult dependents under 26 years of age, comparing changes in 23-to-25-year-olds to 27-to-29-year-olds. Simon et al. (2017) compare states that did and did not expand Medicaid coverage to low-income childless adults.

The studies on the effect of health insurance on risky health behaviors primarily examine effects on smoking, exercise, and drinking. The lack of evidence of ex ante moral hazard may in part be related to the set of outcomes studied, because these health behaviors often do not result in health shocks for many years. To overcome this obstacle I examine risky sex, which has a short lag before resulting in health shocks such as pregnancy and STIs.

In contrast to the lack of consensus about the effect of insurance coverage on risky behaviors, the literature on the economics of sex generally finds that lowering the cost of sex without a condom increases health shocks, particularly STIs (Chesson, 2012). Klick and Stratmann (2007) and Levine (2003) examine state laws that require minors to inform or involve their parents in order to obtain an abortion. These studies show that such laws resulted in fewer abortions, fewer cases of gonorrhea, and fewer pregnancies. Ressler et al. (2006) find that increasing cash welfare payments, which decreased the cost of having a child, increased rates of sexually transmitted infection. Similarly, Ahituv et al. (1996) determine that condom use increased when the cost of unprotected sex (AIDS prevalence/risk of infection) increased.

However, not all studies find that lowering the cost of sex without a condom increased sex-related health shocks. For instance, easier access to emergency contraception did not affect fertility or abortion rates (Gross et al, 2013). In certain contexts, even easier access to condoms did not reduce the number of pregnancies or STIs. Looking at school-based programs that distributed condoms to teens, Buckles and Hungerman (2016) find that these programs increased teen pregnancy, particularly if additional information was not provided with condoms. Conversely, Lovenheim et

al. (2016) find that expansion of school-based health centers, which provide access to prescription birth control and often condoms, led to lower teen fertility. I add to the literature on risky sex by examining a different source of exogenous variation in the cost of risky sex: health insurance expansion.

With some exceptions, existing research on the cost of risky sex focuses on births and diseases, while ignoring the effects on actual behaviors such as use and purchases of condoms. Understanding the effect on condom purchases helps confirm that changes in fertility and infection are due to changes in risky sexual behavior and not due to an unobserved contemporaneous shock. The lack of evidence on these outcomes is primarily driven by data limitations. Questions about use of condoms and prescription contraception are not even included in the surveys most likely to ask about these behaviors, such as recent waves of the Behavioral Risk Factor Surveillance System. I address this gap with proprietary data on condom sales.

2.2. Policy Background

President Obama signed the ACA into law in March 2010. The ACA was the most significant legislative change to the health care system in the 50 years since passage of Medicare and Medicaid (Oberlander, 2010). Unlike Medicare and Medicaid, the ACA is primarily a market-based health insurance expansion.[5] For instance, two of the most well-known aspects of the ACA are the individual mandate and the health insurance exchanges, which caused major changes to the private health insurance system (Kaiser Family Foundation, 2013). The individual mandate requires that every individual have

[5] One major aspect of the ACA that does involve expansion of government-based health insurance is Medicaid coverage of childless adults, a group generally not eligible for Medicaid pre-ACA.

comprehensive health insurance. The insurance exchanges are online marketplaces to compare plans and purchase health insurance. Neither of these policies involve the government directly providing health insurance, instead they leverage and expand the existing private health insurance market.[6]

This study uses an adjustment to private health insurance markets made by the ACA to test for ex ante moral hazard. The zero cost-sharing mandate requires insurance plans to cover prescription contraception with no out-of-pocket cost starting in August 2012 (Health Resources & Service Administration, 2017). At least one version of each form of prescription contraception (e.g. oral, injectable, intrauterine device) must be covered with no out-of-pocket expense, but there is no requirement that branded versions be covered with no cost-sharing if a generic option of the method is available (Centers for Medicare & Medicaid Services, 2015). This policy ensures 47 million women can access prescription contraception and other preventive care with no deductible, co-pay, or co-insurance (Simmons and Skopec, 2012).

The zero cost-sharing mandate affects the intensive margin of health insurance, because it changes the degree of coverage by requiring zero cost-sharing for certain benefits. Importantly, this policy went into effect before much of the ACA, such as the establishment of the insurance exchanges, the requirement that individuals have insurance, or the bulk of Medicaid expansion to childless adults (Senate.gov, 2010),[7] which reduces concern about contemporaneous policy shocks. While the requirement

[6] For more details on these and other aspects of the ACA, see Kaiser Family Foundation (2013).
[7] In fact, many other early aspects of the ACA did not directly affect patients, and instead focused on health care institutions and infrastructure. For more details, see the implementation timeline provided by the U.S. Senate (Senate.gov, 2010).

that young adults be allowed on their parents' insurance began in 2010, I focus my analysis on an older population unaffected by the dependent coverage mandate. Additionally, there is less concern about policy timing endogeneity, because I use a change in federal law instead of state-level policies.

The zero cost-sharing mandate had a meaningful effect on both the out-of-pocket cost of prescription contraception and contraception use. Between 2012 and 2014, the percent of privately insured women who paid $0 out-of-pocket for contraception increased by 30-50 percentage points across methods (oral, injectable, ring, intrauterine device) (Bearak et al., 2016; Sonfield et al., 2015). The median out-of-pocket cost fell from $10 to $0 for oral contraception and from $20 to $0 for intrauterine devices (IUD). Even a few hundred dollars can be meaningful to low-income women, but the reduction in cost was much higher for many women. For example, the cost of an IUD at the 90th percentile dropped from $844 to $15, though some uninsured women and women working for religiously-exempt employers still bear at least some financial burden (Bearak et al., 2016; Sonfield et al., 2015). An analysis of women working in 499 Midwest firms that provide health insurance found this policy caused a 2.3 percentage point or 7.6% increase in prescription contraception use (Carlin et al., 2016).[8]

3. Theoretical Framework

3.1: Moral Hazard

While health insurance provides protection against the financial burden caused by

[8] There is evidence in the behavioral economics literature that reducing the price to $0 can be significantly more effective than reductions to small, non-zero prices (e.g., Shampanier et al., 2007).

health shocks, it can also have unintended consequences known as moral hazard. While much of the literature has emphasized the ex post moral hazard effect due to lower out-of-pocket costs of medical care, I focus on ex ante moral hazard. This effect is the distortion caused by lowering the expected cost of risky or unhealthy behaviors. This increase in risky behavior, in turn, increases the probability of a subsequent negative health shock and increased medical spending. Figure 1 shows that lowering the cost of risky behaviors due to health insurance results in greater demand for risky behaviors. The wedge between the quantity of risky behaviors demanded with and without insurance is ex ante moral hazard.

For example, the expected cost of skateboarding is lower if someone has insurance because his insurance will cover medical expenses in the event of an injury. A person who gains health insurance may be more likely to take up skateboarding because the expected cost is lower. Insurance can also increase risky behaviors through substitution to narrower, medical-based forms of prevention – such as prescription birth control or statins – and away from more comprehensive, behavior-based methods – such as abstinence/condoms or diet/exercise. Insured people have cheaper access to birth control, and removing the risk of unintended pregnancy lowers the expected cost of having sex without a condom. This reduction in cost should result in increased demand for risky behaviors such as unprotected sex and sedentary lifestyles.

3.2. Economic Model of Insurance and Sex Decisions

In this section, I present a two-period model of risky sex behavior involving health insurance, prevention, health shocks, and treatment. While economists have modeled

the effect of certain "costs" on sex decision-making,[9] the effect of insurance on the costs and benefits surrounding sex has not received much attention (Arcidiacono et al., 2012; Gross et al., 2014; Chan et al., 2016). I use my model to map out the potential pathways through which health insurance affects risky sex and health outcomes as well as to derive testable hypotheses. I neither attempt to determine the overall effect on utility of health insurance by solving my model nor to structurally estimate model parameters. The model leads to clear predictions of the effect of lowering the cost of prescription contraception: an increase in STIs and a decrease in fertility. As I discuss in Section 6, the overall effect of insurance at the extensive margin is largely ambiguous.

3.2.1: Choices, Parameters, and Outcomes

The two-period model contains three types of variables: *choices* that are completely within the control of a person, *parameters* that a person cannot control, and *outcomes* that are affected by the choices and parameters but are not deterministic.

There are two groups of choices. The choices in the first period are: to use prescription birth control or not (BCRx), to have sex or not (S), and to use a condom or not (C). These choices influence the probability of an adverse health shock, and distortions to these choices are ex ante moral hazard. The probabilistic outcomes – pregnancy (P) and STI infection (STI) – occur between the two rounds of choices. The choices in the second period are: to bring a pregnancy to term or not (B as in birth), and to treat STIs or not (T).

There are two groups of parameters that a person cannot influence – preferences

[9] These costs include both monetary and implicit costs such as probability of infection.

(θ_{pref}) and policies (θ_{policy}) – that determine the environment in which choices are made. Preferences include desire for a child, inclination to use condoms, as well as utility derived from all choices and outcomes. The policy parameters include whether or not prescription contraception is free, as well as the effect this policy has on the prevalence of STIs in potential sexual partners. Changes in the policy parameter directly affect prices; for instance, the zero cost-sharing mandate decreases the cost of prescription contraception.

3.2.2: Uncertainty in the risk of pregnancy and STI infection

In this model, I assume that any form of prevention (condoms or prescription birth control) is a hundred percent effective,[10] but in the absence of prevention there is uncertainty in outcomes. The probability of pregnancy is assumed to be zero if condoms or birth control is used or if a person abstains from sex. The probability of pregnancy in the absence of prevention is some fixed value π. If a person refrains from having sex or uses condoms, I assume the probability of STI infection is zero. The probability of contracting an STI in the case of unprotected sex is an increasing function of the prevalence of STIs in the potential partner pool: If more people have STIs, then the greater the likelihood that unprotected sex will result in a new case of an STI.

[10] This is a simplification; condoms are about 80% effective in pregnancy prevention, the pill about 90%, and IUDs over 99% (Centers for Disease Control and Prevention, 2017a). Relaxing this assumption by assuming less than hundred percent effectiveness does not result in meaningfully different conclusions.

Pregnancy probability:

$$\Pr(P) = \begin{cases} 0 & \text{if } C=1 \text{ or } BCRx=1 \text{ or } S=0 \\ \pi & \text{if } C=0 \text{ and } BCRx=0 \text{ and } S=1 \end{cases}$$

STI probability:

$$\Pr(STI) = \begin{cases} 0 & \text{if } C=1 \text{ or } S=0 \\ f(STI\,prevalence) & \text{if } C=0 \text{ and } S=1 \end{cases}$$

where $f'(STI\,prevalence) > 0$.

3.2.3: Model of utility maximization

A rational individual faces a two-period utility maximization problem. [11]

$$\max_{\{S,BCRx,C,B,T\}} u_1(S,\ BCRx,\ C\ |\theta_{pref},\ \theta_{policy})\ +\ \beta E[u_2(B,\ T\ |\theta_{pref},\ \theta_{policy};\ P,\ STI)]$$

such that

$$P_{BCRx} * BCRx + P_c * C + Other\ Goods_1 \leq I_1 \qquad \text{(Period 1 Budget Constraint)}$$

and $P_T * T + P_B * B + Other\ Goods_2 \leq I_2$ (Period 2 Budget Constraint)

where

$$\frac{du_1(S,\ BCRx,\ C\ |\theta_{pref}, \theta_{policy})}{dC} < 0$$

$$\frac{du_1(S,\ BCRx,\ C\ |\theta_{pref}, \theta_{policy})}{dBCRx} = 0$$

A person makes choices about sex, prescription birth control, and condoms, which enter her first period utility and depend on her preferences and the current policy

[11] While there are other models of behavior such as the theory of the triune brain (MacLean, 1982), the rational choice model is simple, produces testable hypotheses, and is consistent with previous research showing that risky sex responds to costs such as disease risk (Arcidiacono et al., 2012; Gross et al., 2014; Chan et al., 2016).

environment. The probability of pregnancy and STI infection are based on choices in the first period. Pregnancy and STI status as well as preferences and the policy environment influence birth and STI treatment choice, which enter the second period utility function. Both sets of choices are subject to period-specific budget constraints, where prices are out-of-pocket prices. I assume that income is exogenous and that condoms decrease utility for at least some people and prescription contraception does not.

The utility-maximization problem can also be framed as two distinct optimization problems that are solved by backward-induction.

Stage 2:

$$\max_{\{A,\,B,\,T\}} \; u_2(B,\,T\,|\theta_{pref},\,\theta_{policy};P,\,STI)$$

$$\text{such that } \; P_T * T + P_B * B + Other\ Goods_2 \leq I_2$$

Stage 1:

$$\max_{\{S,\,RxBC,\,C\}} u_1(S,\;BCRx,\;C\,|\theta_{pref},\,\theta_{policy};B',\,T')$$

$$\text{such that } \; P_{BCRx} * BCRx + P_c * C + Other\ Goods_1 \leq I_1$$

First, the actor solves the second period problem by choosing what actions she would take for each pregnancy and STI infection state; she decides if she is willing to give birth and if she will get an STI treated. Then she solves the first period problem and decides what, if any, prevention she will engage in: abstention from sex, use of birth control, and use of condoms. The second-period choice functions are then treated as

parameters in the first period utility function; the decision-maker knows what she will do in the second period given any state and takes those future decisions as given in the first period.

3.2.4: Predictions of Theoretical Model

To derive predictions for the zero cost-sharing for contraception mandate, I determine how each parameter, choice, and outcome affects each other by considering the direction of the marginal effect holding all else constant.[12] First, I determine the effect of the policy on parameters in the utility function, such as the cost of birth control. Next, I consider the effects of parameters on choices, the effects of one choice on other choices, and the effects of outcomes on second stage choices. Last, I consider the effect of choices on outcomes.

Here I discuss only the clear theoretical predictions the model offers for the effect of lowering the cost of prescription contraception. The predicted effect of the zero cost-sharing mandate is a reduction in prevention as measured by condom sales and an increase in STIs, which would provide evidence of ex ante moral hazard. Appendix Tables A1-A5 depict the full set of model interactions for the zero cost-sharing mandate as well as the dependent coverage mandate, which is discussed in Section 6.

The requirement that insurance plans must cover prescription contraception with no out-of-pocket cost to the policyholder is a change in θ_{policy}. This policy lowers the cost and increases the quantity demanded of prescription birth control ($P_{BCRx}\downarrow \Rightarrow BCRx\uparrow$). The use of prescription birth control lowers the expected cost of sex without

[12] This is similar to a partial derivative, but since most parameters, choices, and outcomes are discrete as opposed to continuous, the functions are not differentiable.

a condom by lowering the probability of unintended pregnancy, and results in ex ante moral hazard in the form of more frequent risky sex ($BCRx\uparrow \Rightarrow \Pr(P)\downarrow \Rightarrow S\uparrow,\ C\downarrow$).[13] This increase in the quantity of risky sex demanded causes an increase in the rate of sexually transmitted infections ($S\uparrow,\ C\downarrow \Rightarrow STI\uparrow$). In summary, my model predicts that zero cost-sharing for prescription contraception increases prescription contraception and STI rates and decreases condom-use/sales and pregnancies.

4. Data and Method

4.1: Data

The data for this study come from several sources and are at the state-year level.[14] Each state's insured rate for 25- to 29-year-olds in 2011-12, which serves as the measure of treatment intensity, is derived from the Behavioral Risk Factor Surveillance System (BRFSS). Each year the BRFSS surveys over 400,000 adults and is representative at the state level (Centers for Disease Control and Prevention, 2013). The main benefit of BRFSS is that each state-year has sufficient sample size to precisely estimate the insured rate for 25- to 29-year-olds.

Condom sales for each state-year come from Nielsen Retail Scanner data, which contains sales information provided to Nielsen by retailers.[15] These data have

[13] Even though condoms protect against both STIs and pregnancy, two reasons people may switch are: 1) prescription contraception, particularly long-acting forms such as IUDs, are more effective (Centers for Disease Control and Prevention, 2017a), and 2) some people have a preference for sex without a condom (Ahituv et al., 1996).

[14] Ideally, a panel data set would contain insurance status, demographics, prevention (prescription contraception and condoms), fertility, and STIs. However, no individual-level data set contains the requisite data elements for this analysis.

[15] This is in contrast to Nielsen Consumer Panel Dataset (known as HomeScan), where consumers report purchases to Nielsen. Calculated (or Derived) based on data from The Nielsen Company (US), LLC and marketing databases provided by the Kilts Center for Marketing Data Center at The University of Chicago Booth School of Business. The conclusions drawn from the Nielsen data are those of the

important advantages over many surveys. First, since information is not self-reported, it does not suffer from reporting error, including social desirability bias. Second, Nielsen Retail Scanner data provide information on more condoms in a state-year than any survey. While these data do not cover 100 percent of sales, a large fraction of food, drug, and big-box stores' sales are covered. These data capture over 50% of sales at grocery and drug stores as well as about a third of mass merchandise stores from 35,000 locations (Kilts Center for Marketing, 2017). If changes in which stores are included are uncorrelated with treatment intensity, incomplete coverage will not bias estimated effects. While these data are the most appropriate source available on condoms for my analysis, there are two main limitations: (1) they contain sales of condoms instead of condom use and (2) sales to certain age groups cannot be isolated. However, no survey appropriate for longitudinal analysis or with sufficient sample size in each state-year contains information on condom use.

Less immediate outcomes, such as STI incidence and number of births, come from federal administrative data sources. Information on STIs including chlamydia and gonorrhea are available from the National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Preventions AtlasPlus (Centers for Disease Control and Prevention, 2017c). I focus on chlamydia and gonorrhea because these STIs are primarily found in heterosexuals who may respond to the cost of prescription contraception, while HIV and syphilis are concentrated in men who have sex with men (Centers for Disease Control and Prevention, 2016; Centers for Disease Control and Prevention, 2017b).

researchers and do not reflect the views of Nielsen. Nielsen is not responsible for, had no role in, and was not involved in analyzing and preparing the results reported herein.

State or local regulations require doctors, laboratories, and hospitals to report diagnosed cases of certain illness including STIs to local health departments, who then relay this information to the CDC (Centers for Disease Control and Prevention, 2015). National Vital Statistics provide counts of births in each state and year.

Each outcome is collapsed to the state-year level for 25- to 29-year-olds. While an earlier mandate in the ACA allows dependents to use their parents' insurance applies to young adults up to 26 (through 25) years of age, STI data are only available for pre-determined age groupings (20-24, 25-29, 30-34, etc.). Analyses are performed on 25- to 29-year-olds to isolate the effect of the zero cost-sharing mandate from the earlier policy, though 25-year-olds may contaminate the analysis somewhat. Event study analysis and robustness tests on an older group (30- to 34-year-olds) provide additional evidence that this data limitation is not driving results.

I control for a set of time-varying state-level characteristics: the unemployment rate (total and age-specific) and population (total and age-specific) provided by the Bureau of Labor Statistics;[16] income per capita data from the Bureau of Economic Analysis; a binary measure of strict abortion regulation based on information from the Guttmacher Institute;[17] and state mandates of adult dependent health insurance coverage and of required coverage of prescription contraception from Collins and

[16] "Age-specific" indicates that analyses on 25- to 29-year-olds include controls for the population and unemployment rate of 25- to 29-year-olds.

[17] Data were requested from the Guttmacher Institute. States are assigned to one of four categories – supportive, middle ground, hostile, extremely hostile – based on the number of major abortion restrictions in place during a year. For three examples, see www.guttmacher.org/sites/default/files/images/2000-2014-maps-states.png. I dichotomized categories into hostile (hostile or extremely hostile) or not (supportive or middle ground). Data were unavailable for 2007 and 2009. For the very few states that switched from not hostile to hostile between 2006 and 2008 or 2008 and 2010, I assigned hostile; otherwise 2007 and 2009 values were set to the values of the neighboring years.

Nicholson (2010) and Raissian and Lopoo (2015).

The years of analysis are 2006 to 2014. I start the analysis in 2006 because emergency contraception became available over-the-counter for adults starting in that year (National Conference of State Legislatures, 2012). Over-the-counter emergency contraception could interact with health insurance (the measure of treatment intensity in this study) in important ways. Over-the-counter emergency contraception eliminated the need to interact with a health provider, which was a greater burden to women without insurance. To isolate my analysis from the effect of emergency contraception, I exclude years before the introduction of over-the-counter emergency contraception.

4.2: Method

Identifying the effect of national policies can be difficult, because all states simultaneously experience the policy shock. To identify the effect of the zero cost-sharing mandate, I use a continuous measure of treatment intensity based on the pre-mandate insurance level, specifically the insured rate for 25- to 29-year-olds in 2011-12. By focusing on 25- to 29-year-olds, I am less likely to conflate the estimated effect of the zero cost-sharing mandate with the earlier dependent coverage policy.

The rationale for the treatment measure is that the potential effect of the zero cost-sharing mandate varies with the percent of the population that is insured. The effect of the zero cost-sharing mandate on behavior is stronger in states with high rates of insurance, because the mandate only applies to people who are insured. Consider the extreme cases: a hypothetical state with no insured 25- to 29-year-olds in 2011-12 would have no potential for an exogenous change in the cost of prescription birth

control, while a state where every person is insured would have the potential for a large exogenous change in the cost and use of prescription contraception.

To determine the effect of the zero cost-sharing mandate, I perform non-parametric and parametric event study analyses. The strengths of non-parametric event studies are that they reveal all changes that occur in the event year as well as providing a compelling visual representation. However, by adding a linear parametric assumption I can derive causal estimates with meaningful interpretations, perform statistical inference, and gain statistical power. Additionally, by clearly stating the parametric method's assumptions, I create transparency about when the method is appropriate.

The zero cost-sharing mandate could cause an immediate effect as well as a time-varying effect, both of which are important to capture. The time-varying effect could be due to more people learning about the policy over time. Another potential reason for a time-varying effect is the compounding effect of STI infection: when one case is transmitted, that infection now has the potential to be spread to future partners. See Figure 2 for a stylized event study with a visual representation of the parametric analysis. To capture both the immediate and time-varying effects in the parametric event study analysis, I fit a line for the pre-period represented by the solid blue line before 2012 in Figure 2. I then estimate both a one-time jump/drop that occurs in the year of policy implantation as well as any change in slope.[18]

[18] This is similar to post-estimation in Finkelstein (2007). After estimating event studies, she compares the difference in the event studies between 1970 (five years after Medicare introduction) and 1965 (the year of Medicare introduction) to the difference between 1965 and 1960 (five years before Medicare introduction). This approximately compares the slope in the event study before the policy to the slope in

First, I estimate non-parametric event study models. The estimating equation is

$$\log(Y_{st}) = \beta_0 + \sum_{\substack{t=2006 \\ \neq 2011}}^{2014} \beta_t \left(InsureRate_s * \mathbf{1}(Year_t)\right) + \alpha_j * \mathbf{1}(\text{State}_s) + \delta_t * \mathbf{1}(Year_t) + \beta_x * \boldsymbol{X}_{st} + \epsilon_{st}. \quad \text{(Eq. 1)}$$

By including state and year fixed effects, I control for time-invariant state characteristics and national year-specific changes. The insured rate for 25- to 29-year-olds in 2011-12 is $InsureRate_s$. The relationship between treatment intensity and outcomes in year *t* is β_t, and these coefficients show the pattern in the outcome between states with high and low uninsured rates. Since equation (1) is a log-linear regression and $InsureRate_s$ is a rate between 0 and 1, for a percentage point increase in the treatment intensity (insured rate) there is a β_t percent increase in the outcome in year *t* compared to the base year of 2011.

I also control for a set of time-varying state-specific covariates, $\boldsymbol{X}_{st}$. Controls include the unemployment rate (total and age-specific), population (total and age-specific), income per capita, a dummy for strict regulation of abortion, and dummy variables for state-level mandates similar to the ACA's dependent coverage mandate and mandates of contraception coverage.

In addition to revealing any change that occurs in the treatment year, the non-parametric event studies provide information on the identifying assumptions of the parametric event studies. The two identifying assumptions are: (1) pre- and post-trends

the event study after the study. For an example of a similar parameterization, see Levy et al. (2016) and Wolfers (2006) for a partial parameterization.

are approximately linear, and (2) there are no contemporaneous shocks. The non-parametric event study graphs show that the β_t's progress linearly. The second assumption that there are no contemporaneous changes related to states' pre-policy insurance levels cannot be tested directly. However, I perform falsification tests by estimating the effect of the zero cost-sharing mandate on state characteristics that should be unaffected by this policy change.

The research design for statistical inference is a parametric event study analysis estimating both a one-time immediate effect and a linear time-varying treatment effect using a continuous measure of treatment intensity based on the pre-policy insured rate. To identify the effect of zero cost-sharing, I estimate models of the form:

$$\begin{aligned}\log(Y_{st}) = {} & \beta_0 + \beta_1(InsureRate_s * (t - 2012)) \\ & + \beta_2\big(InsureRate_s * \mathbf{1}(t \geq 2012)\big) \\ & + \beta_3\big(InsureRate_s * (t - 2012) * \mathbf{1}(t \geq 2012)\big) \\ & + \alpha_s * \mathbf{1}(State_s) + \delta_t * \mathbf{1}(Year_t) + \gamma_x * \boldsymbol{X}_{st} + \epsilon_{st} \ . \end{aligned} \qquad \text{(Eq. 2)}$$

β_2 and β_3 are the parameters of interest and represent the estimated effect of the zero cost-sharing mandate on the intercept and the slope, respectively. By including state and year fixed effects, I control for cross-sectional, non-time-varying differences in outcomes; any remaining variation in outcomes is attributed to policy variation.

A pre-period trend is modeled by β_0 and β_1, while β_2 and β_3 represent the deviation from that trend in the post period. The binary variable $\mathbf{1}(t \geq 2012)$ indicates whether the mandate is in effect. Since $InsureRate_s$ is a rate between 0 and 1, for a one percentage point increase in the insured rate, there is a one-time change in the outcome of β_2 percent and an annual increase of β_3 percent. The model includes the

same vector of time-varying state-specific covariates as in Eq. 1, $\boldsymbol{X}_{st}$. The estimated β_2 and β_3 apply to the years in the analysis timeframe but may not persist indefinitely, especially as other aspects of the ACA went into effect.

There are important similarities and distinctions between my parametric event study approach and the traditional difference-in-differences method. Both use different levels of exposure to a treatment (either binary or in my case dose-response) to compare changes before and after policy implementation. In fact, the difference-in-differences method is a special case of my parametric event study analysis with flat pre-trends and no time-varying treatment effects. However, in addition to estimating a one-time change, my approach controls for a linear pre-treatment trend and estimates a linear time-varying treatment effect relative to this trend. This allows for a relaxed identification assumption: instead of requiring groups with different levels of treatment intensity to have the same trend in the absence of the policy, they need only continue on their relative pre-treatment trends. In a stylized event study found in Figure 2, this can be visualized as assuming the solid line on the left before policy implementation would remain on the same path as the lower dotted line if there had been no policy shocks. The deviation from the lower dotted line is the causal effect of the policy.

As with difference-in-differences, the main identifying assumption cannot be directly verified in the parametric event study analysis used in this study. Since the counterfactual is unobservable, I must rely on an ocular test: the pre-trends should be

approximately linear.[19] Linearity is important because higher order functions will have different slopes across the domain, and in those cases, the parametric event study method I am using could find spurious effects. This is similar to the concern with non-parallel pre-trends in the traditional difference-in-differences. Additionally, I must assume there are no contemporaneous shocks that are correlated with the pre-policy insured rate and the outcomes. If both assumptions are met, then any observed changes in intercept or slope in the event studies are due to the policy.

All standard errors are clustered at the state level, and regressions are weighted by the 2011 age-specific population. By weighting according to population, estimates reflect the national average treatment effect.

5. Results

5.1: Summary Statistics

Table 1 contains the mean and standard deviation of measures of treatment intensity and outcome variables, weighted by age-specific state populations. In addition to summary statistics for 25- to 29-year-olds, I also present information on 20- to 24-year-olds who are the age group analyzed for the dependent coverage mandate in Section 6 and 30- to 34-year-olds who are analyzed in a robustness check in Section 7. Treatment measures are in the first panel. About two-thirds of those age 25 to 29 had insurance in 2011-12. Importantly, there is substantial variation across states, with standard deviations in insured rates around 7 percentage points. Figure 3

[19] Though it might be tempting to extend this parametric approach to higher degrees, an ocular test could be difficult to implement with quadratic or higher order functions, especially in terms of detecting causal effects. In such a context, the method used in Wolfers (2006) might be more appropriate.

shows the geographic distribution of the treatment intensities. Geographic variation in insured rates is dispersed. As expected, there are fewer insured people in the South, which has lower average education and income, while there are more insured people in the Northeast and upper Midwest. Summary statistics for outcome variables are shown in the second panel of Table 1. Chlamydia is a very common STI. It is primarily found in women and is often contracted through heterosexual intercourse (Centers for Disease Control and Prevention, 2016), so should be responsive to policies that affect birth control use. Condom sales and births are more common than the diseases examined here.

5.2: Effects on Prevention Investment (Condom Sales)

Figure 4 shows the non-parametric event study analysis for condom sales. Since condom data are not available by age group, I start the analysis for condoms in 2010 to prevent contamination from changes caused by the 2010 dependent coverage mandate.[20] Note that large fractions of people in this age range are condom users. Over a quarter of 25- to 29-year-olds use condoms (Reece et al., 2010), which is consistent with this policy causing a noticeable effect on total condom sales. The figure is consistent with ex ante moral hazard causing a reduction in prevention due to the zero cost-sharing mandate. Comparing before and after 2012 in Figure 4, the trend for log condoms is flat before the policy and drops starting in 2012; the effect of zero cost-sharing is a relative decrease in the purchases of condoms in states that were more treated or had higher insured rates.

[20] While there are only two pre-period years for condom sales, Nielsen data allows for analysis as the quarter level. Event studies at the quarter level are consistent with annual results but are noisier. Results available upon request.

Table 2 presents estimates for the zero cost-sharing mandate on condom sales. The pre-period is 2010-11, and the post-period is 2012-14. The effect of the zero cost-sharing mandate was a 0.092 percent annual decrease in condom sales after 2012 for each percentage point increase in the 2008-09 insured rate. Since the effects are mainly changes in slopes, the impact compounded over time; the zero cost-sharing mandate causes an approximate 0.09 percent decrease in condom sales the first year after policy implementation for a percentage point increase in treatment intensity, and a 0.18 percent decrease in the second year. The point estimate on the change in intercept is the opposite sign as the change in slope, but it is small and not statistically significant. A standard deviation increase in treatment intensity would result in about 9 million fewer condoms sold in the third year after implementation of the zero cost-sharing mandate. The fact that condom sales decreased in response to health insurance coverage is evidence of ex ante moral hazard and is consistent with the theoretical predictions presented in Section 3.

5.3: Effects on Births and STIs

Non-parametric event studies for births and STIs are shown in Figure 5 to Figure 7, and results from the parametric event study analysis are presented in Table 3. Importantly for statistical inference, the figures are consistent with assumption that pre- and post-trends are approximately linear. Figure 5 shows the non-parametric event study analysis for births. There is a sharp decrease in births starting in 2012, which likely reflects increased use of prescription contraception and a decrease in unintended pregnancy. From the parametric event study results in Table 3, the zero cost-sharing mandate did have a statistically significant effect on births. Since the

intensity of treatment is a proportion ranging from 0 to 1 and the outcome is a log, the estimate should be interpreted as a 1/10 of percent annual decrease in births for each percentage point increase in the insured rate. If the pre-policy insured rate among 25- to 29-year-olds was one standard deviation higher, I predict there would have been 58,000 fewer births the third year after the zero cost-sharing mandate went into effect.

Figure 6 and Figure 7 show the non-parametric event studies for chlamydia and gonorrhea, respectively. Again, the figures are consistent with the assumption that pre- and post-trends are approximately linear. The effect of zero cost-sharing on STIs is an important indication of ex ante moral hazard. The non-parametric event study for chlamydia in Figure 6 reverses a monotonic downward trend, with the event study reaching the bottom of a valley the year before policy implementation. This is consistent with my hypotheses and an ex ante moral hazard effect; treatable STI incidence increased more in higher treated states beginning in 2012. The effect of the zero cost-sharing mandate on gonorrhea shown in Figure 7 is less conclusive. However, the pre-trend is fairly flat, and the figure does show a meaningful jump after policy implementation.

The results from a parametric event study analysis of STIs are also in Table 3. The zero cost-sharing mandate caused an increase in chlamydia. A percentage point increase in treatment intensity for 25- to 29-year-olds caused a one-time 0.530% increase and a yearly increase (change in slope) of 0.248% in chlamydia incidence. This implies a standard deviation increase in the insured rate for 25- to 29-year-olds would result in 45,000 more cases of chlamydia three years after the zero cost-sharing mandate. The fact that the zero cost-sharing mandate increased chlamydia reflects the

decrease in prevention as measured by condom sales and the effect of ex ante moral hazard. The estimated effect of the zero cost-sharing mandate on gonorrhea was a large but not statistically significant one-time jump and very little change in slope.

In summary, both the non-parametric and parametric event studies show evidence that the zero cost-sharing mandate caused ex ante moral hazard; investment in prevention as measured by condom sales decreased due to this policy. Since this policy does not offer any countervailing protection against STIs, ex ante moral hazard resulted in increased cases of chlamydia. However, increased access to prescription contraception reduced total births, likely due to fewer unintended pregnancies.

6. Extensive Margin of Health Insurance: Dependent Coverage Mandate

While the zero cost-sharing mandate affected the intensive margin of health insurance, it is important to determine whether an increase in STIs is a feature of other insurance expansions or if comprehensive coverage protects against the spread of disease. To investigate this question I exploit the young adult dependent coverage mandate of 2010. This policy caused an exogenous shock on the extensive margin – the number of people insured – by allowing young adults to join their parents' health insurance.

The dependent coverage mandate required that, starting in September 2010, all insurance plans covering dependents of the primary policyholder must offer coverage to children of the policyholder up to age 26 (Department of Labor, 2017). Before implementation of the dependent coverage mandate, close to 14 million people in their 20s were uninsured (Collins and Nicholson, 2010). The dependent coverage mandate

had an economically meaningful and statistically significant effect on the insured rate for young adults. Appendix Figure A1 shows the pattern of uninsured rates for 18- to 24-year-olds and 25- to 34-year-olds. Before 2010, 18- to 24-year-olds consistently had higher uninsured rates, but experienced a sharp decrease in their uninsured rate starting in 2010. Sommers et al. (2013) estimate that this mandate increased the percent of adults under the age of 26 who are insured by 6.7 percentage points.[21]

While changes at the extensive margin of health insurance should also cause a reduction in prevention, the net impacts on STIs and pregnancy are more ambiguous due to countervailing effects of insurance. One source of ambiguity from the dependent coverage mandate is due to the increased probability that a potential sexual partner has insurance, permitting quick and effective treatment of STIs. If a sexual partner is STI-free, then reductions in prevention in the form of sex without a condom will not result in infection transmission. In the notation of the economic model, if $STI\,prevalence$ decreases and $f'(STI\,prevalence) > 0$, then Pr(*STI)* decreases. Additionally, insurance lowers the cost of having a baby, so the dependent coverage mandate may result in reduction in condom-use and an increase in intended pregnancies among people who would like but could not afford the medical expenses associated with a pregnancy ($P_B\downarrow \Rightarrow C\downarrow \Rightarrow B\uparrow$). So the dependent coverage mandate may cause an increase in intended births but a decrease in unintended births, with an ambiguous effect on net births.

[21] This is a large change compared to other recent policies aimed at increasing insurance rates. For instance, the State Children's Health Insurance Program (SCHIP), which offers public health insurance to low-income but Medicaid-ineligible children, increased coverage by 5.7 percentage points in the target population. However, the net effect on childhood insurance rates was much smaller because of strict income eligibility criteria (LoSasso and Buchmueller, 2004).

I use the same empirical strategy as in my main analysis to examine the effect of the 2010 dependent coverage mandate, but now the treatment intensity is the percent of 20- to 24-year-olds who are uninsured in 2008-09. The intuition is that the more uninsured young adults in a state, the larger the potential increase in the insured rate from this policy. Again, the data are collapsed to the state-year level for 20- to 24-year-olds. While 25-year-olds are eligible to use their parents' insurance due to this policy, most of the data on the outcomes I examine are only available in five-year age groupings (20-24, 25-29), so I focus on 20- to 24-year-olds.

If my empirical strategy shows that the dependent coverage mandate affects outcomes through insurance coverage, then the mandate must increase coverage more in states with lower pre-mandate coverage. To test this hypothesis, I regress the change in insured rate (2011-12 rate minus 2008-09 rate) on the 2008-09 uninsured rate for ages exposed to the dependent coverage mandate using the following model:

$$\Delta \ InsureRate_s = \beta_0 + \beta_{change} UninsureRate0809_s \ + \ \epsilon_{st}$$

Estimates from this model are in Appendix Table A6. Importantly, the effect is large and statistically significant for young adults exposed to the policy. The dependent coverage mandate reduced the uninsured rate for young adults by 4.3 percentage points or about 6 percent.[22] This is comparable to the 6.7 percentage point effect in Sommers et al. (2013). There are several reasons why not all uninsured young adults gain coverage from the dependent coverage mandate, such as uninsured parents or

[22] From Appendix Table A6, 0.139*(1-0.688)=0.043

unwillingness for parents to add a child to their plan. As a falsification test, I conduct the same analysis for older groups who should not be affected by the dependent coverage mandate. Results for older groups are also in Appendix Table A6 and show that the dependent coverage mandate did not affect insurance coverage for these groups; estimates are closer to zero and not statistically significant. The percent change for young adults is at least twice the magnitude as for other age groups.

I conduct similar non-parametric and parametric event study analyses as for the zero cost-sharing mandate. However, there are two changes worth noting: 1) the measure of treatment intensity is now the pre-policy insured rate for 20- to 24-year-olds, and 2) the years of analysis are 2006 to 2012, to isolate the effect of the dependent coverage mandate from the effect of the zero cost-sharing mandate. The omitted year in the non-parametric event studies is 2009.

Effects of the dependent coverage mandate on condom sales are in Appendix Figure A2 and Table 4. Comparing 2006-2009 and 2010-2012 in Appendix Figure A2, condom sales trend upward but flatten out starting in 2010. This indicates that states with high uninsured levels purchased more condoms over time compared to low uninsured states before 2010. In 2010 when high uninsured states experienced larger increases in insurance rates, the relative gains in condom sales stopped. Consistent with the non-parametric event study analysis, the dependent coverage mandate caused a statistically significant negative change in the slope, as shown in Table 4. The effect of the dependent coverage mandate of 2010 was a 0.129 percent annual decrease in condom sales after 2010 for each percentage point increase in the 2008-09 uninsured rate. A standard deviation increase in treatment intensity would result in about 15

million fewer condoms sold in the third year after implementation of the dependent coverage mandate. Like the zero cost-sharing mandate, the dependent coverage mandate caused an ex ante moral hazard effect by reducing investment in prevention.

Estimates for the effect on births and STIs are presented in Table 5 and Appendix Figure A3-A5. The non-parametric event study for births in Appendix Figure A3 and the point estimates in Table 5 indicate the dependent coverage mandate had no net effect on fertility. Since the theoretical model provides ambiguous predictions of the effect of the dependent coverage mandate on STIs, the non-parametric event studies give an impression of which effect is stronger, the protective effect of health care or the reduction in prevention caused by ex ante moral hazard. The estimates for the effect of the dependent coverage mandate indicate that the protective effect of insurance more than makes-up for the reduction in prevention. Appendix Figure A4 and Appendix Figure A5 show the non-parametric event studies for the effect on chlamydia and gonorrhea, respectively. The graph for the effect of the dependent coverage mandate on chlamydia cases in Appendix Figure A4 has an increasing trend before the policy goes into effect, peaks the year of policy implementation, and the trend reverses in the post period. High-uninsured states had increasing incidence of this STI compared to low uninsured states in the pre-period, but this trend reversed in the post-period. As in the graph for the effect on chlamydia, Appendix Figure A5 shows an upward trend in gonorrhea that peaks the year before the policy goes into effect and decreases in the post period.[23]

[23] Like the zero cost-sharing mandate, the dependent coverage mandate is a national policy, which addresses concerns about policy timing endogeneity or regression to the mean driving results.

The results from the parametric event study analysis in Table 5 confirm the results from the non-parametric event study. The dependent coverage mandate caused a downward change in slope for log chlamydia; a percentage point increase in the treatment intensity caused a yearly decrease of 0.256% in chlamydia incidence. An increase of one standard deviation in the uninsured rate for 20- to 24-year-olds would result in almost 87,000 fewer chlamydia cases in the third year after the implementation of the dependent coverage mandate. Likewise, for the effect on gonorrhea, the parametric event study analysis shows a downward change in slope after policy implementation; a percentage point increase in the treatment intensity caused a yearly decrease of 0.410% in gonorrhea cases. A standard deviation increase in treatment intensity would prevent 30,000 gonorrhea cases in the third year after the dependent coverage mandate among 20- to 24-year-olds.

The dependent coverage mandate caused ex ante moral hazard and countervailing effects. This policy resulted in lower investment in prevention as measured by condom sales. However, because the dependent coverage mandate caused a shock to the extensive margin of health insurance, other aspects of insurance protected against STI infection. The net effect on STIs was a reduction in illness. While both the zero cost-sharing and dependent coverage mandate caused ex ante moral hazard, an increase in STIs is not endemic to all insurance expansions.

7. Robustness and Falsification Tests

One possible threat to identification is that health insurance changes how often people interact with healthcare providers, including the frequency of STI testing. Such

a response would change the number of STI diagnoses even if risky behavior remained the same. Appendix Figure A6 and Appendix Figure A7 show non-parametric event studies for the effect of the mandates on routine medical services. Both figures are inconsistent with changes in interactions with health professionals, and thus testing, driving the results for STIs. In fact the direction of the change in this outcome is the opposite as the direction for STIs. One potential factor leading to less frequent contact with doctors after the zero cost-sharing mandate is that women gain access to free long-acting contraception from this policy that can last multiple years (e.g. IUDs), and so they may skip annual wellness visits.

In terms of actual measures of testing, the National Ambulatory Medical Care Survey, which samples doctors' offices and visits, contains information on chlamydia testing. Since this data source is not designed for state-level analysis, I provide suggestive evidence for the dependent coverage mandate based on a comparison between 20- to 25-year-olds (treated group) and 26- 30-year-olds (control group). Appendix Figure A8 shows that both groups generally trend together through the whole period.

I also conduct robustness checks with unweighted regressions, Poisson regressions, excluding early Medicaid expansion states, and on an older group (30- to 34-year-olds). Generally, the results are robust to different specifications, with similar magnitudes and direction. Results for unweighted regressions in Appendix Table A7 and Appendix Table A8 are similar; however, results for the dependent coverage mandate are smaller, and fewer estimates are statistically significant for both mandates. Smaller effects for the unweighted models indicate that more populous

states are more responsive. Though most of the expansion of Medicaid to childless adults occurs after both mandates, some states expanded coverage early. I check if estimates are robust to excluding early expansion states and present these results in Appendix Table A9 and Appendix Table A10. Results are largely robust to excluding states that expanded Medicaid coverage to childless adults early.

In Appendix Table A11, I show the estimated effect of the zero cost-sharing mandate on an older group, 30- to 34-year-olds. This analysis is an important robustness test of the zero cost-sharing mandate, because 25-year-olds are targeted by the 2010 dependent coverage mandate but included in the zero cost-sharing mandate analysis due to data limitations. In addition, many people who are in their early 20s in 2010 age into the 25- to 29-year-old group before 2014. The 30- to 34-year-olds sample does not suffer from either of these concerns. Additionally, it provides insight into heterogeneous effects by age. Comparing the main results for 25- to 29-year-olds in Table 3 to 30- to 34-year-olds in Appendix Table A11 effects for both age groups are similar. The effect on chlamydia is approximately the same and highly significant, while the estimate for fertility is smaller and no longer significant.[24]

Falsification tests of whether the mandates impact other state-level characteristics are reported in Appendix Tables A12-A13. The only coefficient that is statistically significant is the change in slope for the 2012 zero cost-sharing policy on percent of the state that is women. However, the coefficient is quite small, and we would expect

[24] One reason for the difference in fertility response between 25- to 29-year-olds and 30- to 34-year-olds is that during these age ranges, probability of pregnancy from unprotected sex declines (Dunson et al., 2004). Therefore, even though both groups appear to have similar increase in risky sex based on STIs, fertility of the older group is less responsive to this change.

to incorrectly reject the null hypothesis at least once due to type I error. Importantly, HIV and syphilis are not statistically significant, because these two diseases are concentrated in men who have sex with men (Centers for Disease Control and Prevention, 2016; Centers for Disease Control and Prevention, 2017b) and should be less responsive to the mandates, particularly related to female contraception. Another important falsification test is that the non-parametric event studies for the 25- to 29-year-olds and 30- to 34-year-olds[25] trend smoothly through 2010. This indicates that the dependent coverage mandate, which should only affect people under 25, did not affect the older group.

8. Conclusion

This study contributes to the literature by testing for ex ante moral hazard with respect to risky sex. Increased risky sex in response to lower expected costs is consistent with the previous literature on risky sex and the rational choice model of behavior. While previous empirical research generally finds mixed or weak evidence of ex ante moral hazard, there is reason to believe many forms of prevention are responsive to future, not current, insurance status.

I find that a standard deviation increase in treatment intensity results in several million fewer condoms purchased three years after insurance expansions, both at the intensive and extensive margins of insurance. When I isolate the effect of ex ante moral hazard using the zero cost-sharing mandate, a standard deviation increase in

[25] Event studies for 30- to 34-year-olds are available upon request and are similar to event studies for 25- to 29-year-olds.

pre-policy insured rate results in 45,000 more cases of chlamydia three years after the policy went into effect. However, the protective effect of insurance on STI transmission more than makes up for any negative effect of ex ante moral hazard; the dependent coverage mandate caused meaningful reductions in STI incidence.

While I cannot directly quantify the effect of these policies on net utility without making strong assumptions about utility functions, it seems likely both had positive net impacts. While the zero cost-sharing mandate did cause an increase in STIs, decreased unintended births are likely much more meaningful both in financial and non-monetary terms. The benefit of the dependent coverage mandate is more definitive: this mandate reduced STIs. Additionally, based on the reduction in fertility at the intensive margin, the null effect on birth from the dependent coverage mandate is likely due to an increase in intended fertility and a reduction in unintended pregnancies.

An important policy implication of my findings is that insurance has unintended consequences, but in some cases comprehensive insurance coverage can mitigate these problems. Similarly, since lowering the cost of prescription contraception causes substitution away from condoms, one way to prevent risky sex could be to subsidize condoms. Additionally, repeal of one or both policies in this study is a real possibility, and this analysis provides suggestive evidence on the effect of policy termination.

Future work that leverages changes in expectations about future insurance status could shed light on distortions in other health behaviors. Empirically testing the effect of subsidizing condoms on condom use and STI transmission could be important to determine if condom subsidies are a potential tool to counteract the unintended

consequences found in this study. While women generally receive information on STI risk and consequences at initiation of prescription contraception use, male partners may be less well informed; examining the impact of informing both partners about STI risk is another important question for future research.

REFERENCES

Ahituv, Avner, V. Joseph Hotz, and Tomas Philipson. "The responsiveness of the demand for condoms to the local prevalence of AIDS." *Journal of Human Resources* 31.4 (1996): 869-897.

Arcidiacono, Peter, Ahmed Khwaja, and Lijing Ouyang. "Habit persistence and teen sex: Could increased access to contraception have unintended consequences for teen pregnancies?" *Journal of Business & Economic Statistics* 30.2 (2012): 312-325.

Barbaresco, Silvia, Charles J. Courtemanche, and Yanling Qi. "Impacts of the Affordable Care Act dependent coverage provision on health-related outcomes of young adults." *Journal of Health Economics* 40 (2015): 54-68.

Bearak, Jonathan M., et al. "Changes in out-of-pocket costs for hormonal IUDs after implementation of the Affordable Care Act: an analysis of insurance benefit inquiries." *Contraception* 93.2 (2016): 139-144.

Buckles, Kasey S., and Daniel M. Hungerman. "The Incidental Fertility Effects of School Condom Distribution Programs." National Bureau of Economic Research w22322 (2016).

Carlin, Caroline S., Angela R. Fertig, and Bryan E. Dowd. "Affordable Care Act's Mandate Eliminating Contraceptive Cost Sharing Influenced Choices of Women with Employer Coverage." *Health Affairs* 35.9 (2016): 1608-1615.

Centers for Disease Control and Prevention. "Frequently Asked Questions about Changes to the Behavioral Risk Factor Surveillance System." (2013). www.cdc.gov/surveillancepractice/reports/brfss/brfss_faqs.html. Accessed 6 August 2015.

Centers for Disease Control and Prevention. "National Notifiable Diseases Surveillance System." (2015). www.cdc.gov/nndss/. Accessed 7 June 2017.

Centers for Disease Control and Prevention. "Sexually Transmitted Disease Surveillance 2015." (2016). Atlanta: U.S. Department of Health and Human Services.

Centers for Disease Control and Prevention. "Contraception." (2017a). www.cdc.gov/reproductivehealth/contraception/index.htm. Accessed 7 June 2017.

Centers for Disease Control and Prevention. "HIV Among Men in the United States." (2017b). www.cdc.gov/hiv/group/gender/men/index.html. Accessed 18 July 2017.

Centers for Disease Control and Prevention. "NCHHSTP AtlasPlus." (2017c). https://www.cdc.gov/nchhstp/atlas/index.htm. Accessed 3 March 2017.

Centers for Medicare & Medicaid Services “FAQs About Affordable Care Act Implementation (Part XXVI).” (2015). www.dol.gov/sites/default/files/ebsa/about-ebsa/our-activities/resource-center/faqs/aca-part-xxvi.pdf. Accessed 10 June 2017.

Chesson, Harrell W. "Sexonomics: a commentary and review of selected sexually transmitted disease studies in the economics literature." *Sexually Transmitted Diseases* 39.3 (2012): 161-166.

Chetty, Raj. "Moral hazard versus liquidity and optimal unemployment insurance." *Journal of Political Economy* 116.2 (2008): 173-234.

Collins, Sara R., and Jennifer L. Nicholson. "Rite of passage: Young adults and the Affordable Care Act of 2010." (2010). New York: The Commonwealth Fund.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. "Early impacts of the Affordable Care Act on health insurance coverage in Medicaid expansion and non-expansion states." *Journal of Policy Analysis and Management* 36.1 (2017): 178-210.

Cummins, David, and Sharon Tennyson. "Moral hazard in insurance claiming: evidence from automobile insurance." *Journal of Risk and Uncertainty* 12.1 (1996): 29-50.

Dave, Dhaval, and Robert Kaestner. "Health insurance and ex ante moral hazard: evidence from Medicare." *International Journal of Health Care Finance and Economics* 9.4 (2009): 367-390.

de Preux, Laure B. "Anticipatory ex ante moral hazard and the effect of Medicare on prevention." *Health Economics* 20.9 (2011): 1056-1072.

Department of Health and Human Services. "A Report of the Surgeon General: How Tobacco Smoke Causes Disease: What It Means to You.” (2010). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Department of Labor. "Young Adults and the Affordable Care Act: Protecting Young Adults and Eliminating Burdens on Businesses and Families FAQs." (2017). www.dol.gov/agencies/ebsa/about-ebsa/our-activities/resource-center/faqs/young-adult-and-aca. Accessed June 10 2017.

Doleac, Jennifer L., and Anita Mukherjee. "The Moral Hazard of Lifesaving Innovations: Naloxone Access, Opioid Abuse, and Crime." (2018). http://jenniferdoleac.com/wp-content/uploads/2018/03/Doleac_Mukherjee_Naloxone.pdf. Accessed April 3 2018.

Dunson, David B., Donna D. Baird, and Bernardo Colombo. "Increased infertility with age in men and women." *Obstetrics & Gynecology* 103.1 (2004): 51-56.

Ehrlich, Isaac, and Gary S. Becker. "Market insurance, self-insurance, and self-protection." *Journal of Political Economy* 80.4 (1972): 623-648.

Finkelstein, Amy. "The aggregate effects of health insurance: evidence from the introduction of Medicare." *The Quarterly Journal of Economics* 122.1(2007): 1-37.

Finkelstein, Amy. *Moral Hazard in Health Insurance*. Columbia University Press, New York (2015).

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, and Katherine Baicker. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127.3 (2012): 1057-1106.

Gross, Tal, Jeanne Lafortune, and Corinne Low. "What happens the morning after? The costs and benefits of expanding access to emergency contraception." *Journal of Policy Analysis and Management* 33.1 (2014): 70-93.

Health Resources & Service Administration. “Affordable Care Act Expands Prevention Coverage for Women’s Health and Well-Being.” (2017). www.hrsa.gov/womensguidelines. Accessed 10 June 2017.

Hersher, Rebecca. “Gonorrhea is becoming untreatable, U.N. health officials warn.” NPR.com, (2016). http://www.npr.org/sections/thetwo-way/2016/08/30/491969011/u-n-health-officials-warn-gonorrhea-is-becoming-untreatable. Accessed 6 June 2017.

Kaiser Family Foundation. “Minimum Contraceptive Coverage Requirements Clarified by HHS Guidance.” KFF.com, (2017). http://www.kff.org/womens-health-policy/issue-brief/minimum-contraceptive-coverage-requirements-clarified-by-hhs-guidance/. Accessed 6 June 2017.

Kaiser Family Foundation. “Summary of the Affordable Care Act.” KFF.com, (2013). http://www.kff.org/health-reform/fact-sheet/summary-of-the-affordable-care-act/. Accessed 6 August 2017.

Kaplan, Thomas S., and Robert Pear. "House Passes Measure to Repeal and Replace the Affordable Care Act." *The New York Times* (2017).

Kilts Center for Marketing. “Nielsen Datasets.” (2017). https://research.chicagobooth.edu/nielsen/datasets#simple2. Accessed 9 June 2017.

Klick, Jonathan, and Thomas Stratmann. "Abortion access and risky sex among teens: parental involvement laws and sexually transmitted diseases." *Journal of Law, Economics, and Organization* 24.1 (2008): 2-21.

Levine, Phillip B. "Parental involvement laws and fertility behavior." *Journal of Health Economics* 22.5 (2003): 861-878.

Levy, Helen, Thomas C. Buchmueller, and Sayeh Nikpay. "Health reform and retirement." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* (2016): gbw115.

LoSasso, Anthony T., and Thomas C. Buchmueller. "The effect of the state children's health insurance program on health insurance coverage." *Journal of Health Economics* 23 (2004), 1059–1082.

Lovenheim, Michael F., Randall Reback, and Leigh Wedenoja. "How Does Access to Health Care Affect Teen Fertility and High School Dropout Rates? Evidence from School-based Health Centers." National Bureau of Economic Research w22030, (2016).

MacLean, Paul D. "Evolution of the psychencephalon." *Zygon* 17.2 (1982): 187-211.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz. "Health insurance and the demand for medical care: evidence from a randomized experiment." *The American Economic Review* 77.3 (1987): 251-277.

National Conference of State Legislatures. "Emergency Contraception State Laws." (2012). http://www.ncsl.org/research/health/emergency-contraception-state-laws.aspx Accessed 9 June 2017.

Oberlander, Jonathan. "Long time coming: why health reform finally passed." *Health Affairs* 29.6 (2010): 1112-1116.

Pauly, Mark V. "The economics of moral hazard: comment." *The American Economic Review* 58.3 (1968): 531-537.

Raissian, Kerri M., and Leonard M. Lopoo. "Mandating Prescription Contraception Coverage: Effects on Contraception Consumption and Preventive Health Services." *Population Research and Policy Review* 34.4 (2015): 481-510.

Reece, Michael, et al. "Condom use rates in a national probability sample of males and females ages 14 to 94 in the United States." *The Journal of Sexual Medicine* 7.s5 (2010): 266-276.

Ressler, Rand W., Melissa S. Waters, and John Keith Watson. "Contributing factors to the spread of sexually transmitted diseases." *American Journal of Economics and Sociology* 65.4 (2006): 943-961.

Senate.gov. “Implementation Timeline.” (2010). https://www.dpc.senate.gov/healthreformbill/ healthbill65.pdf. Accessed 15 July 2017.

Shampanier, Kristina, Nina Mazar, and Dan Ariely. "Zero as a special price: The true value of free products." *Marketing Science* 26.6 (2007): 742-757.

Simmons, Adelle, and Laura Skopec. “47 Million Women Will Have Guaranteed Access to Women's Preventive Services with Zero Cost-sharing Under the Affordable Care Act.” United States, Department of Health and Human Services, 2012.

Simon, Kosali, Aparna Soni, and John Cawley. "The Impact of Health Insurance on Preventive Care and Health Behaviors: Evidence from the First Two Years of the ACA Medicaid Expansions." *Journal of Policy Analysis and Management* (2017).

Sommers, Benjamin D., Thomas Buchmueller, Sandra L. Decker, Colleen Carey, and Richard Kronick. "The Affordable Care Act has led to significant gains in health insurance and access to care for young adults." *Health Affairs* 32.1 (2013): 165-174.

Sonfield, Adam, Athena Tapales, Rachel K. Jones, and Lawrence B. Finer. "Impact of the federal contraceptive coverage guarantee on out-of-pocket payments for contraceptives: 2014 update." *Contraception* 91.1 (2015): 44-48.

Supreme Court of the United States. *Burwell, Secretary of Health and Human Services, et al, vs Hobby Lobby Stores, Inc, et al.* Docket No. 13-354 (2014). https://www.supremecourt.gov/Search.aspx?FileName=/docketfiles/13-354.htm. Accessed 21 August 2016.

Thomas, Adam, and Emily Monea. "The high cost of unintended pregnancy." *CCF Brief* 45.5 (2011): 2-7. https://www.brookings.edu/wp-content/uploads/2016/06/07_unintended_pregnancy_thomas_monea.pdf. Accessed 21 February 2018.

Wolf, Richard. "Trump administration reversing Obamacare's birth control mandate." *USA Today* (2017). https://www.usatoday.com/story/news/politics/2017/05/31/trump-administration-reversing-obamacares-birth-control-mandate/102346078/. Accessed 21 February 2018.

Wolfers, Justin. "Did unilateral divorce laws raise divorce rates? A reconciliation and new results." *The American Economic Review* 96.5 (2006): 1802-1820.

World Health Organization. "Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics." Geneva: World Health Organization (2017). http://www.who.int/medicines/publications/global-priority-list-antibiotic-resistant-bacteria/en/.
Accessed 21 February 2018.

Zweifel, Peter, and Willard G. Manning. "Moral Hazard and Consumer Incentives in Health Care." Chapter 8 in Joseph Newhouse and Anthony Culyer (eds.), The Handbook of Health Economics Vol. 1. Elsevier, New York (2000): 409-459.

Figures

Figure 1: Ex ante Moral Hazard (Effect of Insurance on Risky Behavior)

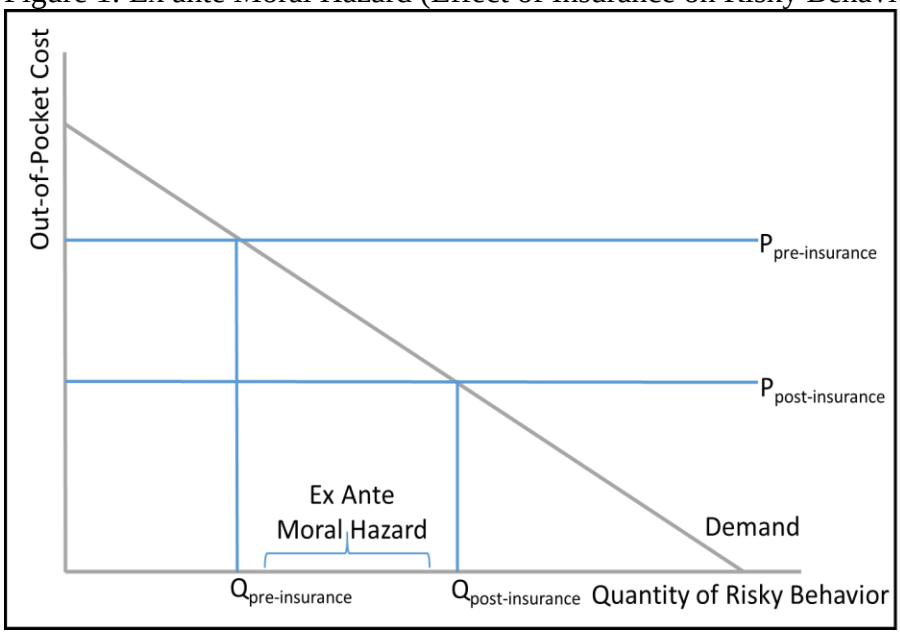
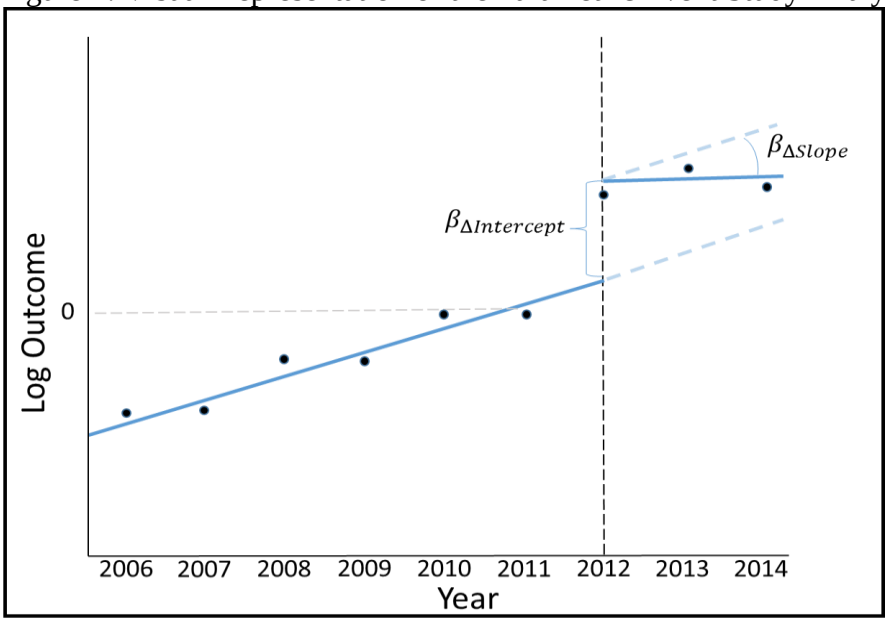


Figure 2: Visual Representation of the Parametric Event Study Analysis



Note: In this stylized event study, the continuous trends assumption can be visualized by assuming the solid line on the left (before policy implementation) would continue on the same path as the lower dotted line if there had been no policy shocks. The deviation from the lower dotted line is the causal effect of the policy.

Figure 3: Geographic Representation of Insured Rates, 25-29 Year Olds, 2011-12

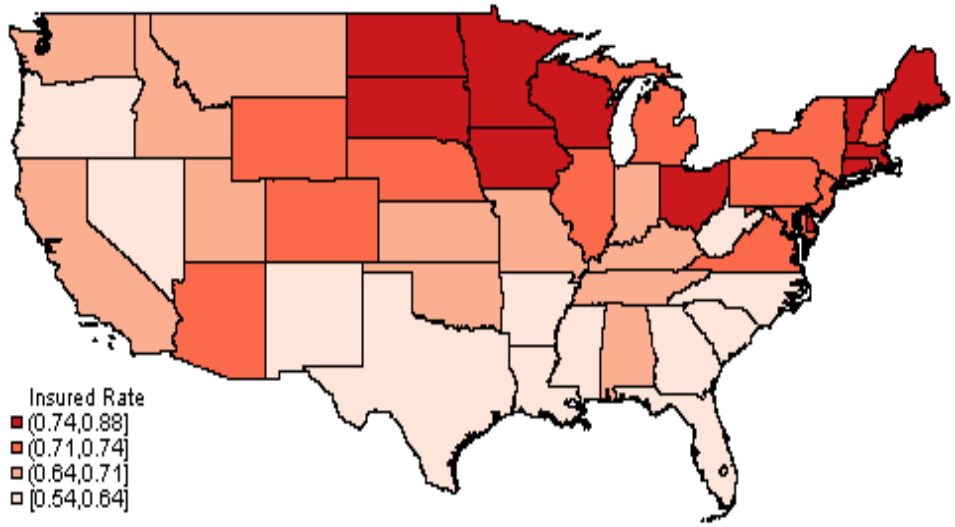
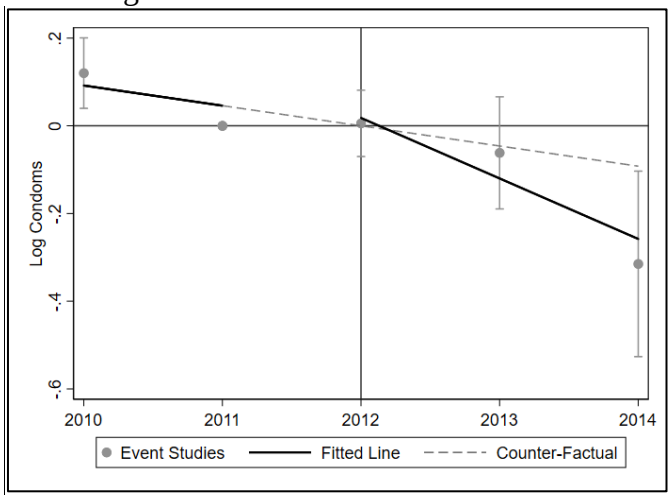
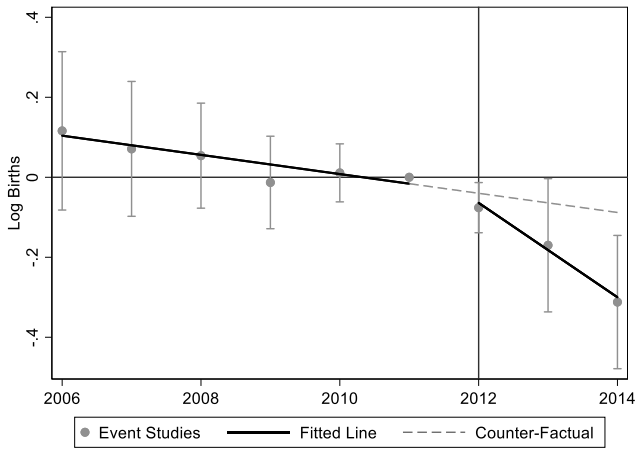


Figure 4: Zero Cost-Sharing Mandate –
Log Condoms



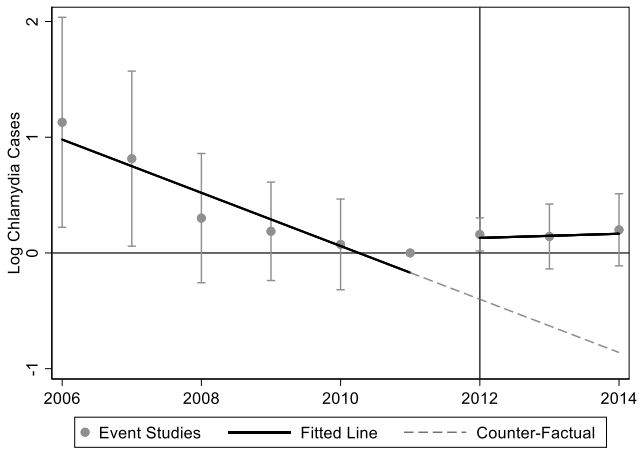
Note: Whiskers 95% confidence intervals
Treatment intensity: 2011-12 insured rate, 25-29 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total), population (total),
income per capita, strict state regulation of abortion,
and state mandates

Figure 5: Zero Cost-Sharing Mandate – Log Births



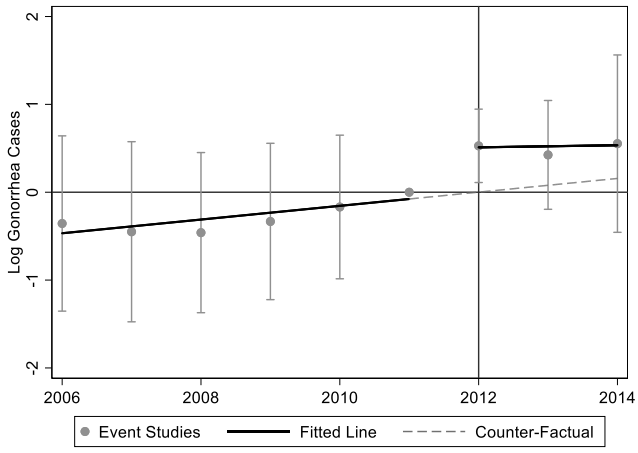
Note: Whiskers 95% confidence intervals
Treatment intensity: 2011-12 insured rate, 25-29 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure 6: Zero Cost-Sharing Mandate – Log Chlamydia Cases



Note: Whiskers 95% confidence intervals
Treatment intensity: 2011-12 insured rate, 25-29 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure 7: Zero Cost-Sharing Mandate – Log Gonorrhea Cases



Note: Whiskers 95% confidence intervals
Treatment intensity: 2011-12 insured rate, 25-29 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Tables

Table 1: Summary Statistics

Variable	20-24 Year Olds	25-29 Year Olds	30-34 Year Olds
	Treatment Intensity[26]		
Uninsured Rate	0.328	-	
(Avg. 2008-09; age-specific)	(0.083)	-	
Insured Rate	-	0.677	0.730
(Avg. 2011-12; age-specific)	-	(0.072)	(0.080)
	Outcomes		
Condoms	9,687,457	9,687,457	9,687,457
(Total)	(9,274,934)	(9,274,934)	(9,274,934)
Log Condoms	15.628	15.628	15.628
(Total)	(1.025)	(1.025)	(1.025)
Births	43.594	52,369	46,899
	(36,514)	(43,164)	(40,882)
Log Births	10.314	10.508	10.350
	(0.910)	(0.894)	(0.962)
Chlamydia Cases	22,056	9,755	4,257
	(18,093)	(8,807)	(4,106)
Log Chlamydia Cases	9.615	8.750	7.87
	(0.962)	(1.002)	(1.049)
Gonorrhea Cases	4,708	2,639	1,423
	(3,356)	(2,122)	(1,275)
Log Gonorrhea Cases	8.02	7.420	6.762
	(1.171)	(1.159)	(1.189)

Notes: Weighted by age-specific state populations
Standard deviation in parentheses (SD)

[26] Correlation between insured rates for 20- to 24-year-olds and 25- to 29-year-olds is 0.79.

Table 2: Effect of Zero Cost-Sharing Mandate on Log Condom Sales

	Log Condom Sales
Pre-Period Slope	-0.046
	(0.038)
Change in Intercept	0.018
	(0.052)
Change in Slope	-0.092*
	(0.052)
Predicted effect of SD increase in treatment in third year of policy	-8,997,710

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total), population (total), income per capita, strict state regulation of abortion, and state mandates
Years: 2010-2014
Predicted effect = Treatment intensity SD * (Avg. yearly occurrences) * ($\beta_{Intercept\Delta}$ + (3 * $\beta_{slope\Delta}$))

Table 3: Effect of Zero Cost-Sharing Mandate (25-29 Year Olds)

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Pre-Period Slope	-0.024	-0.230***	0.078
	(0.020)	(0.098)	(0.101)
Change in Intercept	-0.024	0.530***	0.510
	(0.043)	(0.171)	(0.345)
Change in Slope	-0.094**	0.248**	-0.066
	(0.037)	(0.122)	(0.206)
Predicted effect of SD increase in treatment in third year of policy	-57,971	44,772	2,963

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2010-2014
Predicted effect = Treatment intensity SD * (Avg. yearly occurrences) * ($\beta_{Intercept\Delta}$ + (3 * $\beta_{slope\Delta}$))

Table 4: Effect of Dependent Coverage Mandate on Log Condom Sales

	Log Condom Sales
Pre-Period Slope	0. 131
	(0.044)
Change in Intercept	0.031
	(0.053)
Change in Slope	-0.109***
	(0.037)
Predicted effect of SD increase in treatment in third year of policy	-12,347,025

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level,
weighted by 2011 state-age population
Controls: unemployment rate (total), population (total), income per capita,
strict state regulation of abortion, and state mandates
Years: 2006-2012
Predicted effect = Treatment intensity SD
* (Avg. yearly occurrences)
* ($\beta_{Intercept\Delta}$ + (3 * $\beta_{slope\Delta}$))

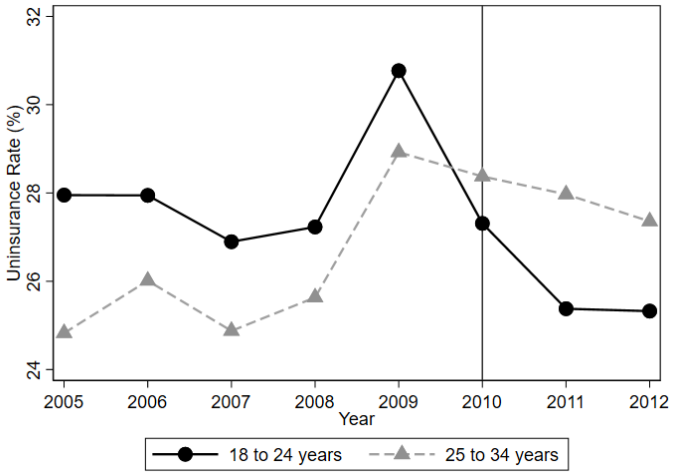
Table 5: Effect of Dependent Coverage Mandate (20-24 Year Olds)

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Pre-Period Slope	0.036	0.178	0.221
	(0.024)	(0.112)	(0.127)
Change in Intercept	-0.031	-0.224	-0.324
	(0.038)	(0.215)	(0.221)
Change in Slope	0.015	-0.256*	-0.410**
	(0.021)	(0.136)	(0.202)
Predicted effect of SD increase in treatment in third year of policy	2,674	-87,047	-29,634

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific),
income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2012
Predicted effect = Treatment intensity SD * (Avg. yearly occurrences) * ($\beta_{Intercept\Delta}$ + (3 * $\beta_{slope\Delta}$))

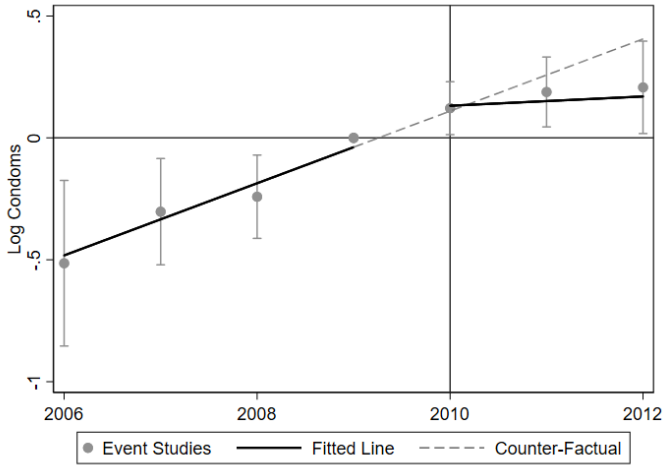
Appendix Figures

Figure A1: Trends in Uninsured Rate by Age Group



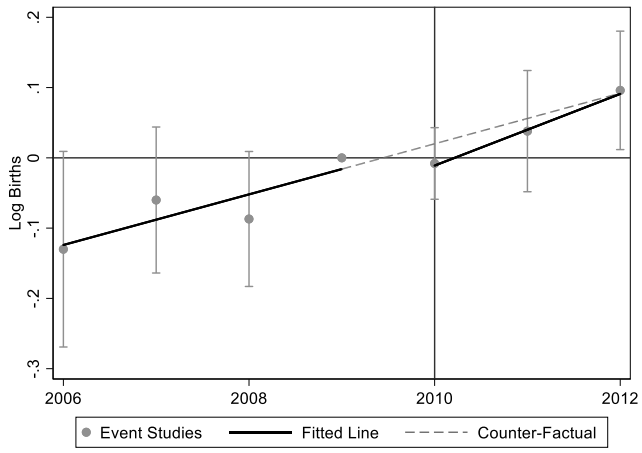
Data Source: Current Population Survey, Table HI01. Health Insurance Coverage Status and Type of Coverage by Selected Characteristics, various years

Figure A2: Dependent Coverage Mandate – Log Condoms



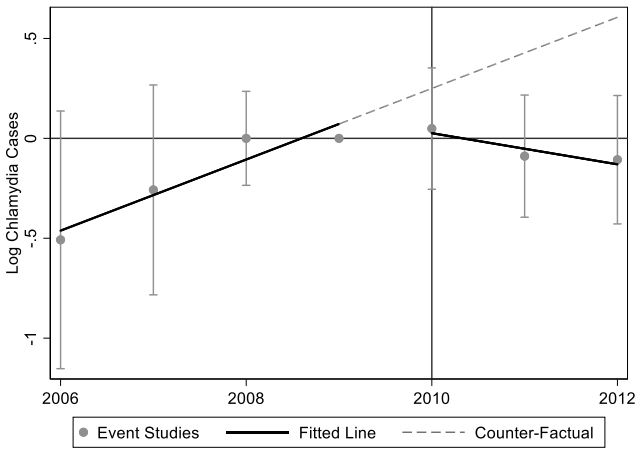
Note: Whiskers 95% confidence intervals
Treatment intensity: 2008-09 uninsured rate, 20-24 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total), population (total), income per capita, strict state regulation of abortion, and state mandates

Figure A3: Dependent Coverage Mandate – Log Births



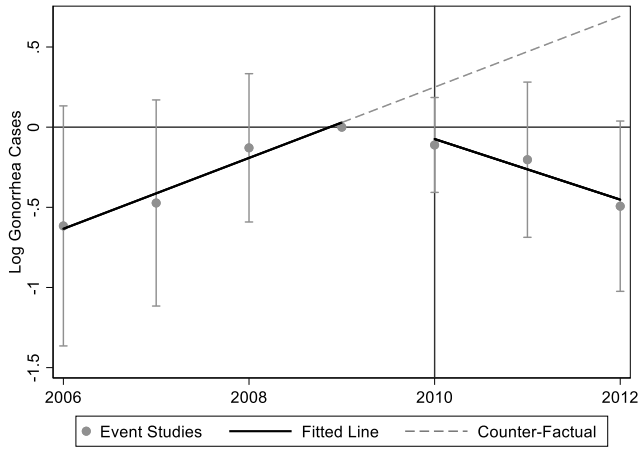
Note: Whiskers 95% confidence intervals
Treatment intensity: 2008-09 uninsured rate, 20-24 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure A4: Dependent Coverage Mandate – Log Chlamydia Cases



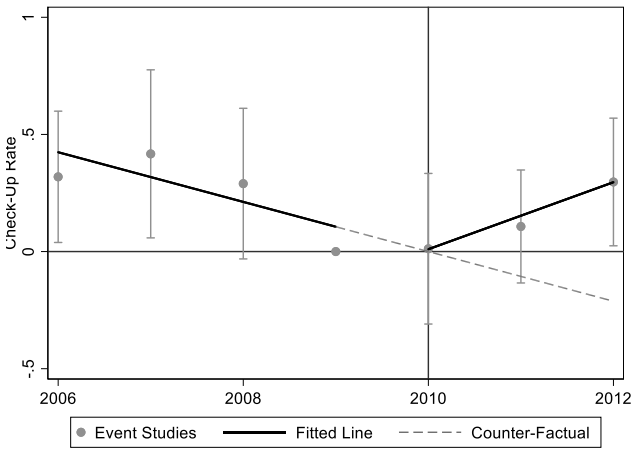
Note: Whiskers 95% confidence intervals
Treatment intensity: 2008-09 uninsured rate, 20-24 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure A5: Dependent Coverage Mandate – Log Gonorrhea Cases



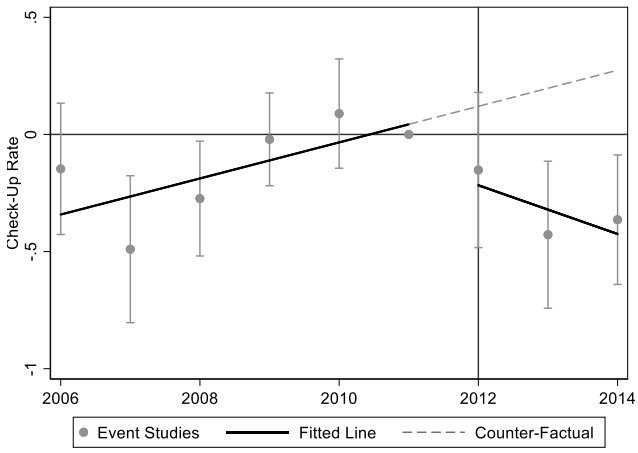
Note: Whiskers 95% confidence intervals
Treatment intensity: 2008-09 uninsured rate, 20-24 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure A6: Dependent Coverage Mandate – Check-Up Rate



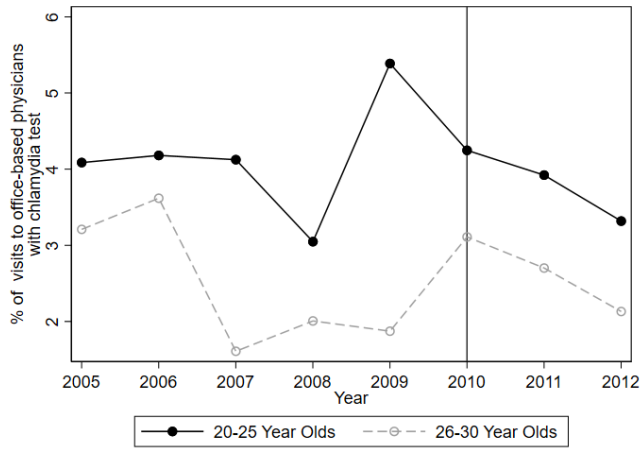
ote: Whiskers 95% confidence intervals
Treatment intensity: 2008-09 uninsured rate, 20-24 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure A7: Zero Cost-Sharing Mandate – Check-Up Rate



Note: Whiskers 95% confidence intervals
Treatment intensity: 2011-12 insured rate, 25-29 year olds
Cluster at state-level, weighted by 2011 state-age population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates

Figure A8: Chlamydia Testing by Year



Data source: National Ambulatory Medical Care Survey

Appendix Tables

Table A1: Effect of Policies on Model Parameters

	Dependent Coverage	Zero Cost-Sharing
Birth Costs	-	0
STI Treatment Costs	-	0
Birth Control Costs	0	-
STI Prevalence	+/-	+/-

Table A2: Effect of Model Parameters on Choices

	Birth Costs	STI Treatment Costs	Birth Control Costs	STI Prevalence
Birth Control	+	0	-	0
Sex	-	0	0	0
Condom	+	0	0	0
Birth	-	0	0	0
STI Treatment	0	-	0	0

Table A3: Effect of Choices on Choices

	Birth Control	Sex	Condom	Birth	STI Treatment
Birth Control	1	+	-	+	0
Sex	+	1	+	-	+
Condom	-	+	1	+	-
Birth	0	0	0	1	0
STI Treatment	0	0	0	0	1

Table A4: Effect of Outcomes on Choices

	Pregnancy	STI
Birth Control	0	0
Sex	0	0
Condom	0	0
Birth	+	0
STI Treatment	0	+

Table A5: Effect of Choices on Outcomes

	Birth Control	Sex	Condom	Birth	STI Treatment
Pregnancy	-	+	-	0	0
STI	0	+	-	0	0

Table A6: Effect of Dependent Coverage Mandate on Insurance Rate

Age Groups (Year Olds)			
20-25	26-30	31-35	41-45
0.139**	0.087	-0.084	-0.057
(0.057)	(0.073)	(0.082)	(0.064)
2008-09 Insurance Rate			
0.688	0.769	0.829	0.851
Percent Change			
6.3%	2.6%	-1.7%	-1.0%

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE)
Percent change = (estimated effect * (1 - baseline)) / baseline. For 20-25 year olds, (0.139 * (1-0.688)) / 0.688

Table A7: Unweighted Effect of Dependent Coverage Mandate (20-24 Year Olds)

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Change in Intercept	-0.057	-0.224	-0.486
	(0.050)	(0.147)	(0.385)
Change in Slope	0.033	-0.211	-0.172
	(0.029)	(0.133)	(0.283)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2012

Table A8: Unweighted Effect of Zero Cost-Sharing Mandate (25-29 Year Olds)

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Change in Intercept	-0.012	0.501**	0.394
	(0.043)	(0.223)	(0.583)
Change in Slope	-0.053	0.125	-0.044
	(0.043)	(0.132)	(0.367)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2014

Table A9: Effect of Dependent Coverage Mandate (20-24 Year Olds), Excluding Earlier Medicaid Expansion States

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Change in Intercept	-0.058	-0.432**	-0.289
	(0.053)	(0.181)	(0.327)
Change in Slope	-0.017	-0.403**	-0.366
	(0.029)	(0.155)	(0.236)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level, weighted by 2011 state age-specific pop.
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2012

Table A10: Effect of Zero Cost-Sharing Mandate (25-29 Year Olds), Excluding Earlier Medicaid Expansion States

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Change in Intercept	-0.007	0.534**	0.231
	(0.034)	(0.206)	(0.313)
Change in Slope	-0.082**	0.199	-0.243
	(0.033)	(0.156)	(0.225)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level, weighted by 2011 state age-specific pop.
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2014

Table A11: Effect of Zero Cost-Sharing Mandate (30-34 Year Olds)

	Log Birth Cases	Log Chlamydia Cases	Log Gonorrhea Cases
Change in Intercept	0.020	0.460***	0.244
	(0.041)	(0.142)	(0.357)
Change in Slope	-0.036	0.338***	-0.062
	(0.044)	(0.102)	(0.213)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state- level, weighted by 2011 state age-specific pop.
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2014

Table A12: Falsification Test - Effect of Dependent Coverage Mandate

	Log Syphilis Cases	Log HIV Cases[27]	% Women	% White	Unemp. Rate	Income Per Capita	Strict Abortion Reg.
Change in Intercept	0.265	0.152	0.000	-0.018	0.337	1231.5	-0.092
	(0.583)	(0.375)	(0.002)	(0.012)	(0.792)	(1402.0)	(0.184)
Change in Slope	-0.144	-0.240	0.002	0.008	-2.101	-1618.0	-0.318
	(0.448)	(0.269)	(0.004)	(0.008)	(1.609)	(1980.4)	(0.193)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level, weighted by 2011 state age-specific population
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2012, HIV starts in 2008

Table A13: Falsification Test - Effect of Zero Cost-Sharing Mandate

	Log Syphilis Cases	Log HIV Cases[28]	% Women	% White	Unemp. Rate	Income Per Capita	Strict Abortion Reg.
Change in Intercept	0.571	0.216	-0.005	-0.013	0.317	-1165.53	0.412
	(0.705)	(0.25)	(0.005)	(0.016)	(1.537)	(2766.241)	(0.452)
Change in Slope	0.498	-0.121	0.003**	-0.006	1.181	-1406.594	-0.053
	(0.335)	(0.151)	(0.001)	(0.007)	(1.311)	(889.75)	(0.099)

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE), cluster at state-level
Controls: unemployment rate (total and age-specific), population (total and age-specific), income per capita, strict state regulation of abortion, and state mandates
Years: 2006-2014, HIV starts in 2008

[27] Age group is 25-34 year olds
[28] Age group is 25-34 year olds

CHAPTER 2

The Effect of Weight on Mental Health: New Evidence Using Genetic IVs*

Abstract: Average body mass index (BMI) and depression prevalence grew over the last several decades, increasing medical expenditures. There is a strong correlation between obesity and depression but limited evidence on the causal effect of weight on mental health. I use an index of genetic risk for high BMI as a source of exogenous variation in weight to provide novel evidence on the effect of weight on mental health. This is one of the first studies to use genetics as an instrument for BMI and to examine the causal relationship between weight and depression. Results are mixed; I find a meaningful and significant effect of weight on suicidal ideation but no effects on counselling and an index of depression. The effect on suicidal ideation is concentrated in white females. From respondent and interviewer opinions of respondent attractiveness, social stigma is a mechanism through which weight affects mental health for white women.

*For a final version of this paper, visit www.sciencedirect.com/science/article/pii/S0167629617303223. Citation: Willage, Barton. The Effect of Weight on Mental Health: New Evidence Using Genetic IVs. *Journal of Health Economics*. 3 (2018): 113-130.

This research uses data from Add Health, a program project directed by Kathleen Mullan Harris and designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill, and funded by grant P01-HD31921 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Information on how to obtain the Add Health data files is available on the Add Health website (http://www.cpc.unc.edu/addhealth). No direct support was received from grant P01-HD31921 for this analysis.

1. Introduction

In the United States, there has been an across the board shift in the weight distribution. The prevalence of obesity (BMI ≥ 30) has increased dramatically from less than 20% in the early 1960s to over 30%, with a similar decrease in healthy weight (BMI 18.5–24.9) (Centers for Disease Control and Prevention, 2003; Fryar et al., 2012). As seen in Appendix Figure A1 and Appendix Figure A2, no group has avoided increases in obesity (Fryar et al., 2012). Commensurate with the rise in weight has been a marked increase in the rate of depression, from 3.3% in 1991-92 to 7.1% in 2001-02 (Compton et al., 2006). Increases in weight and frequency of mental illness are imposing large costs on society. Medical expenditures on weight-related care have reached $200 billion per year in the United States, and Americans spend $100 billion annually on mental health treatments (Cawley and Meyerhoefer, 2012; Mark et al, 2007). These medical expenditures cause negative externalities to other members of insurance risk pools and taxpayers in the case of public insurance. Additionally, mental health may serve as an important pathway through which weight affects labor market outcomes. Mental illness also can have negative spillovers to family members, in terms of decreased earnings and damaged relationships which can lead to divorce (Lépine and Briley, 2011).

The rise in weight – particularly the severity and the prevalence of obesity – has led to a great deal of interest in the effects of weight on physical health (Fryar et al., 2012; World Health Organization, 2017). However, there is much less research about the causal effects of weight on mental health. The influence of body mass index (BMI) on mental health has important ramifications for how we measure the overall impact of the ongoing obesity epidemic. This paper aims to help fill the gap in the literature on the effect of weight on mental health by using a genetic index of risk for high body mass index to instrument for BMI. In addition, this research addresses the possible

mechanisms – body image, disrupted social relationships including stigma, and health problems – through which weight affects mental health. Understanding the empirical relevance of these various mechanisms is important in shaping policies to address depression.

Not only have the increases in weight and depression coincided over time, but the correlational evidence is quite clear: there is a strong positive association between high BMI and depression (Luppino et al., 2010; Pratt and Brody, 2014; Scott et al., 2010; Zhao et al., 2009). The correlation between weight and depression could operate through three non-mutually exclusive potential pathways: (1) higher BMI harms mental health, (2) poor mental health causes weight gain – such as through changes in diet or as a side effect of anti-depressants, or (3) a third factor – such as parental SES – affects both mental health and weight. This paper tests the first pathway, whether being heavier harms mental health. It is important to distinguish between these three relationships. For instance, if higher BMI has a negative effect on mental health, it has important ramifications for how we calculate the cost of the obesity epidemic. Similarly, efforts to reduce BMI can have important spillovers that improve mental health; as communities, policymaker, and employers invest more resources to fight increasing BMIs, it will be useful to consider the mental health benefits of any success. Additionally, paying special attention to overweight and obese people during mental health screenings could better identify individuals with increased risk of mental health problems.

Because weight is not randomly assigned and typically is based on a number of factors that researchers cannot observe, overcoming problems associated with endogeneity is very difficult in this literature. Consequently, there is a dearth of studies that use credible causal methods to estimate the effect of weight on mental health. Two recent papers make progress on this difficult endogeneity problem by

using biological relatives' BMI/obesity as instruments for respondents' weight. Sabia and Rees (2015) find that increased weight adversely affects mental health, primarily for females, while Bjørngaard et al. (2015) find higher weight increases depression but not anxiety. While this instrument provides a clear advancement relative to prior associational work, it cannot disentangle the causal effects of BMI from the effects stemming from other common factors within the household.

This paper contributes to the literature by leveraging individual level genetic information to examine the effect of weight on mental health within an instrumental variables (IV) strategy. The use of genetic information – derived from biological samples and laboratory analysis – as an instrument for BMI is in contrast to much of the previous literature that relied on biological relatives to proxy for genetics (Bjørngaard et al., 2015; Sabia and Rees, 2015). I use the National Longitudinal Study of Adolescent to Adult Health (Add Health), which contains data on height and weight, information on mental health, and an index of genetic predisposition for high BMI. The BMI Genetic Risk Score, the instrument used in this study, is based on 31 individual pieces of genetic information known as Single Nucleotide Polymorphisms (SNPs) that are related to elevated BMI. The SNPs included in the BMI Genetic Risk Score are based on a meta-analysis of genome-wide association studies (GWAS) and verification by geneticists (Speloites et al., 2010).

The central identifying assumption underlying my approach is that the BMI Genetic Risk Score is only related to mental health through weight. The main concern is that the SNPs in the BMI Genetic Risk Score are correlated with other variables, such as family socio-economic status, which also independently affect mental health. I show that the instrument is unrelated to a series of individual background characteristics, such as parental education and income, and I also control for parental obesity to account for concerns related to the correlation of parent and child obesity

and mental health. By using a genetic index as the instrument, I use only the variation in BMI related to those genes in the index (Speloites et al., 2010). While one can never completely prove that the exclusion restriction holds, it allows me to relax the assumption associated with using a biological relative as an instrument for obesity, that shared environmental factors or genetics unrelated to weight are uncorrelated with mental health outcomes.

My results for the effect of weight on mental health are mixed. I find an economically meaningful and statistically significant effect of weight on suicidal ideation at the 10% level, but not use of counseling services or a depression index. The effect on suicidal ideation is primarily in white females. A concentration of effects in females is consistent with Sabia and Rees (2015) as well as with the literature on the effect of weight on labor market outcomes (Averett, 2014). However, I find a weaker relationship between weight and mental health than previous research using sibling BMI as an instrument. When I examine mechanisms, I find that social stigma, as measured by the interviewer's opinion of the respondent's attractiveness, is a mechanism through which weight affects mental health for this group.

This paper is organized as follows: Section 2 presents the data, including the instrument and the measures of BMI and mental health; Section 3 explains the research design and method; Section 4 presents empirical results; Section 5 explores potential mechanisms; and Section 6 concludes with a discussion.

2. Data

This paper uses data from the National Longitudinal Study of Adolescent to Adult Health (Add Health). Add Health, a national panel study, began data collection when

participants were in 7th through 12th grade. The first two waves of data collection occurred while respondents were adolescents in 1995 and in 1996; waves 3 and 4 were collected during early adulthood in 2001-2002 and in 2008 when respondents were aged 18-34. Add Health contains a wide variety of information, including details about the respondent's school reported by an administrator and information about family and household reported by an adult household member, as well as neighborhood and community characteristics from administrative datasets such as the Centers for Disease Control Sexually Transmitted Disease File.

Importantly, Add Health collected information about height, weight, and mental health in all four waves. The BMI Genetic Risk Score was collected in wave 4 and is time invariant. Most control variables, including parental characteristics, were collected at baseline and are treated as fixed. This information is in Appendix Table A1 as well as summary statistics on the data aggregated across waves. Respondents are quite young and results might not extrapolate to older populations. However, young adults have higher rates of depression than older adults (Compton et al., 2006), so this maybe the most appropriate group to examine.

I calculate the respondent's BMI, the main variable of interest in this analysis, using measures of height and weight collected by Add Health field interviewers in waves 2-4 and self-reported information in wave 1 when measured values are not available.[29] Add Health contains a rich set of mental health outcomes, including

[29] Estimates with obesity (BMI>30) and waist circumference available upon request and have similar conclusions. Results based on obesity are not presented because the statistical methods used assume continuous treatment. Also, empirical evidence suggests the effects on mental health are across the weight distribution, so dichotomizing BMI overstates the effect of obesity on mental health; the instrument affects mental health not just through obesity but through variation in BMI within obese and

information on counseling, suicidal ideation, suicide attempt, injury from suicide attempt as well as a subset of questions from the Center for Epidemiologic Studies Depression Scale (CES-D), which I standardized to have mean of zero and standard deviation of one.[30,31,32] The CES-D is an established tool to measure symptoms of depression and has been tested in multiple settings for validity and reliability (Lewinsohn et al., 1997; Radloff, 1977). Suicidal ideation can be used as a useful measure of utility. One theory is that a fully rational individual who wants to commit suicide has a negative expected lifetime utility (Hamermesh and Soss, 1974). Another theory is that professed suicidal ideation can serve as an expression of low utility (unhappiness), perhaps in order to extract sympathy and resources (Cutler et al., 2001). To examine potential mechanisms, I use information on respondents' own sense of attractiveness as a measure of self-image and interviewer-rated attractiveness

---

non-obese categories. Effects across the BMI distribution is consistent with the research on wages, particularly for women (Gregory and Ruhm, 2011). Results based on waist circumference are not presented because the genetic index is designed for BMI rather than waist circumference.

[30] Completed suicide information is not yet available at this time, because the genetic information that serves as the instrument was collected in the most recent wave of the Add Health data. Only 2.3% of respondents report a suicide attempt in the past year, and less than 1% injured themselves during an attempt. This lack of variation in these outcomes requires a much larger sample to detect the effect of weight, so estimates on these outcomes are not presented here, but available upon request.

[31] Counseling is based on the question: "In the past year, have you received psychological or emotional counseling?"
Suicidal ideation is based on the question: "During the past 12 months, did you ever seriously think about committing suicide?"
The depression index is standardized to be mean 0, standard deviation 1. Based on how often the following were true in the past week: You were bothered by things that don't usually bother you; You felt that you could not shake off the blues, even with help from your family and your friends; You felt you were just as good as other people; You had trouble keeping your mind on what you were doing; You felt depressed; You felt that you were too tired to do things; You enjoyed life; You felt sad; You felt that people disliked you. Each question takes on a value of one to three, and the index is the sum of the values.

[32] The pair-wise correlations between mental health outcomes examined here are: corr(suicidal ideation, counseling)=0.189; corr(suicidal ideation, depression index)=0.305; corr(counseling, depression index)=0.173.

as a proxy for stigma.[33]

An important characteristic of the Add Health study is the sampling design. Add Health randomly sampled clusters of high schools and collected data from a subsample of students in those schools. This sample design enabled Add Health to develop data about school environment and peer networks. Add Health oversampled certain subpopulations of interest, including the children of college-educated black adults and individuals with physical disabilities, who may have a relatively small presence in a school. Generally, when a study uses a sample design that is not a simple equal probability sample, analyses should account for both clustering and oversampling.

In this study, I cluster at the school level but use no sample weights. I cluster at the school level as opposed to the level of treatment assignment (individual or family) because all individuals and almost 80 percent of families are completely subsumed within schools. By clustering at the larger level, I account for the correlation within the smaller clusters, and according to Cameron and Miller (2015) "[t]he consensus is to be conservative and avoid bias and use bigger and more aggregate clusters." Add Health personnel recommended that no available sample weights be used in the analysis of the BMI Genetic Risk Score sample (National Longitudinal Study of Adolescent to Adult Health, personal communication, 2015). Additionally, if the sample weights are unrelated to the outcomes conditional on controls, unweighted and

33 Respondent-rated attractiveness is based on the question "How attractive are you?" Takes value 1 if response "not at all attractive" or "slightly attractive", 0 if "moderately attractive" or "very attractive". Interviewer-rated attractiveness based on the field interviewers answer to the question "How physically attractive is the respondent?" Takes value 1 if response "very unattractive", "unattractive" or "about average"; 0 if "moderately attractive" or "very attractive".

weighted regressions are both consistent, but weighting can reduce precision (Solon et al., 2015). The results from this analysis are thus local to the sample population and are not representative at the national level.

While the literature on the causal effect of weight on mental health is thin, many studies have used instruments for BMI to look at the effect of weight on other outcomes, such as employment and wages (Averett, 2014; Averett and Korenman, 1996; Cawley, 2004). Most of these studies use the BMI of a biological relative to instrument for the respondent's BMI. The relative's BMI serves as a proxy for shared genetics that are related to weight (Averett, 2014). Researchers have started to use respondents' own genetic information as collecting and analyzing genetic samples has become more common in social science datasets (Böckerman et al., 2017; Tyrrell et al., 2016; von Hinke et al., 2016).

The BMI Genetic Risk Score, the instrument in this study, is based on 31 Single Nucleotide Polymorphisms[34] previously identified as associated with increased BMI by Speloites et al. (2010). The score can range from 0 to 62, as a person can have 0, 1, or 2 risky alleles for each of the 31 SNPs.[35] To determine which SNPs were associated with increased BMI, Speloites et al. (2010) first performed a meta-analysis

[34] There are four nucleotides: cytosine, adenine, guanine, and thymine. DNA (deoxyribonucleic acid) is composed of a sequence of these nucleotides. If a nucleotide occurs in a particular position with at least moderate frequency (more than one in one hundred people) then this location-nucleotide combination is a SNP (for instance, if more than one percent of the population has cytosine in a certain location in their DNA, that would be a SNP) (Nature.com, 2014).

[35] As of June 14, 2017 the online documentation for the BMI Genetic Risk Score is incorrect (http://www.cpc.unc.edu/projects/addhealth/documentation/restricteduse/datasets/GRS_BMI.pdf). It states that the score is standardized (mean 0, standard deviation 1) by race. However, based on communication with Add Health staff, this description applies to how the data were used in Domingue et al. (2014) (National Longitudinal Study of Adolescent to Adult Health, personal communication, 2016). The fact that the BMI Genetic Risk Score is not standardized is confirmed in a frequency table at the bottom of the online documentation.

of 46 genome-wide association studies. She and her co-authors then performed their own analysis on the 19 SNPs from the meta-analysis and found an additional 12 SNPs associated with BMI. In wave 4, Add Health collected the genetic information used to calculate the BMI Genetic Risk Score using saliva samples. Saliva samples were collected from 1,946 respondents, 1,886 have valid genetic data, and 1,733 individuals have all information required for valid BMI Genetic Risk Scores, though not all respondents with BMI Risk Scores were present in all waves of data. Of the 1,733 respondents with valid BMI Genetic Risk Scores, there are 595 sibling pairs, 30 sibling trios, 1 sibling quad, and 449 stand-alone respondents. The Siblings Pair dataset is based on approximately 1,600 individual from whom genetic samples were collected, which is described in the BMI Genetic Risk Score documentation (National Longitudinal Study of Adolescent to Adult Health, 2017).

3. Method

I obtain estimates of correlation between BMI and mental health using ordinary least squares models. While OLS models are important to see associations between weight and mental health, they can suffer from bias in the form of reverse causality (mental illness causes weight gain) and unobserved confounders such as self-control. Von Hinke et al. (2016) carefully outline an argument for the credibility of a genetic index as an instrument for weight based on the instrumental variable assumptions: independence, exclusion, non-zero correlation of instrument and treatment, and monotonicity. I address the same issues and concerns in a new context: mental illness. I combine the discussion of independence and exclusion in terms of "validity";

likewise, I combine monotonicity and non-zero correlation of instrument and treatment as "strength". I discuss the empirical evidence regarding the IV assumptions in the Results section.

While the issue of instrument strength can be empirically examined, instrument validity – that the BMI Genetic Risk Score is only related to mental health by affecting weight – is not as clear-cut. One benefit of genetic instruments involves the concept of Mendelian randomization; genetics are randomly passed from parents to children (Lawlor et al., 2008). Therefore, after controlling for parental genetics, children's genetics are as-if randomly assigned. While I am unable to condition on parents' genes, I do control for parental characteristics including self-reported obesity. Potential threats to the validity assumption that I cannot investigate with the data in Add Health include pleiotropy (when a gene influences multiple characteristics) and biological mechanisms, though I address these issues with evidence from existing research.

In terms of pleiotropy, if a SNP that is related to weight also is directly related to mental health outcomes, the SNP would not be a valid instrument. As I cannot examine this concern with the data available, I follow von Hinke el al. (2016) by checking if the SNPs in the BMI Genetic Risk Score are also found to be related to mental illness in GWAS. Based on literature reviews, the SNPs in the BMI Genetic Risk Score are not directly linked to mental illness such as major depression, bipolar disorder, or schizophrenia (Flint and Kendler, 2014; Gurung and Prata, 2015; Ripke et al, 2013). In aggregate, if BMI affects mental health, then there is an indirect link between the SNPs and mental health; however, with current information on genetic

mechanisms, there is not a definitive method to disentangle a potential direct effect of the SNPs on mental health from an indirect effect through weight. While not much is known about the biological mechanisms of SNPs generally, the current evidence indicates that the SNPs in the BMI Genetic Risk Score impact weight through mechanisms that are consistent with the IV assumptions, such as energy homeostasis (involved in hunger) and gastric inhibitory polypeptide (involved in insulin secretion) (Locke et al., 2015).

To estimate the causal effect in the IV approach, I estimate two stage least squares (2SLS) regression models. The estimating equations are:

First Stage:

$$BMI_{it} = \alpha_0 + \alpha_{riskscore} GeneticRiskScore_i + \boldsymbol{\alpha_x X_{itjk}} + \zeta_{\boldsymbol{itjk}}$$

Second Stage:

$$Y_{it} = \beta_0 + \beta_{BMI} \widehat{BMI_{it}} + \boldsymbol{\beta_x X_{itjk}} + \epsilon_{\boldsymbol{itjk}},$$

where each observation occurs at time *t* for individual *i*, who is a member of family *j*, and sampled from school *k*. Therefore, each observation nests within individuals who are nested within both schools and families, and most families are nested within schools (78%).

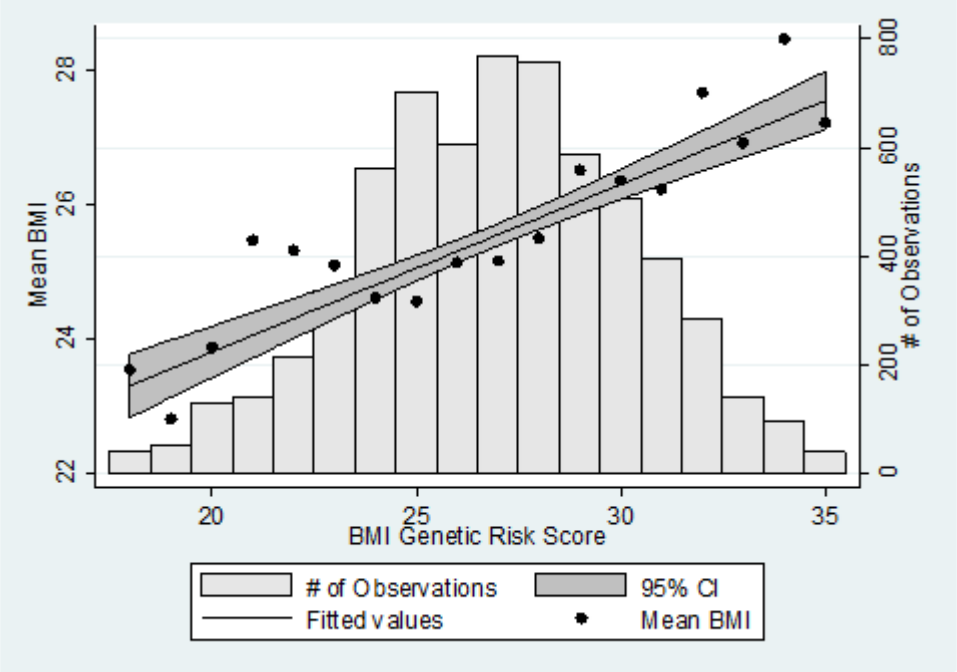
In the first stage, the endogenous regressor of interest (BMI) is regressed on the instrument (BMI Genetic Risk Score). In the second stage, the outcome (depression index, counseling, suicide ideation) is regressed on the fitted value of BMI from the first stage. I also estimate reduced form models in which I regress mental health

outcomes directly on the BMI Genetic Risk Score. These reduced form models relax the functional form assumptions. Though these models have no clear policy implication, they are included as standard good form. To estimate the mechanism by which BMI affects mental health, I estimate the same 2SLS models and use interviewer- and self-rated attractiveness as the outcomes. In all models, I control for individual and area characteristics to increase precision. Controls include age, age squared, native born, wave, race/ethnicity, parent education, self-reported parent obesity (continuous measure not available), county unemployment rate, gender, intelligence (picture vocabulary and number recall).

4. Results

As previously mentioned, the credibility of the instrument rests on two main assumptions. The first assumption is that the instrument is highly correlated with BMI. By construction, the BMI Genetic Risk Score is correlated with BMI. Further, this assumption can be tested empirically. As seen in Figure 1, the BMI Genetic Risk Score and BMI have a strong positive relationship. The mean BMI Genetic Risk Score in the sample is 27 and the standard deviation is 3.44, so for every five SNPs (or about 1.4 standard deviations) that are associated with increased BMI, BMI increases by 2 units. For a 6-foot tall individual, a BMI increase of 2 units corresponds to a weight gain of 15 pounds. Additionally, I present first stage F-statistics from the two stage least squares models. All F-statistics are above the traditional cut-off of 10 for weak instruments (Stock and Yogo, 2005), and many F-statistics are in the 20s and 30s, indicating that the BMI Genetic Risk Score is a strong instrument for BMI.

Figure 1: Relationship between BMI and Genetic Risk Score and Number of Observations by Risk Score



Data Source: The National Longitudinal Study of Adolescent to Adult Health
BMI Genetic Risk Score range restricted due to small numbers of observations at extreme values

Figure 1 also presents evidence of monotonicity. While the relationship between the BMI Genetic Risk Score and BMI is somewhat noisy, the relationship appears largely linear and increasing; a higher BMI Genetic Risk Score is associated with higher average BMI. The one exception is for BMI Genetic Risk Score from 21 to 23. This is likely not caused by an actual violation of the monotonicity assumption, but due to the fact that the mean BMI for each bin is fairly imprecise due to few observations. As a robustness test, those with BMI Genetic Risk Score below 24 are excluded.[36]

[36] First stage estimates increase, while reduced form estimates are substantively the same but are not statistically significant due to loss of precision (results available upon request).

Linearity between outcomes and the predicted BMI based on first-stage regressions (Appendix Figures A3-A5) can be assessed from appendix graphs. These graphs are important because the effect of predicted BMI on outcomes is estimated in the second stage of 2SLS models and assumes linearity. Relationships appear linear, though with limited data points. From Appendix Figure A3 there is some indication that a relationship between predicted BMI and suicidal ideation is largely caused by effects at low predicted BMI. However, very few respondents have predicted BMI below 24 (3%), and so estimates are unlikely to be driven by changes at low predicted BMI. This is reflected in the fact that probability of suicidal ideation at other predicted BMI values lay close to the fitted line. Additionally, effect sizes are largely robust to dropping respondents with BMI Genetic Risk Scores under 24, which includes all respondents with low predicted BMI.[37]

Table 1 presents first stage estimates of the effect of the BMI Genetic Risk Score on BMI.[38] There are several important takeaways. First, estimated effects of the BMI Genetic Risk Score are stable across racial and gender groups: a unit increase in the BMI Genetic Risk Score increases BMI by approximately 0.3 units. Second, all F-statistics are over the traditional cut-off of 10, and often above 20. Third, the marginal R-squared, the additional percent of total variance in BMI explained by controlling for the BMI Genetic Risk Score, ranges from 2-5% (in square brackets).

[37] See footnote 10 for discussion about results when respondents with BMI Genetic Risk Scores below 24 are excluded.

[38] Ideally, the data would have sufficient sample size to include household fixed effects. In this case, the identifying variation would be based solely on differences between siblings. However, including household fixed effects leaves very little variation and uses much of the statistical power. While the estimated relationship is fairly stable, F-statistics drop to below ten, and the partial R^2 falls to about 1%.

Table 1: First Stage - Effect of the Genetic Risk Score on BMI (Mean=25.6)

Panel A: Males and Females Pooled			
	Full Sample	Whites	Blacks
Genetic Risk Score	0.308***	0.337***	0.353***
	(0.0401)	(0.0463)	(0.0730)
	N=5486	N=2876	N=1841
	F=59.194	F=52.912	F=23.702
	[0.030]	[0.043]	[0.030]
Panel B: Males			
	Full Sample	Whites	Blacks
Genetic Risk Score	0.318***	0.348***	0.355***
	(0.0583)	(0.0686)	(0.106)
	N=2677	N=1408	N=868
	F=30.024	F=25.923	F=11.851
	[0.036]	[0.049]	[0.039]
Panel C: Females			
	Full Sample	Whites	Blacks
Genetic Risk Score	0.296***	0.311***	0.323***
	(0.0524)	(0.0694)	(0.0976)
	N=2809	N=1468	N=973
	F=31.554	F=20.166	F=10.399
	[0.025]	[0.034]	[0.021]

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Partial R^2 in square brackets
Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

The second IV assumption – validity or that the instrument is unrelated to the outcomes except through BMI – cannot be tested directly. The two ways this assumption could fail are: (1) the effect of the instrument on the outcome operates through a mechanism besides BMI (exclusion) and (2) some third factor causes changes in the instrument and the outcome (independence). I empirically assess both of these possible threats to validity by examining the relationship between the BMI Genetic Risk Score and other predictors of depression.

To assess the relationship between the BMI Genetic Risk Score and other characteristics, I compare respondents with high and low genetic risk. This is a common approach to investigate the validity of an instrument (McClellan et al., 1994). Appendix Table A2 contains the mean and standard error of the mean of characteristics that are related to BMI and mental health for those with low (below the median of 27) and high (above the median) BMI Genetic Risk Scores in wave 1. In addition to race (addressed in regards to respondent characteristics below), the only significant difference among parental characteristics is that those with higher BMI risk are more likely to have obese parents, which is expected because parents and children share genes. A concern with the heritability of risk of high BMI is that parents of heavier children are heavier and thus might earn less. In this case, the BMI Genetic Risk Score could violate the validity assumption by impacting heavier children's mental health through socioeconomic status. However, childhood household income and parents' education are similar across levels of BMI risk.

Another concern is that parental weight can impact parental mental health, and depressed parents could affect their children's mental health. The concern is that shared genes related to BMI result in correlated child/adult weight, and that higher adult weight could affect child mental health. The effect of parent mental health, net of the effect of parent weight, on child mental health would not violate the IV assumptions. In order to prevent bias from shared genetic disposition for high BMI, I control for a measure of parental obesity in all models as well as parental education. Ideally, I would control for parent mental health to increase precision, but such measures are not available in Add Health.

Appendix Table A2 also shows differences in picture vocabulary test and grade point average that are statistically significant, but the differences are quite small in magnitude. A difference of 2.3 in picture vocabulary test is only 1/6 of a standard deviation, and the difference in GPA is less than 1/10 of a standard deviation. Additionally, academic achievement is endogenous (affected by weight), and so differences in GPA are plausibly a consequence of higher BMI's effect on mental health (Sabia and Rees, 2015). There is a four percentage point difference in prevalence of native born, but both groups are almost completely born in the United States. I include picture vocabulary and native born in all models. Importantly those with higher BMI risk have higher BMI but very similar heights; estimated effects of BMI will be based on variation in weight instead of height. Whites and blacks differ in base rates of the SNPs of interest, because most genetic studies of the effect of SNPs on BMI use white samples as well as due to potential population stratification (different prevalence of genes caused by historic separation of the ancestors of white and black individuals). Thus, I perform all analyses separately for whites and blacks, as suggested by von Hinke et al. (2016).

Appendix Table A3 shows a composite measure of the potential bias from violating the validity assumption for the black and white pooled sample. First, predicted outcomes are based on regressing a mental health measure on observable characteristics. Then the predicted value for the mental health outcome, which is a linear combination of observable characteristics, is regressed on the genetic risk score controlling for parent obesity. The estimates in Table A3 are the association between the BMI Genetic Risk Score and variation in observable characteristics predictive of

mental health issues.

Though the relationships are statistically significant, the correlations are very small. For suicidal ideation, the correlation is 1/3 the reduced form estimate and 1/10 the two stage least squares estimate of the effect of BMI, and thus any bias from violating the validity assumption is unlikely to explain the results for suicidal ideation. The other correlations are also very small and comparable in magnitude to the reduced form estimates, which are not statistically significantly different from zero. Importantly, once the analysis is stratified by race (Appendix Table A4 and Appendix Table A5) the estimates are much smaller and many of the associations are not statistically significant; the significant relationships found in Appendix Table A3 load on race. Since the main results are also stratified by race, the balance in Appendix Tables A4 and A5 further mitigate concerns about the validity assumption.

Since results from reduced form linear probability models are similar to the marginal effect obtained from probit models, I present linear probability models, which are easier to interpret.[39] Tables 2-4 present ordinary least squares (left panel) associations and two stage least squares (right panel) estimates of the effect of BMI on mental health outcomes. Ordinary least square estimates are the association between BMI and the outcome. The 2SLS estimates are the causal effect of a one-unit increase in BMI on the outcome. Comparing OLS and 2SLS estimates can reveal potential bias from using an endogenous measure of treatment.[40] However, differences could also be

[39] For example, see Appendix Table A6 for LPM and Appendix Table A9 for probit marginal effects.
[40] The direction of the bias is ambiguous ex ante. The direction of omitted variable bias depends on the correlation between the omitted variable with both the treatment and the outcome, while bias from reverse causality – mental health affecting weight –could be up or down due to the effect of depression on diet (American Psychiatric Association, 1994).

due to the fact OLS correlations are average treatment effects (ATE) and IV estimates are local average treatment effects (LATE) based only on variation in BMI due to genetic predisposition.[41] If high BMI caused by genetics is more difficult to reverse or prevent, then intuitively the LATE in this context (effect of BMI on mental health specifically caused by genetics) would be larger than the correlation.

Table 2 presents OLS and 2SLS estimates of the effect of BMI on suicide ideation. Effects of suicidal ideation are of particular interest, because it is the most severe measure of mental health problems I examine. From the right panel, BMI has a statistically and economically significant effect on suicidal ideation. Consistent with research on BMI's effect on labor market outcomes and mental health, the effects are concentrated in women and particularly white women (Averett, 2014; Sabia and Rees, 2015). Each unit increase in BMI causes a 0.87 percentage points increase in suicidal ideation, off a mean of 9.7 percent. The effect is almost double for white women at 1.4 percentage points. For women and whites, the estimated effects are much larger than the correlations. The estimated correlations and estimated effects of BMI are smaller and not statistically significant for males and blacks. However, IV point estimates are still positive and of moderate magnitude for all groups; a one-unit increase in BMI causes half a percentage point increase in suicidal ideation for men. Additionally, the estimate for black women is almost as large as for white women but is not statistically significant, perhaps due to a smaller sample size in that subpopulation.

[41] While OLS estimates are average treatment effects and 2SLS estimates are local average treatment effects, the 2SLS estimates are based on variation across the BMI distribution (see Figure 1).

Table 2: OLS and 2SLS - Effect of BMI on Suicidal Ideation (Mean=0.097)

	Ordinary Least Squares			***Two Stage Least Squares***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00152**	0.00191	0.00148	0.00876*	0.0103**	0.00490
	(0.000720)	(0.00118)	(0.00122)	(0.00469)	(0.00508)	(0.00657)
	N=5437	N=2857	N=1818	N=5437	N=2857	N=1818
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.000543	-0.000446	-0.000241	0.00546	0.00552	0.00335
	(0.000869)	(0.00124)	(0.00174)	(0.00548)	(0.00727)	(0.00685)
	N=2648	N=1398	N=852	N=2648	N=1398	N=852
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00283**	0.00357*	0.00258*	0.0116*	0.0140*	0.0136
	(0.00110)	(0.00186)	(0.00146)	(0.00703)	(0.00795)	(0.0100)
	N=2789	N=1459	N=966	N=2789	N=1459	N=966

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Standard errors stable to two-way clustering (school and family, school and individual), see Appendix Table A10

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health

While suicidal ideation is a fairly severe indication of depression, changes in the depression index and use of counseling services can reflect effects of weight on people who have relatively good mental health. The estimated effect of BMI on the depression index and on use of counseling service in the past year are small and exclusively not statistically significant. The largest 2SLS estimate for the effect of BMI on the depression index (Table 3) is 0.031 for white females. While the estimated effect for women is much larger than the estimated correlation of 0.011 based on OLS, it is still a small fraction of the standard deviation of the depression index (SD=1). A 5'4" women who weighs 130 pounds (~22 BMI) could increase her weight by 50 pounds (~8 BMI), and the resulting increase in her depression index would be much smaller than a standard deviation in this outcome (8 BMI x 0.0308 ≈ 0.25). Results for a dichotomized depression score (above vs below average) also result in small estimates that are not statistically significant (results available upon request).

Table 3: OLS and 2SLS - Effect of BMI on Depression Index (Mean=0.0)

	Ordinary Least Squares			***Two Stage Least Squares***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00507*	0.00589	0.00745*	0.00283	0.00267	0.0200
	(0.00304)	(0.00452)	(0.00412)	(0.0198)	(0.0258)	(0.0349)
	N=5477	N=2871	N=1838	N=5477	N=2871	N=1838
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.000145	-0.00248	0.00866	-0.00903	-0.0259	0.0329
	(0.00400)	(0.00528)	(0.00693)	(0.0260)	(0.0381)	(0.0459)
	N=2671	N=1404	N=866	N=2671	N=1404	N=866
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00857**	0.0112*	0.00855*	0.0148	0.0308	0.0103
	(0.00406)	(0.00627)	(0.00503)	(0.0267)	(0.0327)	(0.0498)
	N=2806	N=1467	N=972	N=2806	N=1467	N=972

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Standard errors stable to two-way clustering (school and family, school and individual), see Appendix Table A11Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

The results for the effect on using counseling services (Table 4) tell a similar story, though men are more responsive to weight for this outcome. Again the IV estimates are of greater magnitude than the OLS estimates, but both 2SLS and OLS estimates are generally small compared to the standard deviation of the outcome (SD=0.279). However, the effects on men are of economically significant magnitude but are imprecise. The largest estimated effect is for black males, for whom a one-unit increase in BMI would cause a not statistically significant 1.58 percentage point increase in the probability of using counseling services in the past year. The estimated effect of BMI on counseling is often of the opposite sign from the correlation – indicating potential bias of an endogenous measure of BMI – though all estimated effects and correlations are not significant.

The lack of effect of BMI on these two indicators of less severe[42] mental health problems – depression index and counseling – suggests that those of fair-to-good psychological well-being are unresponsive to weight. In order to test this empirically, I examine the effect of BMI on suicidal ideation for those with high depression index scores (discussed below). Since the data only contain information on the extensive margin of counseling, I am unable to examine the effect of BMI on frequency of counselling. If the intensive margin is responsive to BMI, these estimates could understate the effect of BMI, particularly for people in worse mental health.

[42] This not to say depression cannot be a serious health concern, but that use of counseling services at the extensive margin – the marginal person goes from no use to some – and the topics in the depression index – such as inability to shake off the blues or increased irritability – intuitively fall to the left of suicidal ideation in terms of severity.

Table 4: OLS and 2SLS - Effect of BMI on Counseling Use (Mean=0.085)

	Ordinary Least Squares			***Two Stage Least Squares***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.000233	0.000833	-0.000460	0.00316	-0.000782	0.00493
	(0.000743)	(0.00113)	(0.00108)	(0.00402)	(0.00478)	(0.00639)
	N=5484	N=2874	N=1841	N=5484	N=2874	N=1841
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.000845	-0.000440	-0.000829	0.00768	0.00286	0.0158
	(0.000867)	(0.00131)	(0.00193)	(0.00601)	(0.00678)	(0.0114)
	N=2677	N=1408	N=868	N=2677	N=1408	N=868
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00136	0.00180	0.000486	-0.00158	-0.00749	0.00235
	(0.00110)	(0.00176)	(0.00128)	(0.00627)	(0.00885)	(0.00823)
	N=2807	N=1466	N=973	N=2807	N=1466	N=973

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Standard errors stable to two-way clustering (school and family, school and individual), see Appendix Table A12

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Tables A6-A8 contain estimates of reduced form IV models. Though these models have no clear policy implication, they are included for completeness. In these models, the outcome is regressed on the instrument (BMI Genetic Risk Score) along with a set of controls. The estimates represent the effect of a one-unit increase in the BMI Genetic Risk Score or the presence of one more allele associated with high BMI on mental health outcomes. The results are consistent with the 2SLS regressions above. The BMI genetic risk score has a statistically significant and economically meaningful effect on suicidal ideation, particularly for white females, but not for the black sample or on men. The estimated effects on counseling and the depression index are small and not significantly different form zero.

I also ran a number of analyses to examine heterogeneous effects. Since statistically significant effects of BMI are found only on suicidal ideation, this suggests that BMI influences people who are already in relatively worse mental health. In order to examine this empirically, I estimate the effect of BMI on suicidal ideation for those with above average depression index scores (the index is standardized to mean 0). In Appendix Table A13 are results for the high depression index subsample as well as the full sample. Results are consistent with the hypothesis that individuals with worse mental health are more responsive to weight: Point estimates are much high for the high depression index subsample, and effects are still concentrated in white women.

Two other questions of interest remain: Is there heterogeneity by age and how does the BMI Genetic Risk Score compare to using the traditional instrument of sibling BMI? Results for adults (age>18) are compared to the full sample in Appendix

Tables A14-A16. Again, BMI only has statistically significant effects on suicidal ideation, and the effect continues to be concentrated in white women. Interestingly, the effect of BMI on suicidal ideation is larger for adults, which should address concerns that effects are driven by endogenous misreporting of height and weight in wave 1.

In Appendix Tables A17-A19, I compare estimates using sibling BMI as an instrument to estimates using the BMI Genetic Risk Score. In order to have more comparable analyses, I limit the sample to respondents with BMI Genetic Risk Scores and at least one sibling with BMI information.[43] When restricting the sample, estimates based on the BMI Genetic Risk Score are less precise and similar to the main analysis, except the effects on black women are somewhat unstable.

Pooled gender estimates for suicidal ideation are not statistically significant using sibling BMI as an instrument, because effects for males have decreased. Estimates using the BMI Genetic Risk Score show a moderate if imprecise increase in suicidal ideation for men, but instrumenting with sibling BMI results in reversal of sign – a moderate decrease in suicidal ideation. Effects for white women are similar between instruments, but effects are only statistically significant using sibling BMI in the restricted sample. The estimated effect of weight on the depression index is larger and becomes significant for white women using sibling BMI, and the effect for black men is large but very imprecise. Regardless of instrument, there is no significant effect on counseling use, and many estimates remain precise zeros. Analysis on unrestricted

[43] In order for an observation to be included, the respondent and a sibling must have BMI information in the same wave. If a sibling pair has BMI information in one wave of data collection, they will be included for that wave. If one sibling is missing information in a different wave, they will not be included in that wave.

samples are available in Appendix Tables A20-A22. When using sibling BMI as an instrument, results are similar but smaller for the full sample, and effects on the depression index are no longer statistically significant for white women.

5. Mechanism of BMI Effect on Mental Health

I now explore potential mechanisms through which weight affects mental health. The psychological literature notes three possible mechanisms by which weight can cause worse mental health: body image, disrupted social relationships including stigma, and health problems. First, body image may be important in societies with beauty premiums; being overweight can lower self-esteem and self-worth as the feeling of being attractive decreases (Granberg, 2011). Second, stigma towards heavier people can contribute to depression. Research indicates that heavier people, particularly women, have less successful marriage/romantic partner market outcomes (Averett and Korenman, 1996) and are treated differently in any number of other social interactions. Lastly, high BMI can decrease mental well-being through weight's detrimental effect on physical health. Obesity causes cardiovascular disease and diabetes (Colditz et al., 1995; Hubert et al., 1983), which in turn can decrease the ability to engage in enjoyable activities and to effectively take care of oneself. Co-morbidities increase stress through the financial burden of increased medical care costs. For instance, the lifetime medical cost of developing diabetes can reach $100,000 (Zhuo et al., 2014).

It is important to note that stigma and self-image are two different issues. Stigma harms mental health primarily through worse treatment, such as difficulty

making friends (Granberg, 2011). While stigma could affect self-image if heavier people know they are treated differently because of their weight, stigma may manifest itself in ways that do not affect self-image or in ways that the heavier person does not realize are related to weight/attractiveness. For instance, if stigma plays a role in the effect of weight on labor market outcomes for women such as in Cawley (2004), women may be unaware that their weight and appearance are factors. Therefore, stigma may impact a person and their mental health but not necessarily through self-image. I am using interviewer-rated attractiveness of the respondent as a proxy for stigma.

Given the mix evidence of weight's effect on mental health, investigating the impact of weight on stigma and self-image is important; if these potential mechanisms are unaffected, the evidence that weight harms mental health would be further weakened. Additionally, determining the mechanism by which weight affects mental health can help target policies. For instance, groups that suffer from poor body image could be more responsive to mental health interventions. While broader informational campaigns on the social effects of weight may increase awareness among those who stigmatize obesity, and reduce judgmental behavior and statements. I hypothesize that health problems are unlikely to be a driving factor as respondents are quite young, 12 to 34 years old.[44] Therefore, I examine body image and stigma as viable mechanisms

[44] Appendix Table A23 provides confirmation. First, the prevalence of poor health and diabetes is very low; half a percent of respondents report being in poor general health and two percent have diabetes. Second, the effect of BMI on self-described health and on diabetes (only available in waves 3 and 4) are very small. The exception is for the moderate protective effect of BMI on black respondents', which is the opposite direction one would expect for a mechanism between weight and mental health. Importantly there is no statistically significant effect on diabetes, an objective health outcome highly related to weight.

for this sample. Table 5 contains estimates of the effect of BMI on one's body image, and Table 6 contains estimates of the effect of BMI on the interviewer's perspective of a respondent's attractiveness, which I use as a proxy for stigma. I use this measure to proxy for stigma as I assume that weight's role in what is considered attractive is an important part of the stigma associated with weight.

For white females, the group in which the effect of weight on mental health is concentrated, the mechanism appears to be social stigma. Interviewer-rated attractiveness is highly responsive to weight for this group, while self-rated attractiveness is largely unaffected; this suggest that stigma and differential treatment may manifest in ways that white women are unaware relate to weight/attractiveness. Additionally, women's self-image could be increasingly protected from the effects of weight due to both public service and marketing campaigns promoting positive body image, which have tended to focus on women. For instance, almost 20 years ago the American Dietetic Association began promoting body size acceptance (Parham, 1999), while the Dove Campaign For Real Beauty started in the mid-2000s (Bahadur, 2014). Such campaigns may help protect self-image, but could also reduce reporting of poor self-image due to weight. If a shift towards body acceptance prevents women from reporting poor self-image, the measure used in this study may not capture true problems with self-image.

A unit increase in BMI causes a 2.35 percentage point increase in the probability of the interviewer finds a white female unattractive. In contrast, the point estimate for the effect of BMI on self-rated attractiveness is only 0.32 percentage points. For neither outcome do white males have statistically significant results, which

is consistent with white males' earnings and mental health being largely unaffected by weight, as shown in this research and other studies examining mental health and labor market outcomes (Averett, 2014; Sabia and Rees, 2015).

Black males' self-image and exposure to social stigma are both responsive to increased BMI. The meaningful effect of BMI on mental health for white females and lack of effect for black males in light of the large effects on potential mechanisms is somewhat surprising. While these mechanisms do not manifest themselves in statistically significant effects on the measures of mental health examined here for black males, I estimate BMI has a meaningful but imprecise effect on counseling for this group. Self-image and stigma may also serve as mechanisms for mental health outcomes not examined in this study for this group. Another possible explanation for this difference between black males and white females is that feeling attractive and being perceived as attractive may have a greater effect on the mental health of white women.

Table 5: 2SLS and OLS - Effect of BMI on Self-Rated Unattractiveness (Mean=23.8%)

	Ordinary Least Squares			***Two Stage Least Squares***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00750***	0.0113***	0.00550***	0.0100	0.000792	0.0282***
	(0.00128)	(0.00187)	(0.00170)	(0.00773)	(0.0110)	(0.00897)
	N=2653	N=1404	N=883	N=2653	N=1404	N=883
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00572***	0.00547	0.00886***	0.0153	-0.00803	0.0410***
	(0.00201)	(0.00330)	(0.00271)	(0.0113)	(0.0150)	(0.0152)
	N=1279	N=685	N=408	N=1279	N=685	N=408
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00873***	0.0145***	0.00343	0.00243	0.00320	0.0156
	(0.00192)	(0.00307)	(0.00279)	(0.0116)	(0.0161)	(0.0112)
	N=1374	N=719	N=475	N=1374	N=719	N=475

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Average marginal effect in square brackets [average marginal effect]
Controls: Age, age squared, native born, wave, race/ethnicity, parent education,
county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Table 6: 2SLS and OLS - Effect of BMI on Interviewer-Rated Unattractiveness (Mean=52.9%)

	Ordinary Least Squares			***Two Stage Least Squares***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0154***	0.0191***	0.0125***	0.0108	0.00936	0.0272**
	(0.00122)	(0.00168)	(0.00185)	(0.00720)	(0.00904)	(0.0122)
	N=5476	N=2873	N=1836	N=5476	N=2873	N=1836
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0113***	0.0140***	0.00876**	0.00101	-0.00718	0.0455**
	(0.00188)	(0.00269)	(0.00361)	(0.0102)	(0.0153)	(0.0198)
	N=2672	N=1408	N=865	N=2672	N=1408	N=865
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0180***	0.0226***	0.0150***	0.0189**	0.0235**	0.0135
	(0.00152)	(0.00203)	(0.00263)	(0.00962)	(0.0112)	(0.0181)
	N=2804	N=1465	N=971	N=2804	N=1465	N=971

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Average marginal effect in square brackets [average marginal effect]
Controls: Age, age squared, native born, wave, race/ethnicity, parent education,
county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

6. Conclusion

Overall, I conclude that the evidence on the causal relationship between weight and mental health is mixed. While my findings do not indicate that weight impacts use of counselling services or a depression index, I find a statistically and economically significant effect of weight on suicidal ideation. This effect is concentrated in certain groups, particularly white women. A concentration of effects of weight in white women is consistent with the research on labor market outcomes (Averett, 2014; Cawley, 2004) and the smaller literature on the effect of weight on mental health (Bjørngaard et al., 2015; Sabia and Rees, 2015). However, my findings are less strong than the previous research on mental health that uses sibling BMI as an instrument for own BMI (Sabia and Rees, 2015). Based on respondent and interviewer opinions of respondent attractiveness, social stigma is a possible mechanism by which high BMI affects mental health for white women.

The effect on suicidal ideation is statistically significant at a cut-off of 0.10. While there are many regressions investigating heterogeneous effects by race and gender, there are three main hypothesis tests on the pooled sample (one for each mental health outcome). If there were only statistically significant effects in one subsample and none for the pooled sample, then there may be concerns about multiple comparisons (that a significant finding is due to the combination of many tests and type I error). However, after accounting for three comparisons, the adjusted p-value for the effect of BMI on suicidal ideation for the pooled sample is slightly above the

traditional cut-off.[45]

These findings add to the scant literature that uses econometric methods to examine the causal effect of weight on depression. The effect of weight on mental health also has important implications for the total cost of the ongoing obesity epidemic in the United States. This causal relationship could serve as an important mechanism of weight's effect on labor market outcomes. Additionally, this paper is one of the first economic studies to use measures of genetic information to instrument for BMI. Another contribution is the empirical examination of mechanisms by which weight impacts mental health outcomes, which can have policy implications for educating society about the consequences of stigmatizing weight.

This study is limited to a young sample under 35 years old, so future work should consider the effects of weight on an older population as well as how potential mechanisms operate differently for this group. As more genetic data respondents become available in survey data, particularly in data sets with parent and offspring pairs, the credibility of genetic indices as instruments will continue to improve, because offspring genetic risk is as-if randomly assigned after fully controlling for parental genetic risk. This advancement will further improve genetic information as instruments. Larger data sets with genetic information will provide the required

[45] Using the procedure discussed in Tukey et al. (1985), the adjust p-value for the 2SLS effect on suicidal ideation for the pooled sample is 0.105. The adjusted p-value is

$$p_{adj} = 1 - \left(1 - p_{unadj}\right)^{\sqrt{k}} = 1 - (1 - 0.062)^{\sqrt{3}} = 0.105,$$

where p_{unadj} is the unadjusted p-value and k is the number of tests (3). This adjustment is appealing because it intuitively balances type I and type II error. This method takes into account multiple hypothesis tests and results in adjusted p-values that approach 1 as the number of tests becomes large. However, it is less conservative then the Bonferroni adjustment of multiplying p-values by k, which can result in adjusted p-values over 1.

statistical power to analyze the effect of weight on more severe but less common outcomes such as suicide attempts and completed suicides.

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders,* Fourth Edition. Washington, DC: American Psychiatric Association.

Averett, S. (2014). Obesity and labor market outcomes. *IZA World of Labor*, 32.

Averett, S. and Korenman, S. (1996). The economic reality of the beauty myth. *Journal of Human Resources*, 31(2), 304-330.

Bahadur, N. (2014). Dove 'Real Beauty' Campaign Turns 10: How A Brand Tried To Change The Conversation About Female Beauty. *Huffington Post*. https://www.huffingtonpost.com /2014/01/21/dove-real-beauty-campaign-turns-10_n_4575940.html. Accessed 11 October 2017.

Bjørngaard, J., Carslake, D., Nilsen, T., Linthorst, A., Smith, G., Gunnell, D., and Romundstad, P. (2015). Association of body mass index with depression, anxiety and suicide – an instrumental variable analysis of the HUNT Study. *PloS One*, 10(7), e0131708.

Böckerman, P., Cawley, J., Viinikainen, J., Lehtimäki, T., Rovio, S., Seppälä, I., Pehkonen, J., and Raitakari, O. (2016). *The Effect of Weight on Labor Market Outcomes: An Application of Genetic Instrumental Variables* (No. w22200). National Bureau of Economic Research.

Böckerman, P., Viinikainen, J., Vainiomäki, J., Hintsanen, M., Pitkänen, N., Lehtimäki, T., Pehkonen, J., Rovio, S., and Raitakari, O. (2017). Stature and long-term labor market outcomes: evidence using Mendelian randomization. *Economics & Human Biology*, 24, 18-29.

Cameron, A., and Miller, D. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, *50*(2), 317-372.

Cawley, J. (2004). The impact of obesity on wages. *Journal of Human Resources*, 39(2), 451-474.

Cawley, J. (2007). The labor market impact of obesity. In Z. Acs and L. Alan (Ed.) *Obesity, business and public policy*, (pp. 76–88) Northampton, MA: Edward Elgar Publishing.

Cawley, J. and Meyerhoefer, C. (2012) The medical care costs of obesity: an instrumental variables approach. *Journal of Health Economics,* 31(1), 219-230.

Centers for Disease Control and Prevention. (2003). Healthy weight, overweight, and obesity among U.S. adults. www.cdc.gov/nchs/data/nhanes/databriefs/adultweight.pdf. Accessed 14 June 2017.

Colditz, G., Willett, W., Rotnitzky, A., and Manson, J. (1995). Weight gain as a risk factor for clinical diabetes mellitus in women. *Annals of Internal Medicine,* 122(7), 481-486.

Compton, W., Conway, K., Stinson, F., and Grant, B. (2006). Changes in the prevalence of major depression and comorbid substance use disorders in the United States between 1991–1992 and 2001–2002. *The American Journal of Psychiatry,* 163(12), 2141-2147.

Cutler, D., Glaeser, E., and Norberg, K. (2001). Explaining the rise in youth suicide. In J. Gruber (Ed.) *Risky behavior among youths: an economic analysis* (pp. 219-270). University of Chicago Press.

Ding, W., Lehrer, S., Rosenquist, J., and Audrain-McGovern, J. (2009). The impact of poor health on academic performance: New evidence using genetic markers. *Journal of Health Economics,* 28(3), 578-597.

Domingue, B. Belsky, D., Harris, K., Smolen, A., McQueen, M., and Boardman, J. (2014). Polygenic risk predicts obesity in both white and black young adults. *PloS one*, 9(7), e101596.

Flint, J. and Kendler, K. (2014). The genetics of major depression. *Neuron,* 81(3), 484-503.

Fletcher, J., and Lehrer, S. (2011). Genetic lotteries within families. *Journal of Health Economics*, 30(4), 647-659.

Fryar, C., Carroll, M., and Ogden, C. (2012). Prevalence of overweight, obesity, and extreme obesity among adults: United States, trends 1960–1962 through 2009–2010. NCHS Health E-Stat. http://www.cdc.gov/nchs/data/hestat/obesity_adult_09_10/obesity_adult_09_10.htm. Accessed March 18, 2017.

Granberg, E. (2011). Depression and obesity. In J. Cawley (Ed.). *The Oxford handbook of the social science of obesity* (pp. 329–349). Oxford University Press.

Gregory, C., and Ruhm, (2011). C. Where Does the Wage Penalty Bite? In M. Grossman and N. Mocan (Eds.) *Economic Aspects of Obesity*, (pp. 315-347). University of Chicago Press.

Gurung, R. and Prata, D. (2015). What is the impact of genome-wide supported risk variants for schizophrenia and bipolar disorder on brain structure and function? A systematic review. *Psychological Medicine*, 45(12), 2461-2480.

Hamermesh, D. and Soss, N. (1974). An economic theory of suicide. *Journal of Political Economy*, 82(1), 83-98.

Hubert, H., Feinleib, M., McNamara, P., and Castelli, W. (1983). Obesity as an independent risk factor for cardiovascular disease: a 26-year follow-up of participants in the Framingham Heart Study. *Circulation*, 67(5), 968-977.

Lawlor, D., Harbord, R., Sterne, J., Timpson, N., and Smith, G.D. (2008). Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Statistics in Medicine*, 27(8), 1133-1163.

Lépine, J., and Briley, M. (2011). The increasing burden of depression. *Neuropsychiatric Disease and Treatment*, 7(Suppl 1), 3-7.

Lewinsohn, P., Seeley, J., Roberts, R., and Allen, N. (1997). Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among community-residing older adults. *Psychology and Aging*, 12(2), 277-287.

Locke, A., Kahali, B., Berndt, S., Justice, A., Pers, T., Day, F., ... and Croteau-Chonka, D. (2015). Genetic studies of body mass index yield new insights for obesity biology. *Nature*, 518(7538), 197-206.

Luppino, F., de Wit, L., Bouvy, P., Stijnen, T., Cuijpers, P., Penninx, B., and Zitman, F. (2010). Overweight, obesity, and depression: a systematic review and meta-analysis of longitudinal studies. *Archives of General Psychiatry*, 67(3), 220-229.

Mark, T., Levit, K., Coffey, R., McKusick, D., Harwood, H., King, E., Bouchery, E., Genuardi, J., Vandivort-Warren, R., Buck, J., and Ryan, K. (2007). National expenditures for mental health services and substance abuse treatment, 1993–2003. Rockville, MD: Substance Abuse and Mental Health Services Administration.

McClellan, M., McNeil, B., and Newhouse, J. (1994). Does more intensive treatment of acute myocardial infarction in the elderly reduce mortality? Analysis using instrumental variables. *JAMA*, 272(11), 859-866.

National Longitudinal Study of Adolescent to Adult Health. (2015). Personal communication.

National Longitudinal Study of Adolescent to Adult Health. (2016). Personal communication.

National Longitudinal Study of Adolescent to Adult Health. (2017). Wave IV BMI Genetic Risk Score. www.cpc.unc.edu/projects/addhealth/documentation/restricteduse/datasets/GRS_BMI.pdf. Accessed 14 June 2017.

Nature.com. (2014). SNP. www.nature.com/scitable/definition/single-nucleotide-polymorphism-snp-295. Accessed 7 July 2017.

Norton, E., & Han, E. (2008). Genetic information, obesity, and labor market outcomes. *Health Economics*, 17(9), 1089-1104.

Parham, E. (1999). Promoting body size acceptance in weight management counseling. *Journal of the Academy of Nutrition and Dietetics*, 99(8), 920-925.

Pratt, L. and Brody, D. (2014). Depression and obesity in the US adult household population, 2005–2010. *NCHS Data Brief,* 167:1–8.

Radloff, L. (1977). The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement*, 1(3), 385-401.

Ripke, S., O'Dushlaine, C., Chambert, K., Moran, J., Kähler, A., Akterin, S., ... and Kim, Y. (2013). Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nature Genetics*, 45(10), 1150-1159.

Sabia, J. and Rees, D. (2015). Body weight, mental health capital, and academic achievement. *Review of Economics of the Household*, 13(3), 653-684.

Sankoh, A. J., Huque, M. F., & Dubey, S. D. (1997). Some comments on frequently used multiple endpoint adjustment methods in clinical trials. *Statistics in Medicine*, 16(22), 2529-2542.

Scott, K., McGee, M., Wells, J., and Browne, M. (2008). Obesity and mental disorders in the adult general population. *Journal of Psychosomatic Research*, 64(1), 97-105.

Solon, G., Haider, S., and Wooldridge, J. (2015). What are we weighting for? *Journal of Human Resources*, 50(2), 301-316.

Speliotes, E., Willer, C., Berndt, S., Monda, K., Thorleifsson, G., Jackson, A., ... and Randall, J. (2010). Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nature Genetics*, 42(11), 937-948.

Stock, J. and Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. Andrews and J. Stock (Eds.) *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, (pp. 80-108). Cambridge University Press.

Tukey, J., Ciminera, J., and Heyse, J. (1985). Testing the statistical certainty of a response to increasing doses of a drug. *Biometrics*, 41(1), 295-301.

Tyrrell, J., Jones, S., Beaumont, R., Astley, C., Lovell, R., Yaghootkar, H., Tuke, M., Ruth, K. S., Freathy, R., Hirschhorn, J., Wood, A., Murray, A., Weedon, M., and Frayling T. (2016). Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. *BMJ* 352: i582.

von Hinke, S., Smith, G., Lawlor, D., Propper, C., and Windmeijer, F. (2016). Genetic markers as instrumental variables. *Journal of Health Economics*, 45, 131-148.

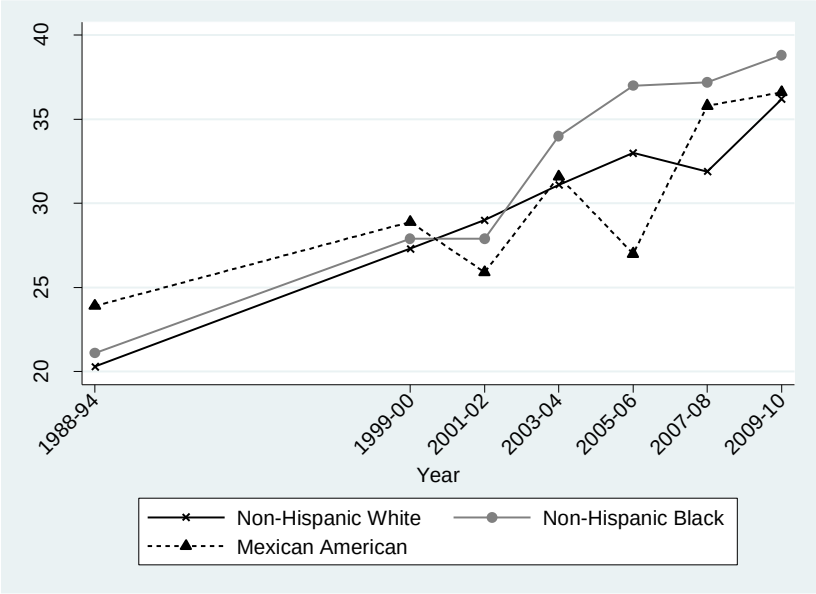
World Health Organization (2017). Health Topics: Obesity. http://www.who.int/topics/obesity/en/. Accessed March 18, 2017.

Zhao, G., Ford, E., Dhingra, S., Li, C., Strine, T., and Mokdad, A. (2009). Depression and anxiety among US adults: associations with body mass index. *International Journal of Obesity*, 33(2), 257-266.

Zhuo, X., Zhang, P., Barker, L., Albright, A., Thompson, T., and Gregg, E. (2014). The lifetime cost of diabetes and its implications for diabetes prevention. *Diabetes Care*, 37(9), 2557-2564.

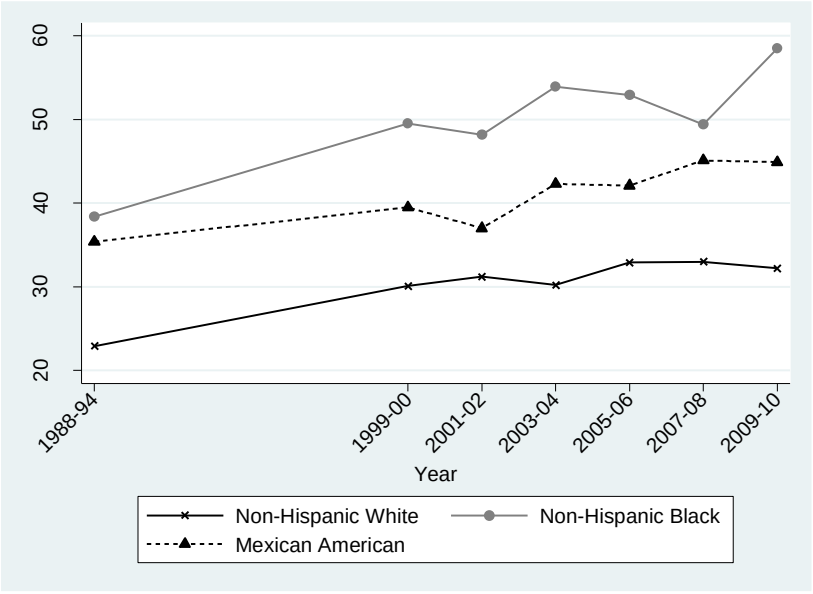
Appendix

Appendix Figure A1: Trends in obesity rates among adults by race, men



Source: Fryar et al., 2012. Data: NHANES III; Continuous NHANES (1999-00 through 2009-10)

Appendix Figure A2: Trends in obesity rates among adults by race, women



Source: Fryar et al., 2012. Data: NHANES III; Continuous NHANES (1999-00 through 2009-10)

Appendix Table A1: Summary Statistics and Data Availability by Wave

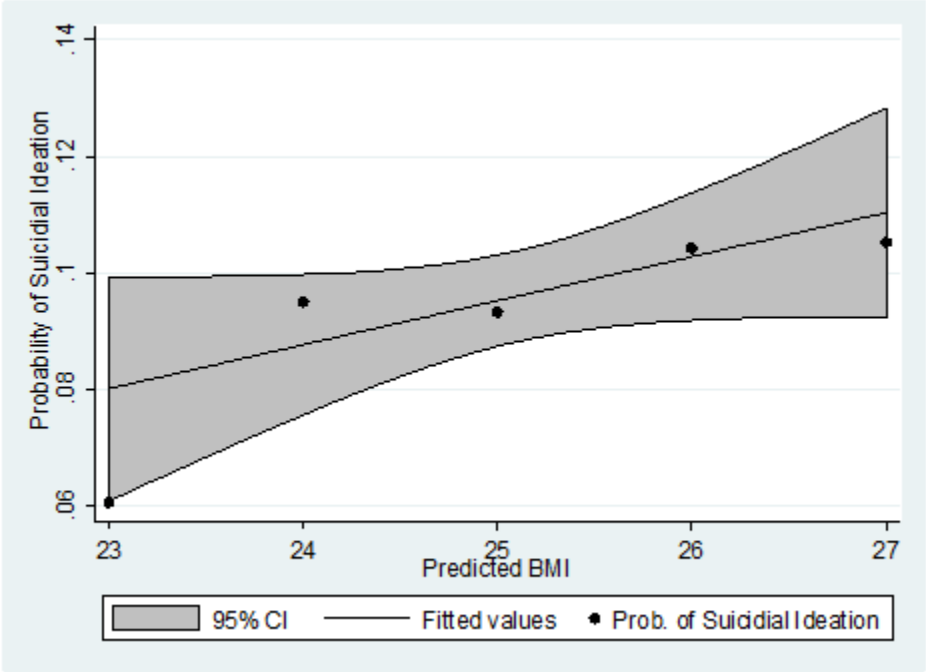
	Availability					
	Wave 1	Wave 2	Wave 3	Wave 4	Mean	(SD)
BMI Genetic Risk Score				X[1]	27.0	(3.4)
BMI	X[2]	X	X	X	25.6	(6.8)
Suicidal Ideation	X	X	X	X	0.097	(0.296)
Depression Index	X	X	X	X	0.0	(1.0)
Counseling Use	X	X	X	X	0.085	(0.279)
Self-Rated Unattractiveness			X	X	0.238	(0.426)
Interviewer-Rated Unattractiveness	X	X	X	X	0.529	(0.499)
Father Obese	X[3]				0.117	(0.321)
Mother Obese	X[3]				0.229	(0.420)
Father Ed. (HS Grad or Less)	X[3]				0.488	(0.500)
Mother Ed. (HS Grad or Less)	X[3]				0.556	(0.497)
Native Born	X[1]				0.961	(0.194)
Black	X[1]				0.351	(0.477)
White	X[1]				0.495	(0.500)
Hispanic	X[1]				0.108	(0.310)
Gender	X[1]				0.479	(0.500)
Age	X	X	X	X	21.1	(5.5)
Intelligence (Number Recall)				X[3]	4.065	(1.525)
Intelligence (Picture Vocab)	X[3]				0.097	(0.296)

[1] Time-invariant; [2] Self-reported; [3] Treating as time-invariant

Standard deviations in parentheses (SD)

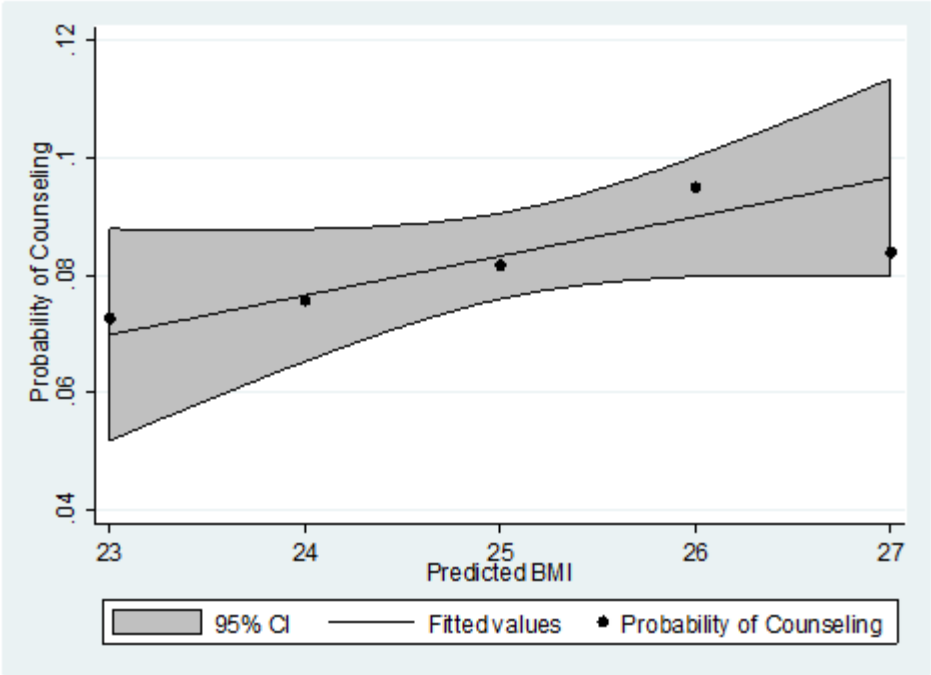
Data: The National Longitudinal Study of Adolescent to Adult Health

Appendix Figure A3: Relationship between Suicidal Ideation and Predicted BMI



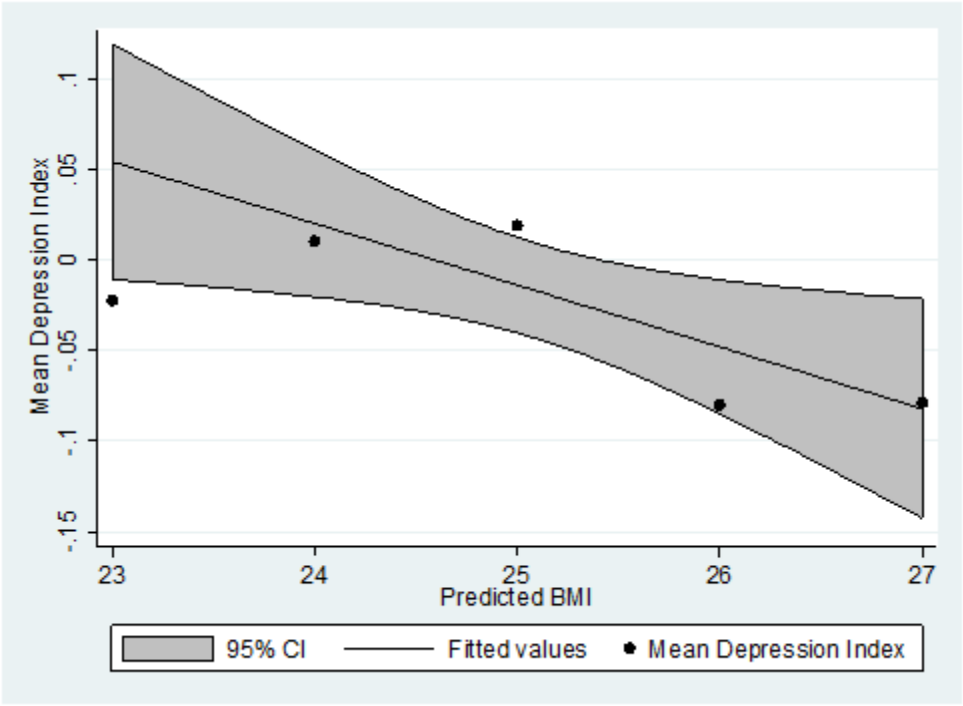
Data Source: The National Longitudinal Study of Adolescent to Adult Health
Predicted BMI range restricted due to small numbers of observations at extreme values

Appendix Figure A4: Relationship between Counseling and Predicted BMI



Data Source: The National Longitudinal Study of Adolescent to Adult Health
Predicted BMI range restricted due to small numbers of observations at extreme values

Appendix Figure A5: Relationship between Depression Index and Predicted BMI



Data Source: The National Longitudinal Study of Adolescent to Adult Health
Predicted BMI range restricted due to small numbers of observations at extreme values

Appendix Table A2: Balance Between High and Low BMI Risk

	Low BMI Risk	High BMI Risk	Difference
Parental Characteristics			
Parent Income ($1,000)	41.518	43.063	1.545
	(2.223)	(1.559)	p=0.581
Less than HS Mother	0.171	0.167	-0.004
	(0.013)	(0.014)	p=0.819
Less than HS Father	0.147	0.151	0.004
	(0.015)	(0.016)	p=0.863
HS Grad Mother	0.382	0.393	0.012
	(0.017)	(0.018)	p=0.639
HS Grad Father	0.336	0.348	0.012
	(0.02)	(0.021)	p=0.686
Some Post HS Mother	0.195	0.213	0.019
	(0.013)	(0.015)	p=0.362
Some Post HS Father	0.182	0.205	0.023
	(0.016)	(0.018)	p=0.344
College Mother	0.197	0.177	-0.02
	(0.014)	(0.014)	p=0.306
College Father	0.234	0.213	-0.021
	(0.018)	(0.018)	p=0.404
More Than College Mother	0.056	0.049	-0.007
	(0.008)	(0.008)	p=0.564
More Than College Father	0.102	0.085	-0.017
	(0.013)	(0.012)	p=0.348
Obese Mother	0.209	0.247	0.038*
	(0.014)	(0.017)	p=0.078
Obese Father	0.075	0.162	0.086***
	(0.009)	(0.014)	p<0.001
Alcoholism Mother	0.031	0.037	0.006
	(0.006)	(0.007)	p=0.523
Alcoholism Father	0.153	0.157	0.003
	(0.013)	(0.014)	p=0.861
Diabetes Mother	0.030	0.036	0.005
	(0.006)	(0.007)	p=0.568
Diabetes Father	0.055	0.063	0.008
	(0.008)	(0.01)	p=0.527

Continue next page

Data: The National Longitudinal Study of Adolescent to Adult Health; Wave 1 (2001-02)
*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE)

Appendix Table A2: Balance Between High and Low BMI Risk (cont)

	Low BMI Risk	High BMI Risk	Difference
Intelligence and Academic Achievement			
Picture Vocab Test	96.987	99.255	2.268***
	(0.472)	(0.520)	p=0.001
Number Recall	4.027	4.096	0.069
	(0.05)	(0.054)	p=0.350
Grade Point Average	2.768	2.695	-0.073*
	(0.024)	(0.028)	p=0.048
Less Than High School	0.111	0.101	-0.011
	(0.01)	(0.011)	p=0.480
High School	0.193	0.181	-0.012
	(0.013)	(0.014)	p=0.539
Some Post HS	0.363	0.387	0.024
	(0.016)	(0.018)	p=0.299
JR College	0.07	0.066	-0.004
	(0.008)	(0.009)	p=0.754
College	0.168	0.171	0.003
	(0.012)	(0.014)	p=0.866
More Than College	0.096	0.094	-0.001
	(0.01)	(0.011)	p=0.922
Demographics			
Native Born	0.942	0.982	0.041***
	(0.008)	(0.005)	p<0.001
White	0.407	0.6	0.193***
	(0.016)	(0.018)	p<0.001
Black	0.432	0.257	-0.175***
	(0.016)	(0.016)	p<0.001
Hispanic	0.100	0.119	0.019
	(0.01)	(0.012)	p=0.205

Continue next page

Data: The National Longitudinal Study of Adolescent to Adult Health; Wave 1 (2001-02)
*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE)

Appendix Table A2: Balance Between High and Low BMI Risk (cont)

	Low BMI Risk	High BMI Risk	Difference
	Body Measures		
BMI (self-reported) (Wave 1)	22.216	23.393	1.176***
	(0.141)	(0.188)	p<0.001
Obese (self-reported) (Wave 1)	0.055	0.11	0.055***
	(0.007)	(0.011)	p<0.001
Weight (lbs) (Wave 4)	182.186	195.457	13.271***
	(1.556)	(2.004)	p<0.001
Height (inches) (Wave 4)	66.917	67.076	0.159
	(0.132)	(0.142)	p=0.413
	Mental Health Outcomes		
Suicidal Ideation (Wave 1)	0.136	0.146	0.010
(past 12 months)	(0.011)	(0.013)	p=0.553
Counseling (Wave 1)	0.095	0.114	0.019
(past 12 months)	(0.009)	(0.011)	p=0.192
Depression Index (Wave 1)	6.372	6.220	-0.151
	(0.138)	(0.158)	p=0.469

Data: The National Longitudinal Study of Adolescent to Adult Health; Wave 1 (2001-02)
*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE)

Appendix Table A3: Association between BMI Risk Score and Predicted Outcomes; Black and White Pooled Sample

Suicidal Ideation	Counseling	Depression Index
0.0009***	0.0015***	0.0075***
(0.0002)	(0.0003)	(0.0026)
N=5517	N=5566	N=5559

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Observables included: Age, age squared, native born, wave, race/ethnicity, parent education, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Second stage also includes parent obesity

Appendix Table A4: Association between BMI Risk Score and Predicted Outcomes; White Sample

Suicidal Ideation	Counseling	Depression Index
0.0005*	0.0003	0.0003
(0.0003)	(0.0003)	(0.0029)
N=2898	N=2915	N=2912

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Observables included: Age, age squared, native born, wave, race/ethnicity, parent education, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Second stage also includes parent obesity

Appendix Table A5: Association between BMI Risk Score and Predicted Outcomes; Black Sample

Suicidal Ideation	Counseling	Depression Index
-0.0003	-0.0001	0.0007
(0.0005)	(0.0003)	(0.0049)
N=1840	N=1863	N=1860

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Observables included: Age, age squared, native born, wave, race/ethnicity, parent education, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Second stage also includes parent obesity

Appendix Table A6: Reduced Form Regressions - Effect of Genetic Risk Score on Mental Health Outcomes; Black and White Pooled Sample

	Suicidal Ideation	Counseling	Depression Index
Panel A: Males and Females Pooled			
Genetic Risk Score	0.00271*	0.000804	0.000583
	(0.00139)	(0.00122)	(0.00605)
	N=5517	N=5566	N=5559
Panel B: Males			
Genetic Risk Score	0.00190	0.00234	-0.00319
	(0.00170)	(0.00178)	(0.00825)
	N=2679	N=2708	N=2702
Panel C: Females			
Genetic Risk Score	0.00330	-0.000644	0.00412
	(0.00208)	(0.00190)	(0.00776)
	N=2838	N=2858	N=2857

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A7: Reduced Form Regressions - Effect of Genetic Risk Score on Mental Health Outcomes; White Sample

	Suicidal Ideation	Counseling	Depression Index
Panel A: Males and Females Pooled			
Genetic Risk Score	0.00340*	-0.000566	0.000485
	(0.00177)	(0.00163)	(0.00877)
	N=2898	N=2915	N=2912
Panel B: Males			
Genetic Risk Score	0.00212	0.000879	-0.0101
	(0.00256)	(0.00235)	(0.0131)
	N=1415	N=1425	N=1421
Panel C: Females			
Genetic Risk Score	0.00413*	-0.00272	0.00950
	(0.00246)	(0.00271)	(0.0100)
	N=1483	N=1490	N=1491

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A8: Reduced Form Regressions - Effect of Genetic Risk Score on Mental Health Outcomes; Black Sample

	Suicidal Ideation	Counseling	Depression Index
Panel A: Males and Females Pooled			
Genetic Risk Score	0.00203	0.00175	0.00685
	(0.00234)	(0.00221)	(0.0124)
	N=1840	N=1863	N=1860
Panel B: Males			
Genetic Risk Score	0.00118	0.00557	0.0119
	(0.00239)	(0.00358)	(0.0165)
	N=858	N=874	N=872
Panel C: Females			
Genetic Risk Score	0.00484	0.000802	0.00242
	(0.00331)	(0.00274)	(0.0163)
	N=982	N=989	N=988

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A9: Reduced Form Probit Regressions Margins - Effect of Genetic Risk Score on Mental Health Outcomes; Black and White Pooled Sample

	Suicidal Ideation	Counseling
Panel A: Males and Females Pooled		
Genetic Risk Score	0.00262*	0.000786
	(0.00135)	(0.00119)
	N=5517	N=5566
Panel B: Males		
Genetic Risk Score	0.00190	0.00235
	(0.00164)	(0.00169)
	N=2679	N=2708
Panel C: Females		
Genetic Risk Score	0.00337*	-0.000795
	(0.00205)	(0.00183)
	N=2838	N=2858

*p<0.1, **p<0.05, ***p<0.01
Standard errors in parentheses (SE); Standard errors clustered at the school level
Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A10: 2SLS Standard Errors by Clusters, Suicidal Ideation

Panel A: Males and Females Pooled			
	Full Sample	Whites	Blacks
BMI	(0.00469)	(0.00508)	(0.00657)
	[0.00472]	[0.00498]	[0.00666]
	{0.00469}	{0.00508}	{0.00657}
Panel B: Males			
	Full Sample	Whites	Blacks
BMI	(0.00548)	(0.00727)	(0.00685)
	[0.00547]	[0.00726]	[0.00663]
	{0.00548}	{0.00727}	{0.00685}
Panel C: Females			
	Full Sample	Whites	Blacks
BMI	(0.00703)	(0.00795)	(0.0100)
	[0.00710]	[0.00799]	[0.0106]
	{0.00703}	{0.00795}	{0.0100}

Standard errors clustered at the school level in parentheses (SE)
Standard errors clustered at the school and family level in square brackets [SE]
Standard errors clustered at the school and individual level in braces {SE}
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A11: 2SLS Standard Errors by Clusters, Depression Index

Panel A: Males and Females Pooled			
	Full Sample	Whites	Blacks
BMI	(0.0198)	(0.0258)	(0.0349)
	[0.0203]	[0.0264]	[0.0362]
	{0.0198}	{0.0258}	{0.0349}
Panel B: Males			
	Full Sample	Whites	Blacks
BMI	(0.0260)	(0.0381)	(0.0459)
	[0.0266]	[0.0390]	[0.0475]
	{0.0260}	{0.0381}	{0.0459}
Panel C: Females			
	Full Sample	Whites	Blacks
BMI	(0.0267)	(0.0327)	(0.0498)
	[0.0277]	[0.0339]	[0.0519]
	{0.0267}	{0.0327}	{0.0498}

Standard errors clustered at the school level in parentheses (SE)
Standard errors clustered at the school and family level in square brackets [SE]
Standard errors clustered at the school and individual level in braces {SE}
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A12: 2SLS Standard Errors by Clusters, Counseling Use

Panel A: Males and Females Pooled			
	Full Sample	Whites	Blacks
BMI	(0.00402)	(0.00478)	(0.00639)
	[0.00406]	[0.00480]	[0.00643]
	{0.00402}	{0.00478}	{0.00639}
Panel B: Males			
	Full Sample	Whites	Blacks
BMI	(0.00601)	(0.00678)	(0.0114)
	[0.00600]	[0.00672]	[0.0115]
	{0.00601}	{0.00678}	{0.0114}
Panel C: Females			
	Full Sample	Whites	Blacks
BMI	(0.00627)	(0.00885)	(0.00823)
	[0.00630]	[0.00896]	[0.00802]
	{0.00627}	{0.00885}	{0.00823}

Standard errors clustered at the school level in parentheses (SE)
Standard errors clustered at the school and family level in square brackets [SE]
Standard errors clustered at the school and individual level in braces {SE}
Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A13: 2SLS - Effect of BMI on Suicidal Ideation, High Depression Subsample and Full Sample (Mean=0.097)

	High Depression Index Subsample			***Full Sample***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0244**	0.0318***	0.00909	0.00876*	0.0103**	0.00490
	(0.00980)	(0.00965)	(0.0120)	(0.00469)	(0.00508)	(0.00657)
	N=2345	N=1059	N=901	N=5437	N=2857	N=1818
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0128	0.0162	0.00933	0.00546	0.00552	0.00335
	(0.0103)	(0.0147)	(0.0152)	(0.00548)	(0.00727)	(0.00685)
	N=1006	N=448	N=390	N=2648	N=1398	N=852
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0323**	0.0415***	0.0190	0.0116*	0.0140*	0.0136
	(0.0145)	(0.0141)	(0.0151)	(0.00703)	(0.00795)	(0.0100)
	N=1339	N=611	N=511	N=2789	N=1459	N=966

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

High depression subsample restricted to those with scores above the mean of 0

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health

Appendix Table A14: 2SLS - Effect of BMI on Suicidal Ideation, Adults and All Ages (Mean=0.097)

	Adults Only			***All Ages***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0111***	0.0174***	0.00251	0.00876*	0.0103**	0.00490
	(0.00432)	(0.00523)	(0.00615)	(0.00469)	(0.00508)	(0.00657)
	N=3252	N=1711	N=1072	N=5437	N=2857	N=1818
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00890	0.0121	0.00400	0.00546	0.00552	0.00335
	(0.00577)	(0.00860)	(0.00763)	(0.00548)	(0.00727)	(0.00685)
	N=1593	N=848	N=501	N=2648	N=1398	N=852
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0133**	0.0216**	0.00412	0.0116*	0.0140*	0.0136
	(0.00620)	(0.00916)	(0.00901)	(0.00703)	(0.00795)	(0.0100)
	N=1659	N=863	N=571	N=2789	N=1459	N=966

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A15: 2SLS - Effect of BMI on Depression Index, Adults and All Ages (Mean=0.0)

	Adults Only			***All Ages***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0118	0.000953	0.0214	0.00283	0.00267	0.0200
	(0.0184)	(0.0251)	(0.0397)	(0.0198)	(0.0258)	(0.0349)
	N=3280	N=1724	N=1083	N=5477	N=2871	N=1838
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00120	-0.0257	0.0332	-0.00903	-0.0259	0.0329
	(0.0272)	(0.0476)	(0.0461)	(0.0260)	(0.0381)	(0.0459)
	N=1612	N=856	N=509	N=2671	N=1404	N=866
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0238	0.0233	0.00666	0.0148	0.0308	0.0103
	(0.0259)	(0.0302)	(0.0598)	(0.0267)	(0.0327)	(0.0498)
	N=1668	N=868	N=574	N=2806	N=1467	N=972

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A16: 2SLS - Effect of BMI on Counseling Use, Adults and All Ages (Mean=0.085)

	Adults Only			***All Ages***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00255	0.000610	0.00258	0.00316	-0.000782	0.00493
	(0.00373)	(0.00491)	(0.00588)	(0.00402)	(0.00478)	(0.00639)
	N=3286	N=1726	N=1086	N=5484	N=2874	N=1841
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00872	0.00840	0.00950	0.00768	0.00286	0.0158
	(0.00541)	(0.00734)	(0.0104)	(0.00601)	(0.00678)	(0.0114)
	N=1616	N=858	N=511	N=2677	N=1408	N=868
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.00392	-0.0108	0.00102	-0.00158	-0.00749	0.00235
	(0.00627)	(0.00960)	(0.00886)	(0.00627)	(0.00885)	(0.00823)
	N=1670	N=868	N=575	N=2807	N=1466	N=973

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A17: 2SLS - Effect of BMI on Suicidal Ideation, by Instrument (Common Sample) (Mean=0.097)

	Sibling BMI			***BMI Genetic Risk Score***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00408	0.00436	-0.00618	0.0104*	0.00794	0.00790
	(0.00392)	(0.00425)	(0.0156)	(0.00563)	(0.00559)	(0.0115)
	N=3909	N=2353	N=1016	N=3909	N=2353	N=1016
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.00570	-0.00536	-0.0916	0.00602	0.00216	0.00365
	(0.00514)	(0.00501)	(0.388)	(0.00586)	(0.00644)	(0.00709)
	N=1884	N=1148	N=447	N=1884	N=1148	N=447
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0123**	0.0121**	-0.000205	0.0151	0.0120	0.0395
	(0.00546)	(0.00514)	(0.0150)	(0.0111)	(0.0103)	(0.0509)
	N=2025	N=1205	N=569	N=2025	N=1205	N=569

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

Sample restricted to respondents with BMI Genetic Risk Scores

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A18: 2SLS - Effect of BMI on Depression Index, by Instrument (Common Sample) (Mean=0.0)

	Sibling BMI			***BMI Genetic Risk Score***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00150	0.0108	-0.0589	-0.0118	-0.0254	0.0469
	(0.0145)	(0.0157)	(0.0655)	(0.0234)	(0.0285)	(0.0631)
	N=3937	N=2366	N=1026	N=3937	N=2366	N=1026
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.0157	-0.0297	0.218	-0.0350	-0.0560	0.0312
	(0.0223)	(0.0269)	(1.492)	(0.0279)	(0.0389)	(0.0367)
	N=1897	N=1153	N=452	N=1897	N=1153	N=452
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0162	0.0410*	-0.0571	0.0222	0.0119	0.0711
	(0.0198)	(0.0223)	(0.0519)	(0.0409)	(0.0413)	(0.160)
	N=2040	N=1213	N=574	N=2040	N=1213	N=574

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

Sample restricted to respondents with BMI Genetic Risk Scores

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A19: 2SLS - Effect of BMI on Counseling Use, by Instrument (Common Sample) (Mean=0.085)

	Sibling BMI			***BMI Genetic Risk Score***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00412	0.00500	0.00652	0.00593	-0.000911	0.00639
	(0.00324)	(0.00343)	(0.0179)	(0.00505)	(0.00560)	(0.0113)
	N=3943	N=2368	N=1029	N=3943	N=2368	N=1029
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.00212	0.00158	-0.110	0.00765	0.00380	0.00588
	(0.00411)	(0.00464)	(0.571)	(0.00606)	(0.00763)	(0.0109)
	N=1903	N=1157	N=454	N=1903	N=1157	N=454
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00792	0.00734	0.0113	0.00122	-0.00882	0.0108
	(0.00484)	(0.00522)	(0.0171)	(0.00893)	(0.0105)	(0.0261)
	N=2040	N=1211	N=575	N=2040	N=1211	N=575

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

Sample restricted to respondents with BMI Genetic Risk Scores

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A20: 2SLS - Effect of BMI on Suicidal Ideation, by Instrument (Full Sample) (Mean=0.097)

	Sibling BMI			***BMI Genetic Risk Score***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00297	0.00275	-0.000088	0.00876*	0.0103**	0.00490
	(0.00184)	(0.00237)	(0.00634)	(0.00469)	(0.00508)	(0.00657)
	N=44997	N=28109	N=8806	N=5437	N=2857	N=1818
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.00110	-0.00300	-0.00149	0.00546	0.00552	0.00335
	(0.00261)	(0.00313)	(0.0153)	(0.00548)	(0.00727)	(0.00685)
	N=21063	N=13423	N=3779	N=2648	N=1398	N=852
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00567**	0.00697**	0.000961	0.0116*	0.0140*	0.0136
	(0.00245)	(0.00347)	(0.00542)	(0.00703)	(0.00795)	(0.0100)
	N=23934	N=14686	N=5027	N=2789	N=1459	N=966

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health

Appendix Table A21: 2SLS - Effect of BMI on Depression Index, by Instrument (Full Sample) (Mean=0.0)

	Sibling BMI			***BMI Genetic Risk Score***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00368	0.00238	-0.0148	0.00283	0.00267	0.0200
	(0.00695)	(0.00834)	(0.0204)	(0.0198)	(0.0258)	(0.0349)
	N=45335	N=28284	N=8905	N=5477	N=2871	N=1838
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.00762	-0.00853	-0.0125	-0.00903	-0.0259	0.0329
	(0.0104)	(0.0128)	(0.0638)	(0.0260)	(0.0381)	(0.0459)
	N=21278	N=13531	N=3837	N=2671	N=1404	N=866
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.0111	0.0102	-0.0129	0.0148	0.0308	0.0103
	(0.00873)	(0.0104)	(0.0213)	(0.0267)	(0.0327)	(0.0498)
	N=24057	N=14753	N=5068	N=2806	N=1467	N=972

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A22: 2SLS - Effect of BMI on Counseling Use, by Instrument (Full Sample) (Mean=0.085)

	Sibling BMI			***BMI Genetic Risk Score***		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00110	0.00129	0.00483	0.00316	-0.000782	0.00493
	(0.00192)	(0.00224)	(0.00516)	(0.00402)	(0.00478)	(0.00639)
	N=45436	N=28345	N=8927	N=5484	N=2874	N=1841
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.00156	-0.00241	-0.00358	0.00768	0.00286	0.0158
	(0.00271)	(0.00384)	(0.0110)	(0.00601)	(0.00678)	(0.0114)
	N=21332	N=13566	N=3850	N=2677	N=1408	N=868
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.00286	0.00380	0.00832	-0.00158	-0.00749	0.00235
	(0.00248)	(0.00298)	(0.00577)	(0.00627)	(0.00885)	(0.00823)
	N=24104	N=14779	N=5077	N=2807	N=1466	N=973

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

Appendix Table A23: 2SLS - Effect of BMI on Health

	Poor Health *(Mean=0.005)*			***Diabetes*** *(Mean=0.021)*		
Panel A: Males and Females Pooled						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.000311	0.00121	-0.00278**	-0.00146	-0.00267	0.00127
	(0.00110)	(0.00143)	(0.00121)	(0.00269)	(0.00334)	(0.00375)
	N=5486	N=2876	N=1841	N=2676	N=1413	N=890
Panel B: Males						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	0.000834	0.000076	0.000666	-0.000655	-0.00327	-0.000277
	(0.000958)	(0.000908)	(0.000594)	(0.00402)	(0.00612)	(0.00472)
	N=2677	N=1408	N=868	N=1294	N=690	N=413
Panel C: Females						
	Full Sample	Whites	Blacks	Full Sample	Whites	Blacks
BMI	-0.000193	0.00208	-0.00647**	-0.00193	-0.00243	0.00467
	(0.00186)	(0.00251)	(0.00286)	(0.00381)	(0.00371)	(0.00649)
	N=2809	N=1468	N=973	N=1382	N=723	N=477

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

Standard errors in parentheses (SE); Standard errors clustered at the school level

Controls: Age, age squared, native born, wave, race/ethnicity, parent education, parent obesity, county unemployment rate, gender, intelligence (picture vocabulary and number recall)

Data: The National Longitudinal Study of Adolescent to Adult Health (Waves 1-4)

CHAPTER 3

The Long-Run Effects of Student Aid on Health: Leveraging Changes in Social Security Benefits

Abstract

Financial aid for college lowers the cost of higher education, and more affordable college increases college-going, reduces drop-out rates, and improves academic performance. However, there is little evidence on the long-run impact of aid on health. To test the causal effect of aid on weight and general health, I exploit a shock in Social Security benefits that occurred 1981. Minor children of retired or disabled Social Security beneficiaries as well as children with deceased parents are eligible for their own Social Security benefits, and before 1981 college-aged recipients could continue to receive this cash transfer conditional on college enrollment. Using a difference-in-differences strategy, I show that this program reduced women's body mass index and improved general health, but had no effect on men. I then investigate the effects of these benefits on education, which is a likely mechanism between the benefits and health. I find that financial aid improved educational attainment for beneficiaries. Consistent with the effects on health, the effects on education are concentrated in women.

1. Introduction

Education and health are two of the largest categories of spending in the United States. Health expenditures in 2014 were 17% of the gross domestic product ($3 trillion); the average person received $9,500 of medical care (National Center for Health Statistics, 2016). While spending on education is much smaller, Americans still spend $1.2 trillion on education annually, and $500 billion on post-secondary schooling (Snyder et al., 2016). Much of the total spending on college comes from student aid, with at least $125 billion of aid in 2016 (U.S. Department of Education, 2016). In addition to the significant annual cost, education and health are important factors in other long-run outcomes such as labor market success (Currie and Madrian, 1999; Card, 1999).

Empirical research provides compelling evidence that college aid increases college initiation (Angrist, 1993; Dynarski, 2000; Dynarski, 2002; Seftor and Turner, 2002) and improves college performance and completion (Bettinger, 2004; Scott-Clayton, 2011). College aid may also impact health since education and health are strongly related (Grossman, 2008). For instance, college graduates are ten percentage points less likely to be obese than those with no post-secondary education (Schoenborn et al., 2013). Most of the gradient in health behaviors occurs between secondary and post-secondary, but not within secondary school (Schoenborn et al., 2013).

Determining the causal effect of financial aid on health is complicated by the fact that aid is not randomly assigned; many other factors affect both aid receipt and health. One example is parental socio-economic status (SES). Lower parental SES increases availability of need-based aid, but lower SES is associated with worse health

outcomes (Currie and Goodman, 2010). Conversely, high parental SES is correlated with improved primary and secondary school performance (Sirin, 2005), which increases access to merit-based aid.

Investments in health and financial aid are costly but also have great potential to improve lives, so it is crucial to investigate a potential causal relationship between them. A protective effect of financial aid on health has important implications for both the private and social returns to this investment. While health and education both improve labor market outcomes (Currie and Madrian, 1999; Card, 1999), a person who is both healthy and educated might have additional labor market advantages. Improved health also has important externalities, because public insurance pays for a great deal of health care (CMS.gov, 2016).

In this study I examine how an exogenous change in federal aid for college affected weight and overall health.[46] Minor children of disabled and retired Social Security beneficiaries or of deceased parents who meet Social Security work requirements are eligible to receive their own benefits. Between 1965 and 1981, these children could continue to receive benefits until the age of 22 if they were enrolled in school, including college. The benefits for college-aged children serves the function of financial aid and lowers the out-of-pocket cost of higher education. This form of financial aid is unusual in that it takes the form of a conditional cash transfer (CCT), and the value depends on past parental earnings instead of the cost of education. In fact, the benefit was often larger than tuition at public four-year colleges (Dynarski,

[46] I focus on weight because alcohol consumption has a complicated and often non-monotonic association with education, while most smoking behavior and particularly initiation are determined before college-age (Schoenberg et al., 2013; U.S. Department of Health and Human Services, 2014).

2003). Due to the size of the benefit, when enrollment peaked in the late 1970s, over 800,000 students received benefits each year (DeWitt, 2001).

As originally shown by Dynarski (2003), the sudden termination of benefits for college-aged children in 1981 resulted in a large decrease in educational attainment among beneficiaries. Based on data from the National Longitudinal Survey of Youth 1979 (NLSY79) cohort, Dynarski (2003) identifies probable recipients of Social Security benefits based on the respondent having a deceased father. She also uses the panel nature of the NLSY79 to precisely determine which respondents graduated high school during the timeframe Social Security provided benefits to college students. Using a difference-in-differences approach, she compares the difference in education among benefit recipients before and after the policy changed to the difference in education among non-recipients. She finds this form of financial aid increased college going by 20 percentage points (Dynarski, 2003).

Using data from the National Health Interview Survey (NHIS) linked to Social Security Administration (SSA) records, I exploit the variation in Social Security benefits to determine the causal effect of financial aid on BMI and overall health. My empirical approach is a difference-in-differences strategy similar to the approach Dynarski (2003) used. This model compares changes in health among beneficiaries to changes among non-beneficiaries. Any differential change in health is due to the change in policy if the difference-in-differences assumptions are met. These assumptions are that outcomes for beneficiaries and non-beneficiaries would have parallel trends in the absence of policy variation, and there are no contemporaneous shocks that differentially affect beneficiaries and non-beneficiaries. I also use the same

difference-in-differences strategy to examine the effect of this policy on educational attainment.

The main contribution of this paper is leveraging exogenous variation in Social Security benefits to test for the effects of financial aid on long-run health motivated by Dynarski's (2003) work on the effects of aid on education. I am the first to use changes in Social Security benefits for college-aged children to examine the effect of aid on health. Since much of the health behavior gradient occurs after high school (Schoenborn et al., 2013), changes in college-going is a plausible mechanism between financial aid and health. To investigate this potential mechanism, I also replicate Dynarski's (2003) analysis of this policy on education using a separate data set, the NHIS linked to Social Security records. In addition to increased statistical power from a larger data set, I use actual measures of Social Security receipt as opposed to self-reported paternal death as a proxy, so I also capture children of disabled and retired parents. However, I have less precise information on the timing of high school completion and less granular measures of educational attainment. If financial aid only impacts health through changes in educational attainment, then the effect of aid on health can be interpreted as reduced form intent-to-treat effects of education on health.

I find statistically significant and economically meaningful effects of Social Security benefits on educational attainment. My point estimates are smaller than Dynarski's (2003), but they fall within Dynarski's 95% confidence interval. I also find that the effect on education is highly concentrated among women, perhaps due to men having other sources to finance higher education, such as the GI Bill. In my main analysis, I find that financial aid reduces BMI by 0.8 units and improves the

probability of very good self-reported health by 3 percentage points for women. I find no effect of Social Security benefits on men's health. Since financial aid improves education and health for women, but has little-to-no effect on men, education likely is a primary pathway between aid and health.

This paper proceeds as follows: Section 2 reviews the existing literature. I describe the policy background and institutional details in Section 3. In Section 4, I discuss the data and method. Empirical results are presented in section 5. I conclude with a discussion in Section 6.

2. Existing Literature

This study relates to the literatures on college financial aid and conditional cash transfers, as well as the literature on the relationship between education and health. The theoretical motivation for financial aid increasing educational attainment rests on the assumption that potential college students have downward sloping demand curves for education. Since financial aid decreases the out-of-pocket cost for college, demand will increase. The most relevant research is Dynarski's (2003) study on the effects of Social Security benefits for college students. She finds that these benefits caused a large increase in college-going, however there is little information on the long-run effects of this aid on health. I contribute to this literature by examining if financial aid improves health as well as replicating Dynarski's (2003) results using a larger data set with more precise measures of treatment.

There are two other main categories of financial aid: need-based aid and merit-based aid. Pell Grants are a federal program that provides need-based aid to low-

income students. The primary criterion for receiving a Pell Grant for undergraduate education is financial need, and recipients are not required to repay the Pell Grant Program (U.S. Department of Education, 2015a). Pell Grants increase college initiation as well as improve college retention after the first year (Bettinger, 2004; Seftor and Turner, 2002). Merit-based financial aid programs are often granted at the state level, such as Georgia HOPE and West Virginia PROMISE. These programs benefit high achieving students, and receipt is not based on income. As with need-based aid, both Georgia HOPE and West Virginia PROMISE improve educational attainment (Dynarski, 2000; Scott-Clayton, 2011).

Due to the design of financial aid via Social Security benefits, this study also relates to the literature on conditional cash transfers (CCT). Since Social Security benefits for college students are a continuation of benefits for minors, these benefits are not a function of education costs, but only of parents' earning history (DeWitt, 2001; Dynarski, 2003). The fact that benefit size is independent from costs differs from need- and merit-based aid, so it can be viewed as a cash transfer conditional on college enrollment. The literature on CCTs focuses on the effect of means-tested transfers on parental investment in child health and education primarily in developing countries (see Fiszbein and Shady (2009) for a review of the literature). I contribute to this literature by examining a program that is distinct from much of the existing literature on CCTs: these benefits more generous in terms of benefit size and they are not means-tested.

Several programs in Latin America provide cash transfers to low-income families conditional on investment in child education, and these programs generally improve

school attendance (Attanasio, Fitzsimmons, and Gómez, 2005; Shady and Araujo, 2008; Schultz, 2004). CCTs also increase education participation in Asia and the United Kingdom (Chaudhury and Parajuli, 2008; Middleton et al., 2003). Conditional cash transfers improve recipients' health, including decreasing the probability of child mortality, increased height-for-age, and reduced prevalence of anemia (Fernald et al., 2008; Gertler, 2004; Rasella et al., 2013).

Since financial aid increases educational attainment, and education is highly correlated with health, education is an important potential mechanism between aid and health. Grossman (1972) first developed the theoretical framework for the relationship between human capital (education) and health capital. More education incentivizes people to be healthier because education leads to higher wages. With higher wages, healthy time is more valuable and can be more efficiently converted into other utility-increasing goods and services. Education can also improve how efficiently people produce health. Grossman discusses two forms of efficiency that explain how education leads to better health. Education informs people or helps them inform themselves about what is healthy, so they allocate resources to more effective health inputs (allocative efficiency). Given the same inputs, higher levels of human capital also result in better health: educated people can turn the same set of inputs into a greater quantity of good health (productive efficiency).

Though several studies empirically examine the relationship between education and health, the evidence is mixed and questions remain about the validity of the research designs. Studies fall into two main categories: those that use compulsory schooling laws to examine the effect of one more year of secondary schooling, and

those that use changes to the cost or benefit of attending college. I focus on the second category because it more closely relates to the effect of college financial aid.

De Walque (2007) and Grimard and Parent (2007) use the Vietnam draft as an instrument for college attendance. Conceptually, draft risk increased the benefit of higher education for men as college served as a method to avoid conscription. Both studies find that education reduces smoking. One concern is that over 20% of certain birth cohorts were drafted (De Walque, 2007), so men who avoided military service born in the mid-1940s faced different opportunities than men who did not serve from other birth cohorts. The Vietnam draft may also have influenced health through other pathways, such as marrying and having children in order to obtain a "hardship deferment" from military service (Selective Service System, 2016a; Selective Service System, 2016b). The effects on family formation could in turn drive differences in health (Gardner and Oswald, 2004).

Currie and Moretti (2003) leverage newly opened institutions of higher education in a woman's county to instrument for education's effect on behaviors related to birth outcomes. The intuition is that if colleges are closer to women, women are more likely to attend college. Results show that among mothers, education improves health-related behaviors such as smoking and use of prenatal care. One concern with this strategy is that the presence of colleges can change the number and quality of potential marriage partners, which in turn can affect health though a pathway other than women's education.

3. Policy Background

Children of deceased, disabled, and retired Social Security-eligible parents are also eligible for Social Security benefits. The benefit for children is meant to replace the financial support traditionally provided by parents, so children must be unmarried as well as either under 18 years old or 18-19 years old attending primary or secondary school (DeWitt, 2001). These criteria have been in place since 1981 and were the standard before 1965. Between 1965 and 1981, there was an additional group eligible for benefits: children up to age 22 who were enrolled in post-secondary education (Appendix Table A1).

During the 1965-1981 period, these benefits provided a major financial incentive for Social Security recipients to attend college. Social Security payments to children were quite large (1980 average $5,700 in real 2000 dollars), especially compared to tuition at a public university ($1,700 in real 2000 dollars), and covered most of the tuition at the average private university ($7,600 in real 2000 dollars) (Appendix Table A2). These benefits were more generous than other forms of financial aid such as Pell Grants, which in the 1980-81 school year paid a maximum of $1,800 ($3,500 in real 2000 dollars). The benefit differs from other financial aid in that the dollar amount is not dependent on college costs, but only on the earnings history of the recipient's parents. There is also little or no crowding-out of other aid, because children are unlikely to earn enough to pay taxes on these Social Security benefits and untaxed Social Security benefits are not included in the Free Application for Federal Student Aid (FAFSA) accounting of income (U.S. Department of Education, 2015b; Social Security Administration, 2016b).

As one might expect for such a generous program, many eligible children enrolled. One reason for the high take-up of this program for college-enrolled children was the relative ease of access; since a recipient could continue to receive Social Security benefits they had already enrolled in, there was little marginal cost in terms of time and effort to receive benefits during college. In contrast, need-based aid requires hours of collecting documentation and filling out paperwork.[47] In the first year college students were eligible, 200,000 non-minor students (including secondary students) were covered by Social Security. Enrollment peaked at 865,000 in 1977 and stayed relatively flat until 1981.[48]

Starting in the mid-1970s, social and political pressure to end benefits to college-age children began to build. DeWitt (2001) points to three issues with these benefits that could have contributed to their termination. First, many recipients failed to notify the Social Security Administration when they stopped being full-time students due to reduction in course-load or completely dropping out of college. In many cases, benefits were mistakenly paid to individuals who were no longer full-time students. Second, these benefits were quite expensive during a time when the economy was struggling, which put additional stress on the Social Security program. Third, there was a conceptual shift in the rationale for children's benefits. While child Social Security benefits were meant to replace the income of a parent, policymakers started to view benefits for college students as financial aid. Since other forms of financial aid

[47] See Dynarksi and Scott-Clayton (2006) for discussion of the effects of student aid complexity on aid take-up, and Kleven and Kopczuk (2011) for a discussion of other social benefits.
[48] For comparison, there were 8.6 million total college students in 1970 and 12.1 million in 1980 (Snyder and Dillow, 2014).

were available, some members of Congress argued that Social Security did not need to provide these funds.

Both the Ford and Carter administrations attempted to slowly end Social Security benefits for college students, however legislation ending these benefits did not pass until the Reagan presidency (DeWitt, 2001). Benefits were scheduled to end for college-aged recipients in April 1985, and the program was cut-sharply during the phase-out period. There were no new beneficiaries after May 1982, and payments to existing beneficiaries were reduced by 25% annually. Furthermore, Social Security only made payments during the school year, instead of year-round.

As seen in Figure 1, enrollment for children over 18 dropped precipitously during the phase-out period. Enrollment plateaued at 100,000 in 1985; the remaining beneficiaries are 18 year-olds still enrolled in primary or secondary school (DeWitt, 2001). Dynarski (2003) estimates that the termination of Social Security's payments to college students caused a 22 percentage point decrease in the probability of college attendance among eligible children.

4. Data and Method

4.1. Data

I use restricted data from the 1997 through 2005 National Health Interview Survey (NHIS) linked to Social Security Administration records. The NHIS is a large and nationally representative survey with approximately 90,000 respondents per year (CDC.gov, 2017). The NHIS data contain a wide range of information, including data on education, health, and demographics. For this study, the most important variables

from the NHIS are educational attainment and the outcome variables: body mass index (BMI) and general self-described health. Additionally, the NHIS has date of birth, so I can determine the approximate year respondents graduated from high school.

The Social Security Administration records contain month-by-month information on benefit type. Using this information, I can determine if a NHIS respondent received benefits for minors before her 18th birthday; this is the group who is eligible for college benefits between 1965 and 1981. By interacting a variable indicating receipt of minor benefits with a dummy variable of graduating high school in years Social Security covered college students, I construct a variable indicating eligibility for benefits in college.

I use NHIS sample weights and cluster standard errors at the primary sampling unit to account for the sampling design. Respondents were born between 1957 and 1981, and so graduated high school between the mid-1970s and the late 1990s. I exclude respondents who likely graduated high school before 1974. This group is excluded to ensure that the effects of the Vietnam War and the draft do not contaminate my analysis. Health information was collected between 1997 and 2005. The oldest person in the sample is 48 years old. I exclude respondents under 23 years of age to increase the probability of education completion. The last cohort with access to Social Security benefits during college graduated high school in 1981, so the marginal respondent was born in 1964. The effects of aid on health are long-run, because the marginal cohort is between 33 and 41 years old when they provided health information, one-to-two decades after they would have received financial aid for college.

4.2. Method

Main Analysis – Effect of Social Security Benefits on Health

The empirical strategy for estimating the effect of aid on health is a difference-in-differences. The estimating equation is:

$$Health_i = \beta_0 + \beta_1(SS_i * Policy_i) + \beta_2 SS_i + \boldsymbol{\beta_c}\mathbf{1}(Cohort_i) + \boldsymbol{\beta_y}\mathbf{1}(Year_i) + \boldsymbol{\beta_x}\boldsymbol{X}_i + \epsilon_i. \quad (2)$$

The variable $Health_i$ is a measure of health, specifically BMI or general self-described health. The variable SS_i is an indicator variable that equals 1 if the respondent received Social Security benefits before age 18, and 0 if not. The variable $Policy_i$ is an indicator variable of whether the respondent likely graduated high school in a year that college benefits could be received; this variable equals 1 if the respondent graduated in 1981 or earlier, and equals 0 if she graduated after 1981. The independent variable of interest is the interaction of SS_i and $Policy_i$, which equals 1 if the respondent is a Social Security recipient and she graduated high school in 1981 or earlier. The coefficient of interest is β_1, which is the change in health caused by being eligible for Social Security benefits during high school graduation. I include high school graduation cohort and survey year fixed effects, $\mathbf{1}(Cohort_i)$ and $\mathbf{1}(Year_i)$. I also control for a vector of demographics, $\boldsymbol{X_i}$, and the error term is γ_i.

The identifying assumptions of this model are common trends and no contemporaneous shocks in outcomes that differentially affect treated and control

groups. The common trends assumption is that in the absence of the change to Social Security rules, recipients of benefits and non-recipients would have common trends in health. I will present empirical evidence on common pre-trends in the following section. In terms of contemporaneous shocks, I show that characteristics that should not be affected by education, such as race, trend smoothly through the year of the policy shock. Under this assumption, controlling for high school graduation cohort and receipt of Social Security will isolate the variation in health caused by changes in college aid availability.

Potential Mechanism – Effect of Social Security Benefits on Education

The research design used in this analysis is a difference-in-differences. The empirical strategy is motivated by Dynarski's 2003 research, which showed that variation in Social Security benefits can have a large impact on college attendance. The estimating equation is:

$$Edu_i = \alpha_0 + \alpha_1(SS_i * Policy_i) + \alpha_2 SS_i + \boldsymbol{\alpha_c}\mathbf{1}(Cohort_i) + \boldsymbol{\alpha_y}\mathbf{1}(Year_i) + \boldsymbol{\alpha_x}\boldsymbol{X}_i + \gamma_i \quad . \tag{1}$$

The outcome, Edu_i, is a binary measure of educational attainment: high school graduation, any college, or college graduation. The coefficient of interest is α_1, which is the effect of Social Security benefits on the probability of achieving a certain level of education. Equation (1) is similar to Dynarski's (2003) equation (2), but there are important distinctions. I control for high school graduation cohort and year fixed effects, while she controls for a direct binary measure of high school graduation before

the policy is terminated. Additionally, I have a direct measure of Social Security benefit receipt, while she uses death of the respondent's father as a proxy of benefits.

The identifying assumptions of this model are: 1) treated and control groups have parallel trends in outcomes, and 2) there are no contemporaneous shocks. While I cannot observe the post-period trends in the absence of the policy change, I provide graphical evidence that recipients of child Social Security benefits and non-recipients have similar pre-trends in educational attainment. Under these assumptions, including dummy variables for high school graduation cohort and receiving Social Security benefits as a minor will allow me to identify the variation in education caused by changes in Social Security benefits.

If two additional criteria are met, then the effects of aid on health from the main analysis can be interpreted as reduced form, intent-to-treat effects of education on health. The first requirement is that financial aid affects education, and the second requirement is that education is the only pathway between financial aid and health. The first criterion is addressed by testing the effect of aid on health. The main concern regarding the second requirement is that recipients can use the benefits to directly improve health. However, the marginal respondent is reporting health more than a decade after receiving Social Security benefits, so the direct effect of benefits on health would need to be long-lasting. The net effect of receiving financial aid income during college age is ambiguous because recipients forwent earnings in the labor market. While Social Security benefits for college students were generous compared to other forms of financial aid, the potential lost wages are high; the average earnings among high school graduates was $34,500 for males and $12,000 for females in 1980

measured in constant 2000 dollars (Snyder, 1993).

5. Results

Main Analysis – Effect of Social Security Benefits on Health

I present the estimated effect of financial aid on BMI and general health in Figure 2 and Figure 3 and in Table 1. Estimates for the effect of education on BMI are in Figure 2 and Table 1 Column 1. In the pre-period of Figure 2, Social Security beneficiaries' BMIs are about half of a unit larger than non- beneficiaries', and both groups' BMIs move downward together.[49] However, once Social Security benefits for college stop, the two groups' BMIs diverge; non-beneficiaries continue on their slight downward path, while beneficiaries stop decreasing and increase after benefit termination. The difference-in-differences estimates in Table 1 Column 1 are consistent with Figure 2 and indicate that Social Security benefits for college-aged students decreased BMI about 0.465 units. For a 5'7" tall person, this is approximately a three pound or 0.8 of standard deviation decrease in weight. The effect of aid on BMI loads on women; the effect for women is a statistically significant 0.8 unit decrease in BMI and a much smaller not significant effect for men. The 0.8 decrease in BMI among women is approximately five pounds for a 5'7" tall person.

Results for the effect of aid on general health are presented in Figure 3 and Table 1 Column 2. In the pre-period of Figure 3, Social Security beneficiaries' are four percentage points less likely than non-beneficiaries' to report very good heath, and

[49] While the overall trend in BMI nationally is increasing, BMI is decreasing with cohort year in this graph. The reason is that respondents are all providing BMI information in 1997-2005, and so later cohorts are younger during this time period than earlier cohorts.

both groups have a similar trend. Once college-aged benefits cease, the two trends separate. From Table 1 Column 2, Social Security benefits for college-aged students caused a three percentage point or 5% increase in women reporting very good general health; the point estimate for men is smaller in magnitude, the opposite sign, and not statistically significant. The estimate for pooled genders is close to zero, because the opposite signs on the effects for men and women cancel each other out.

Potential Mechanism – Effect of Social Security Benefits on Education

The trend in attending college for recipients of Social Security benefits and non-recipients are in Figure 4. Both groups have a similar trend for cohorts that graduated high school before benefits for college students terminated, though recipients have lower rates of college going. When Social Security benefits for college students ended in the early 1980s, college going drops significantly for beneficiaries, though it does temporarily rebound some in the mid-1980s. However, the difference between the two groups largely persist in the post period. In Table 2, the effect of Social Security benefits on entering college is five percentage points and very precisely estimated, which is approximately a 10% effect on a mean of approximately 50%.

I also find that benefits increase college graduation rates by three percentage points. The benefits of graduating high school are greater if a student has financial access to college, and so high school graduation rates increase by two percentage points. When the analysis is stratified by gender, it is clear that Social Security benefits primarily affect women, with much smaller and often not significant effects on men. To confirm the heterogeneity by gender, I re-analyzed the data used in

Dynarski's (2003) analysis, the National Longitudinal Survey of Youth, 1979. Using these data, I also find that the effect of Social Security benefits for college-aged children is concentrated among women (Appendix Table A3). The fact that effects on education load on women is unsurprising as males in the early 1980s might have had access to other sources of educational funding, such as military service and the GI Bill. Even today, only about 15% of active military personnel and 18% of GI Bill beneficiaries are women (U.S. Department of Veterans Affairs, 2015; Parker et al., 2017).

As mentioned in the Methods Section, if financial aid impacts education and education is the only mechanism through which aid affects health, then financial aid is a valid instrument for education. I have shown that financial aid strongly impacts education, particularly for college initiation. The F-statistic of obtaining at least some college is 24, which is above the traditional cut-off of 10 (Stock and Yogo, 2005). The fact that the effects of aid on both education and health load on women contributes to the plausibility that education is the only pathway between aid and health, and Social Security benefits for college students do not independently improve long-run health. Another reason to doubt that Social Security benefits directly impact later-life health is that benefits are likely smaller than forgone earnings (Snyder, 1993).

Additional Results

Since I do not know the actual year of high school graduation, I must impute graduation year based on birth date. In the main analysis, I use births before the end of 1964 (turn 17 years old in 1981) as the last group eligible. The reason is that many

states allowed children to start kindergarten at the age of four as long as the child's birthday was before December (Education Commission of the States, 2011). Hawaii, Illinois, Michigan, New York, Rhode Island, and Wisconsin had birthday thresholds in December, and many other states had thresholds in November. However, many high school graduates turned 18 in 1981, so I use births in September 1963 as the cut-off in a robustness check. Results are presented in Column 1, Appendix Tables A4-A5 and are consistent with the main results (Column 3).

A deceased parent is the main source of eligibility for children receiving Social Security benefits (Dynarski, 2003). However, some children are also eligible due to a disabled or retired parent. In the early 1980s, many disabled adults were removed from Social Security (Knight-Ridder Newspapers, 1989). Since such a change in Social Security enrollment of adults could impact children's benefits, I also perform a robustness check excluding children with disabled parents. The remaining children in the sample have deceased or retired parents who were unaffected by the change for the disabled population. The results are presented in Column 2, Appendix Tables A4-A5, and the estimates are robust to dropping children with disabled parents.

Other characteristics that should be unaffected by financial aid trend smoothly through the termination of Social Security benefits to college students. In Appendix Figures A1 and A2, there are not shocks in race or ethnicity around 1981, and so changes in demographics in the treated or control groups are not driving results. While financial aid could impact BMI through changes in weight, height should not be affected by benefits that occur in the late teens and early 20s. In Appendix Figure A3, there is no discernible differential shift in height that occurs in the early 1980s.

However, in Appendix Figure A4, the weight of recipients and non-recipients trend together in the pre-period, but the gap between them widens after termination of benefits to college students.

6. Conclusion

This study contributes to the literature on the long-run effects of financial aid as well as the literature on the relationship between education and health. I use a shock to Social Security benefits as a source of exogenous variation in financial aid. I find that financial aid lowers long-run BMI and improves overall health for women, but I do not find effects on men's health. Based on the strong effect of aid on education, I show education is a mechanism by which aid improves health. I also find that the improvements in education are primarily among women, with little-to-no effects on men. The strong effects of aid on health, and aid affecting only women's education and health, suggests that aid is a valid instrument for education. If financial aid is a valid instrument of education, then my findings indicate that education improves health.

There is great potential to use Social Security benefits for college-aged students to investigate the effects of aid or as an instrument for education, but there are significant data requirements. Since relatively few minors receive Social Security benefits, precise estimates require large datasets. Another obstacle is accurate measures of eligibility for college-aged benefits; determining whether someone received Social Security as a minor is not possible with most surveys. One ideal area for future research is the effect of these benefits on labor market outcomes using administrative

Social Security data for both program eligibility and labor force participation. Using such data, researchers could observe the effect of aid on employment and earnings at one point in time as well as the full life-cycle of earnings.

These findings have important policy implications. First, the returns to financial aid are greater than improved educational outcomes. If increased college attendance was the only benefit that individuals and society consider, then we would be under-investing in financial aid. Improved health has direct impacts on people's utility as well as a positive externality, because all members in an insurance risk pool benefit from improved health. Second, financial aid does increase educational attainment, and education is a mechanism through which aid improves

REFERENCE

Albouy, Valerie, and Laurent Lequien. "Does compulsory education lower mortality?" *Journal of Health Economics* 28.1 (2009): 155-168.

Angrist, Joshua D. "The effect of veterans benefits on education and earnings." *Industrial and Labor Relations Review* 46.4 (1993): 637-652.

Attanasio, O., Emla Fitzsimons, Ana Gomez, Diana Lopez, C. Meghir, and Alice Mesnard (2006). Child Education and Work Choices in the Presence of a Conditional Cash Transfer Programme in Rural Colombia." *Working Paper W06/13,* Institute of Fiscal Studies, London.

Bettinger, E. (2004). How financial aid affects persistence. In C. Hoxby (Ed.), *College choices: The economics of where to go, when to go and how to pay for it* (pp. 207-237). Chicago: The University of Chicago Press

Card, D. (1999). The causal effect of education on earnings. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (1801-1863). North Holland, Amsterdam.

CDC.gov. “About the National Health Interview Survey.” (2017). https://www.cdc.gov/nchs/nhis/about_nhis.htm. Accessed 24 November 2017.

Chaudhury, Nazmul, and Dilip Parajuli. "Conditional cash transfers and female schooling: the impact of the female school stipend programme on public school enrolments in Punjab, Pakistan." *Applied Economics* 42.28 (2010): 3565-3583.

CMS.gov. “National Health Expenditures 2016 Highlights.” (2016). https://www.cms.gov/ Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData /downloads/highlights.pdf. Accessed 19 April 2018.

Clark, Damon, and Heather Roayer. "The effect of education on adult mortality and health: Evidence from Britain." *The American Economic Review* 103.6 (2013): 2087-2120.

Currie, Janet and Brigitte Madrian. (1999) Health, health insurance and the labor market. In O. Ashenfelter and D. Card (Eds.), *Handbook of Labor Economics* (3309-3407). North Holland, Amsterdam.

Currie, Janet, and Enrico Moretti. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *The Quarterly Journal of Economics* (2003): 1495-1532.

Currie, Janet, and Joshua Goodman. (2010). Parental Socioeconomic Status, Child Health and Human Capital. In D. Brewer and P. McEwan (Eds.), *Economics of Education* (156-162). Oxford, UK and San Diego: Elsevier.

De Walque, Damien. "Does education affect smoking behaviors?: Evidence using the Vietnam draft as an instrument for college education." *Journal of Health Economics* 26.5 (2007): 877-895.

DeWitt, Larry. “Research Note #11: The History of Social Security "Student" Benefits.” (2001). https://www.ssa.gov/history/studentbenefit.html. Accessed 11 August 2016.

Dynarski, Susan. "Hope for whom? Financial aid for the middle class and its impact on college attendance." *National Tax Journal* 53.3 (2000): 629-661.

Dynarski, Susan. "The behavioral and distributional implications of aid for college." American Economic Review 92.2 (2002): 279-285.

Dynarski, Susan. "Does Aid Matter? Measuring the Effect of Student Aid on College Attendance and Completion." *American Economic Review* 93.1 (2003): 279-288.

Dynarski, Susan and Judith E. Scott-Clayton. "The Cost of Complexity in Federal Student Aid: Lessons from Optimal Tax Theory and Behavioral Economics." *National Tax Journal* (2006): 319-356.

Fernald, Lia CH, Paul J. Gertler, and Lynnette M. Neufeld. "Role of cash in conditional cash transfer programmes for child health, growth, and development: an analysis of Mexico's Oportunidades." *The Lancet* 371.9615 (2008): 828-837.

Fiszbein, Ariel, and Norbert R. Schady. *Conditional cash transfers: reducing present and future poverty*. Policy Research Report, World Bank Publications, 2009.

Education Commission of the States. “Kindergarten Entrance Ages: A 35 Year Trend Analysis.” (2011). http://www.ecs.org/clearinghouse/93/61/9361.pdf. Accessed 4 March 2018.

Gardner, Jonathan, and Andrew Oswald. "How is mortality affected by money, marriage, and stress?" *Journal of Health Economics* 23.6 (2004): 1181-1207.

Geiss, Linda S., et al. "Prevalence and incidence trends for diagnosed diabetes among adults aged 20 to 79 years, United States, 1980-2012." *Journal of the American Medical Association* 312.12 (2014): 1218-1226.

Gertler, Paul. "Do conditional cash transfers improve child health? Evidence from PROGRESA's control randomized experiment." *American Economic Review* 94.2 (2004): 336-341.

Grimard, Franque, and Daniel Parent. "Education and smoking: Were Vietnam war draft avoiders also more likely to avoid smoking?" *Journal of Health Economics* 26.5 (2007): 896-926.

Grossman, Michael. "On the concept of health capital and the demand for health." *Journal of Political Economy* 80.2 (1972): 223-255.

Grossman, Michael. "The human capital model." in A.J. Culyer, J.P. Newhouse (Eds.) *Handbook of Health Economics* 1 (2000): 347-408.

Grossman, Michael. "The relationship between health and schooling." *Eastern Economic Journal* 34.3 (2008): 281-292.

Jamal, Ahmed, et al. "Current cigarette smoking among adults—United States, 2005–2014." *Morbidity and Mortality Weekly Report* 64.44 (2015): 1233-1240.

Kleven, Henrik Jacobsen, and Wojciech Kopczuk. "Transfer program complexity and the take-up of social benefits." *American Economic Journal: Economic Policy* 3.1 (2011): 54-90.

Knight-Ridder Newspapers. “Social Security Cuts From Reagan Years Being Restored.” (1989). http://articles.chicagotribune.com/1989-12-04/news/8903150162_1_average-monthly-benefit-disabled-rolls. Accessed 4 March 2018.

Lleras-Muney, Adriana. "The relationship between education and adult mortality in the United States." *The Review of Economic Studies* 72.1 (2005): 189-221.

Ma, Jennifer, Baum, S., Pender, M., & Bell, D. *Trends in College Pricing 2015*. New York, NY: The College Board (2015).

Meara, Ellen R., Seth Richards, and David M. Cutler. "The gap gets bigger: changes in mortality and life expectancy, by education, 1981–2000." *Health Affairs* 27.2 (2008): 350-360.

Middleton, S., Maguire, S., Ashworth, K., Legge, K., Allen, T., Perren, K., Battistin, E., Dearden, L., Emmerson, C., Fitzsimons, E. & Meghir, C. (2003) *The Evaluation of the Education Maintenance Allowance Pilots: Three Years Evidence. A Quantitative Evaluation*, DfES RR 499.

National Center for Health Statistics. "Health, United States, 2015: With Special Feature on Racial and Ethnic Health Disparities." Hyattsville, MD. (2016).

Parker, Kim, Anthony Cilluffo, and Renee Stepler. “6 facts about the U.S. military and its changing demographics.” (2017). http://www.pewresearch.org/fact-tank/2017/04/13/6-facts-about-the-u-s-military-and-its-changing-demographics/ Accessed 26 February 2018.

Rasella, Davide, Rosana Aquino, Carlos Santos, Rômulo Paes-Sousa, & Mauricio Barreto. "Effect of a conditional cash transfer programme on childhood mortality: a nationwide analysis of Brazilian municipalities." *The Lancet* 382.9886 (2013): 57-64.

Scott-Clayton, Judith. "On money and motivation a quasi-experimental analysis of financial incentives for college achievement." *Journal of Human Resources* 46.3 (2011): 614-646.

Schoenborn, Charlotte, Patricia Adams, & Jennifer A. Peregoy. "Health behaviors of adults: United States, 2008-2010." *Vital and health statistics* 10.257 (2013): 1-184.

Seftor, Neil S., and Sarah E. Turner. "Back to school: Federal student aid policy and adult college enrollment." *Journal of Human Resources* (2002): 336-352.

Schady, Norbert, and Maria Caridad Araujo. "Cash transfers, conditions, and school enrollment in Ecuador." *Economía* 8.2 (2008): 43-70.

Schultz, T. Paul. "School subsidies for the poor: evaluating the Mexican Progresa poverty program." *Journal of Development Economics* 74.1 (2004): 199-250.

Sirin, Selcuk R. "Socioeconomic status and academic achievement: A meta-analytic review of research." *Review of Educational Research* 75.3 (2005): 417-453.

Selective Service System. “Effects of Marriage and Fatherhood on Draft Eligibility.” (2016a). https://www.sss.gov/About/History-And-Records/Effects. Accessed 10 August 2016.

Selective Service System. “Postponements, Deferments, Exemptions.” (2016b). https://www.sss.gov/About/Return-to-the-Draft/Postponements-Deferments-Exemptions. Accessed 10 August 2016.

Social Security Administration. “Benefits Planner: Income Taxes and Your Social Security Benefits.” (2016a). https://www.ssa.gov/planners/taxes.html. Accessed 15 August 2016.

Social Security Administration. “Benefits for Children.” (2016b). https://www.ssa.gov/pubs/ EN-05-10085.pdf. Accessed 15 August 2016.

Silles, Mary A. "The causal effect of education on health: Evidence from the United Kingdom." *Economics of Education Review* 28.1 (2009): 122-128.

Snyder, Thomas D., Cristobal de Brey and Sally A. Dillow. *Digest of Education Statistics 2014*. National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC. (2016).

Snyder, T.D. (1993). *120 Years of American Education: A Statistical Portrait*. U.S. Dept. of Education, Office of Educational Research and Improvement, National Center for Education Statistics, Washington, DC

Stock, James H., and Motohiro Yogo. "Testing for Weak Instruments in Linear IV Regression." In D. Andrews, J. Stock (Eds.), *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, Cambridge University Press (2005): 80-108.

U.S. Department of Education. *Federal Pell Grant Program*. (2015a). https://www2.ed.gov/programs/fpg/index.html. Accessed 19 April 2018.

Stephens, Melvin, and Dou-Yan Yang. "Compulsory education and the benefits of schooling." *The American Economic Review* 104.6 (2014): 1777-1792.

U.S. Department of Education. *Federal Student Aid Handbook with Active Index 2015–2016*. (2015b). https://ifap.ed.gov/ifap/byAwardYear.jsp?type=fsahandbook&awardyear=2015-2016. Accessed 26 February 2018.

U.S. Department of Education, Federal Student Aid, *Annual Report FY 2016*, Washington, D.C. (2016).

U.S. Department of Health and Human Services. *The Health Consequences of Smoking: 50 Years of Progress. A Report of the Surgeon General.* Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health, 2014. Printed with corrections, January 2014.

U.S. Department of Veterans Affairs. 2015 *Veteran Economic Opportunity Report*. (2015).https://www.benefits.va.gov/benefits/docs/veteraneconomicopportunityreport2015.pdf. Accessed 26 February 2018.

Tables

Table 1: Effect of Social Security College Benefits on Health

	(1) BMI	(2) Very Good Health
Both Genders	-0.465**	0.012
	(0.192)	(0.010)
N	63,869	147,104
Female	-0.792***	0.030**
	(0.286)	(0.014)
N	34,867	76,683
Male	-0.087	-0.003
	(0.252)	(0.014)
N	29,002	71,373

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE)
Controls: Gender, race, citizen, interview year, minor benefit recipient (treat), high school cohort

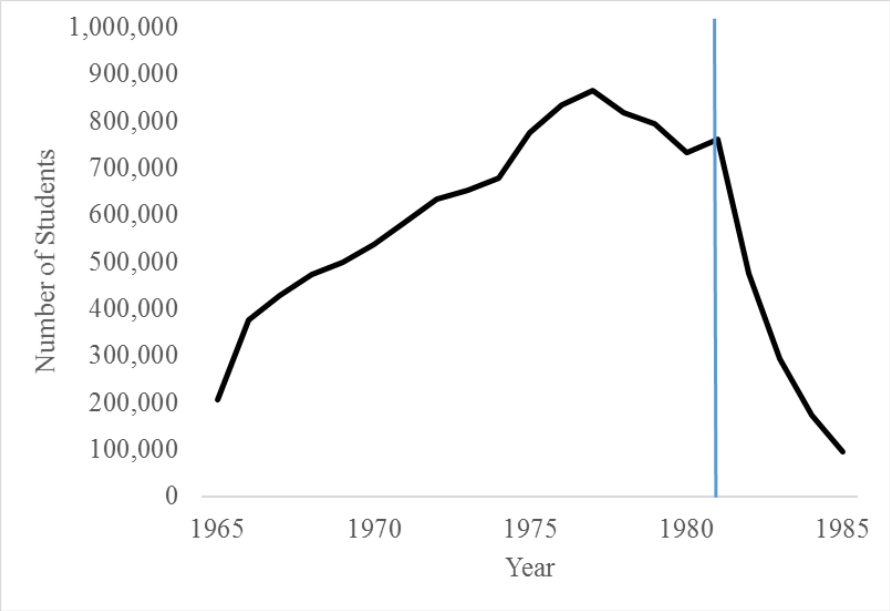
Table 2: Effect of Social Security College Benefits on Education

	Education Outcome		
	(1) Some College	(2) HS Graduation	(3) College Graduation
Both Genders	0.052***	0.021**	0.028***
	(0.011)	(0.008)	(0.009)
F-Statistic	24.29	6.49	10.96
N	147,104	147,104	147,104
Female	0.067***	0.031***	0.044***
	(0.015)	(0.011)	(0.012)
F-Statistic	21.48	8.17	13.27
N	75,731	75,731	75,731
Male	0.037**	0.010	0.013
	(0.015)	(0.012)	(0.012)
F-Statistic	5.70	0.73	1.21
N	71,373	71,373	71,373

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE)
Controls: Gender, race, citizen, interview year, minor benefit recipient (treat), high school cohort

Figures

Figure 1: Number of Social Security Student Benefit Recipients



Data source: DeWitt, 2001

Figure 2: Effect of Social Security College Benefits on BMI

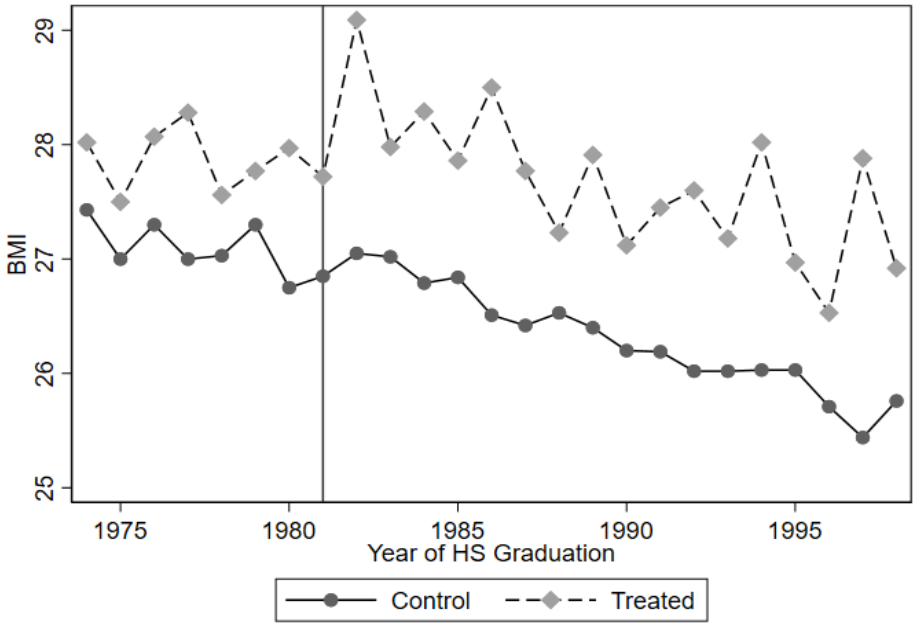


Figure 3: Effect of Social Security College Benefits on General Health

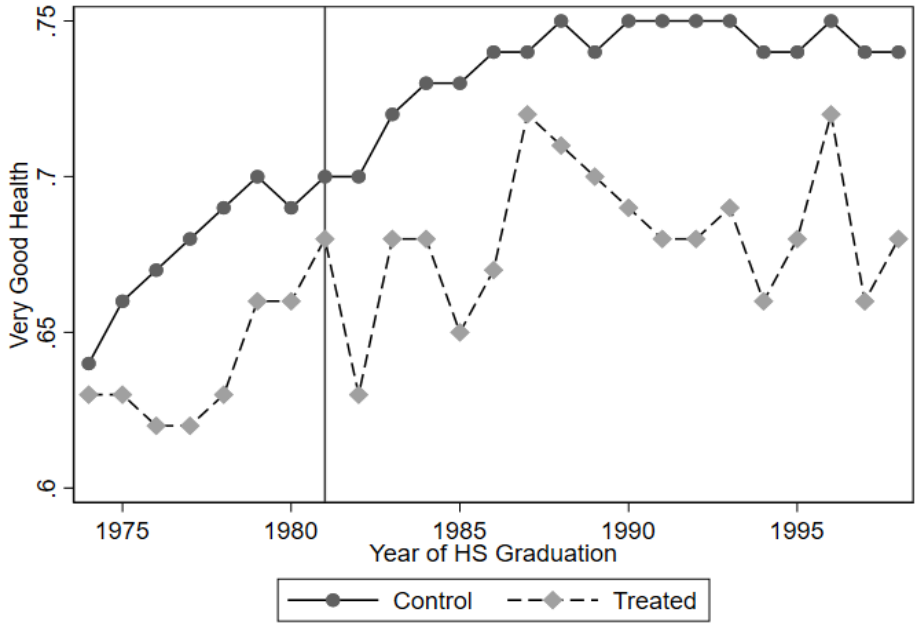
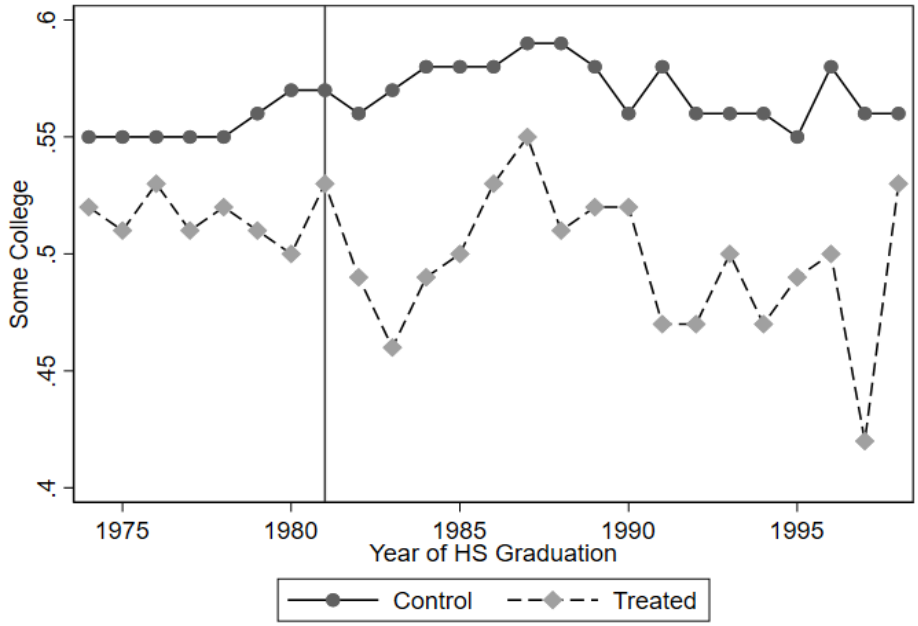


Figure 4: Effect of Social Security College Benefits on Attending College



Appendix Tables

Table A1: Requirements for Receipt of Social Security Children Benefits

Before 1965 and After 1981	**Between 1965 and 1981**
Not married AND	Not married AND
Under 18 OR	Under 18 OR
Under 20 and in secondary school OR	***Under 22 and in any school (including college)*** OR
Disabled since before age 18/22 (1957)	Disabled since before age 22

Source: Social Security Administration (2016b)
Note: Parent must be 1) Entitled to Social Security due to disability or retirement, OR 2) Met work requirement for Social Security but deceased

Table A2: Average Annual Social Security Benefit for Students and College Costs

Year	Avg. Yearly Benefit	4 Year Private Tuition/Fees	4 Year Public Tuition/Fees
1965	$4,380		
1972	$5,420	$8,100	$2,090
1980	$5,720	$7,620	$1,690

Source: DeWitt, 2001; Ma et al., 2015. (2000 $)

Table A3: Effect of Social Security College Benefits on Any College, NLSY79 Data

	Dynarski	My Analysis		
	Gender Pooled	Gender Pooled	Male	Female
	0.182**	0.148*	0.052	0.244*
	(0.096)	(0.086)	(0.118)	(0.125)
N	3,986	3,170	1,604	1,566

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE)

Table A4: Effect of Social Security College Benefits on BMI, Robustness Tests

	Column 1 1963 Birth Cut-off	Column 2 No Disabled Parents	Column 3 Main Results
	-0.577***	-0.455**	-0.465**
	(0.192)	(0.202)	(0.192)
N	64,878	57,102	63,869

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE)
Controls: Gender, race, citizen, interview year, minor benefit recipient (treat), high school cohort

Table A5: Effect of Social Security College Benefits on General Health, Robustness Tests

	Column 1 1963 Birth Cut-off	Column 2 No Disabled Parents	Column 3 Main Results
	0.008	0.012	0.012
	(0.010)	(0.010)	(0.010)
N	148,818	133,631	147,104

*p-value<0.10, **p-value<0.05, ***p-value<0.01
Standard errors in parentheses (SE)
Controls: Gender, race, citizen, interview year, minor benefit recipient (treat), high school cohort

Appendix Graphs

Figure A1: Effect of Social Security College Benefits on Race

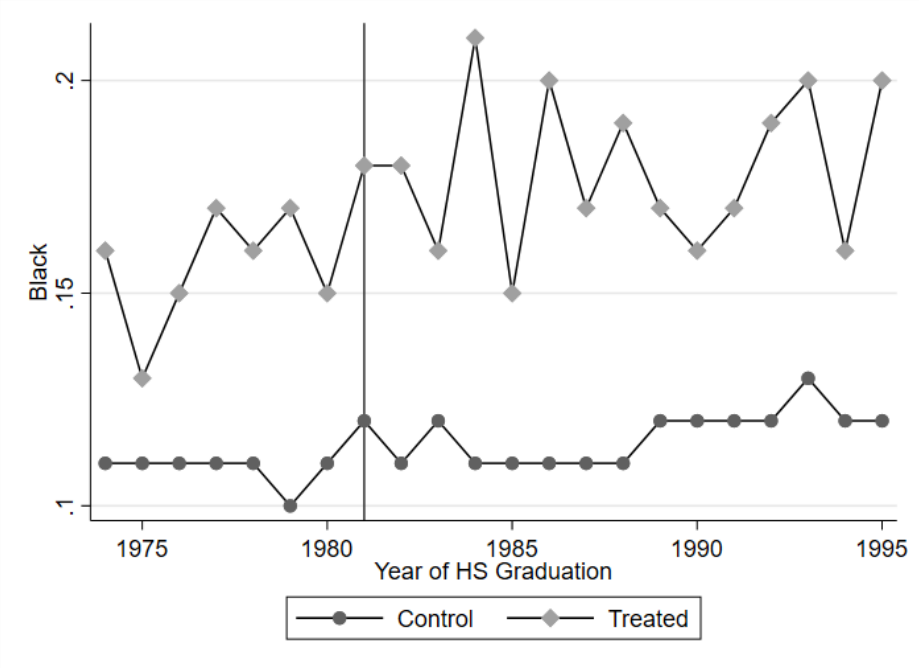


Figure A2: Effect of Social Security College Benefits on Ethnicity

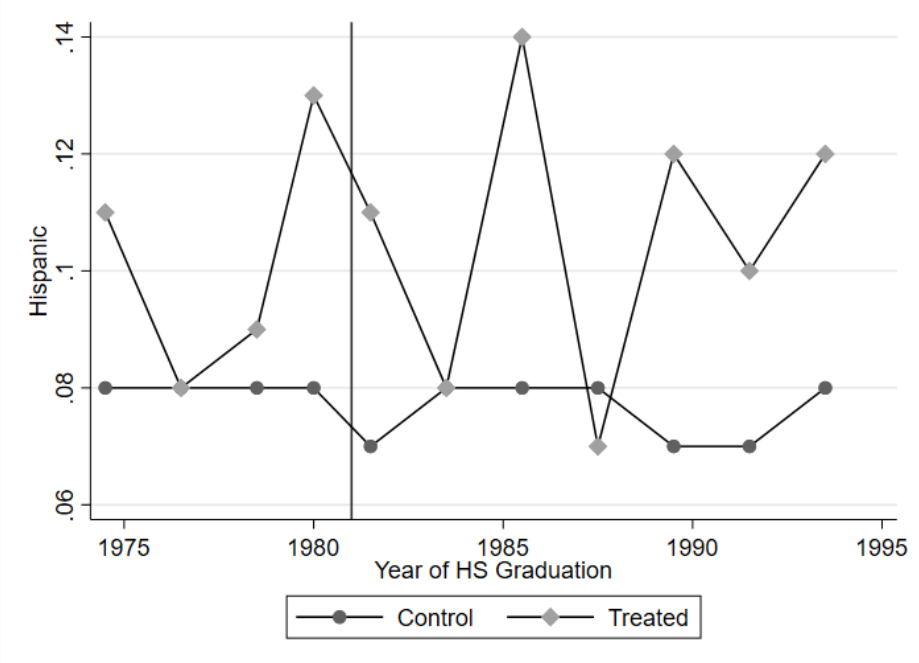


Figure A3: Effect of Social Security College Benefits on Height

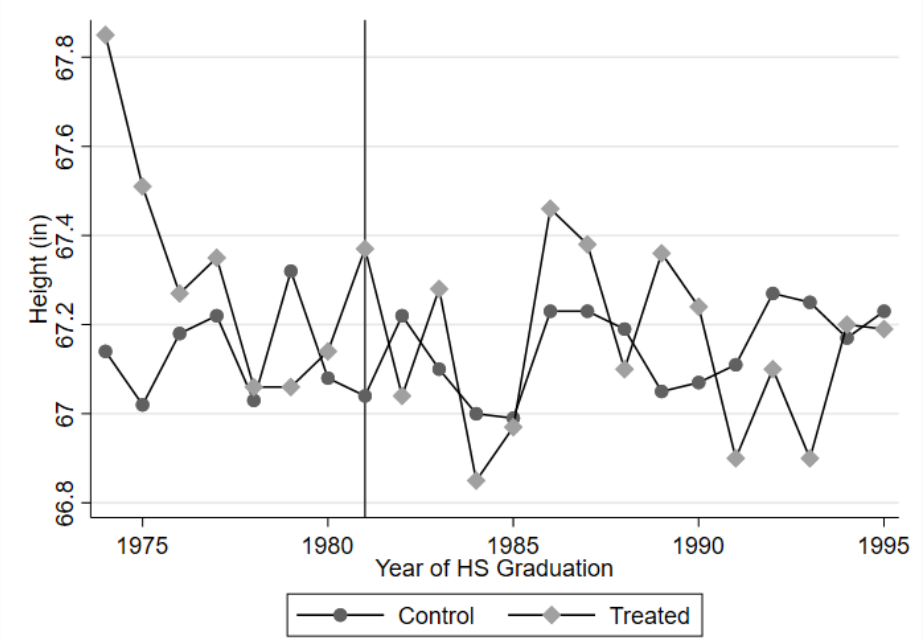


Figure A4: Effect of Social Security College Benefits on Weight