Summer 8-11-2020

Essays on Health and Education Economics

Ishtiaque Fazlul

Georgia State University

ABSTRACT

ESSAYS ON HEALTH AND EDUCATION ECONOMICS

BY

ISHTIAQUE FAZLUL

August 2020

Committee Chair: Dr. James Marton

Major Department: Economics

This dissertation evaluates the impact of various institutions and policies on health and educational outcomes.

The first chapter estimates the effect of a school stipend program for girls' education in Bangladesh on the health of their children. For this study, I use five rounds of a repeated cross-sectional survey, which has a rich set of objectively measured child height and weight information. Using the geographic and cohort variation of the program implementation, I find that the stipend program led to a lower probability of stunting as well as higher height for age and weight for age for the children of the stipend-eligible women. The effects are larger for first births and larger still for women who were induced by the program to acquire more years of schooling.

The second chapter, co-authored with Todd Jones and Jonathan Smith, explores the phenomena of taking an Advanced Placement (AP) course but not the exam. Using data from four large school districts in the metro Atlanta area, we find that 15 percent of the AP courses do not

lead to an exam. Traditionally disadvantaged populations have higher rates of missing the AP exam. Using a rich set of individual-level academic and demographic variables, we predict that a large portion of these courses would lead to a credit-granting score (3 or higher) if the exams were taken. We find evidence that students are more likely to take the AP exam when school districts offer a more generous AP exam subsidy. We find no evidence that a female AP student being paired with a female teacher leads to higher rates of exam-taking, even for courses where females are underrepresented.

The third chapter, co-authored with Charles Courtemanche, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata, estimates the effect of the major components of ACA on various types of insurance coverage during the first two years of the Trump administration (2017 and 2018). During this time, the implementation of different components of the ACA that would impose a burden have been tabled, budgets for outreach have been cut, and a vote to repeal the ACA narrowly failed. All of these things may have affected the impact of the ACA on insurance coverage during 2017 and 2018. In non-expansion states, which should be impacted by the national components of the ACA alone, we see statistically significantly smaller coverage increases in 2017 and 2018 as compared to 2016. In expansion states, which should be impacted both by the national components of the ACA and the Medicaid expansion, we estimate similar gains in coverage in 2016 through 2018. This difference between expansion and non-expansion states is due to smaller year over year increases in coverage because of the national components of the ACA and larger year over year increases because of the Medicaid expansion.

ESSAYS ON HEALTH AND EDUCATION ECONOMICS

BY

ISHTIAQUE FAZLUL

A Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree

of

Doctor of Philosophy

in the

Andrew Young School of Policy Studies

of

Georgia State University

GEORGIA STATE UNIVERSITY

August 2020

Acceptance

This dissertation was prepared under the direction of the candidate's Dissertation Committee. It has been approved and accepted by all members of that committee, and it has been accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics in the Andrew Young School of Policy Studies of Georgia State University.

Dissertation Chair: Dr. James Marton

Committee: Dr. Tim Sass
Dr. Tom Mroz
Dr. Charles Courtemanche

Electronic Version Approved:

Sally Wallace, Dean
Andrew Young School of Policy Studies
Georgia State University
August 2020

Dedication

To my mother, Shamima Begum, and my father, Dr. Fazlul Awal Mollah for their unconditional love and unending support for everything I have done in life

&

To my wife Kazi Kimiya

Acknowledgments

This dissertation would not be possible without the support of my mentors, friends, and family.

I am grateful to my supervisor Dr. James Marton for his mentorship. Thanks for being abundant with your time and support. Your mentorship went beyond teaching me techniques and methods. You taught me how to navigate the publication process, introduced me to academics in your network, and helped me immensely in the job search process. The PhD journey can be frustrating and thankless at times, but your mentorship made the experience rewarding and enjoyable, even.

I am indebted to my dissertation committee members, teachers, and mentors. Thanks, Dr. Tim Sass, for introducing me to Education Economics and bringing me into the Georgia Policy Labs (GPL). You have played an important role in shaping my research interest and have played an essential role in my job placement. I thank Dr. Thomas Mroz and Dr. Charles Courtemanche, for their valuable feedback on my work. I also want to thank Dr. Jonathan Smith for his comments on my work and for teaching me how to write a good introduction. My heartfelt thanks to Dr. Atonu Rabbani for believing in me and rekindling my interest in academia.

Friends are the family we choose ourselves, and I am glad to have chosen you, Tareena Musaddiq and Siyu Pan, as my friends in graduate school. I am fortunate to have an amazingly helpful cohort in the PhD program without whom navigating this process would not be possible. Between the endless study sessions in the first couple of years and the random discussions on identification strategies in the last two, I believe I have learned more from you collectively than any mentor.

I want to thank my second family in Atlanta for their hospitality and love. Thanks, Biplab Kumar Datta, for being my guiding light and Nilanjana Chatterjee for being Nilapu. Thanks, Tania Afsana and Muhammad Jami Husain, for being my anchors in Atlanta; Mudabbir Husain and Sohani Fatehin for being sources of inspiration and life advice; and Zia Haque and Raiyan Hridita for being up for impromptu meetings any hour of the day (or night). This journey would not be as rewarding as it was without all of you, and for that, I am grateful.

I want to express my heartfelt gratitude to ammu and abbu for their love and support in everything I wanted to do in life. You always let me be who and what I wanted to be. Ammu, thanks for listening to my frustrations during all of my academic life. Abbu, thanks for instilling in me the importance of hard work through your words and your own work, and for letting me know that it is OK to be less than perfect. Thanks, Marzina Ahmed, for being my friend, my aunt, my shelter, and for being the example of an independent woman. Thanks, Shahnewaz Kabir, for all the books, movies, music, philosophers, and artists you introduced me to and for kindling my intellectual pursuit in the social sciences and arts.

Kimiya, you are the most effortlessly smart and kind person I know. You proved over and over that you can be whatever you want to be. Thanks for choosing to be by my side through thick and thin for the last fifteen years.

Table of Contents

List of Figures

List of Tables

Introduction

Health and education are two of the most critical components of human capital that lead to higher income and wellbeing. Policies and institutions that affect health and education can have far-reaching consequences in improving quality of life. This dissertation includes three papers that explore how three different policy interventions and institutions concerning education and health affect various relevant outcomes.

The policy interventions and institutions explored in this dissertation affect access to education and health care. The first chapter evaluates the effect of a secondary school stipend program for rural girls in Bangladesh. This program was originally designed to improve access to education for girls in the rural areas of Bangladesh who had low secondary school-participation rates even among girls (girls had a lower secondary school enrollment compared to boys) before the intervention in 1994. The stipend program was designed to increase girls' years of schooling, delay their marriage, and increase their labor market participation. Existing literature shows that it succeeded in all these primary goals. In this chapter, I estimate how this program affects the health of the children of its recipients. The second chapter of this dissertation explores AP exam-taking behavior and the relationship of AP subsidy policies and teacher demographics (specifically AP teacher and student gender match) with exam taking. The AP program in general and the two policy aspects evaluated (subsidy and gender match) are interventions designed to improve access to education. Finally, the program evaluated in the third chapter of this dissertation is the Affordable Care Act (ACA). The ACA is a multi-pronged intervention that includes Medicaid expansions, insurance market reforms, subsidized Marketplace coverage, and an individual mandate with the primary aim of increasing insurance coverage.

The outcomes analyzed include both intermediate as well as final outcomes. Outcomes such as lower childhood stunting, improved height for age and weight for age in a developing country are desirable outcomes by themselves. On the other hand, taking the AP exam after taking AP courses or increases in insurance coverage are intermediate steps to achieving other desirable goals. There is existing literature showing that earning college credit while in high school is related to performance in college and college graduation (Dougherty, Mellor, and Jian, 2006; Morgan and Klaric, 2007; An, 2012; Allen and Dadgar, 2012; Patterson and Ewing, 2013) and higher probability of graduating college in four years (Smith, Hurwitz, and Avery, 2017). There is also evidence that health insurance coverage leads to more healthcare utilization (Manning, 1987; Baickeret al., 2013) and better health (Brook et al., 1983; Finkelstein et al., 2013).

In the first chapter, I examine the effect of an educational intervention designed to increase girls' years of schooling on their children's health. The educational intervention used in this study is a secondary school stipend program for rural girls in Bangladesh. The second generation's outcome measures considered are stunting, wasting, the prevalence of being underweight, and the corresponding z scores for children under the age of five. To solve the endogeneity problem in the relationship between mother's education and child health, I use the mother's birth cohort and geographic variation in program implementation in a difference in differences framework. I find that one more year of mother's eligibility for the stipend program decreases the probability of her child being stunted by one percentage point (2.3 percent). One more year of eligibility for the program also leads to 0.031 and 0.016 standard deviation increase in height for age and weight for age, respectively. Using mother's years of eligibility as an instrument for her years of schooling, I present evidence that for those who were induced by the stipend program to acquire more education, one more year of mother's schooling leads to a 3.6 percentage point (8.4 percent) decrease in the child's probability of stunting. The paper advances the literature by presenting evidence of an intergenerational effect of education in a developing country and also by considering a broader definition of child health compared to birth-weight or infant mortality.

In the second chapter, we explore the phenomena of taking Advanced Placement (AP) courses but not taking the exams, investigate if the students are "leaving credit on the table" by doing so, and identify some of the possible determinants of this behavior. Advanced Placement program (AP) gives high school students in the US an opportunity to take college-level courses and get college credit for them. In general, a student first takes the corresponding AP course, which helps prepare him/her to take the AP exam and provides knowledge in its own right. In order to gain the full benefit of AP, though, the student needs to take and pass the AP exam. The exam, if passed, provides college credit, making it easier for the student to graduate college (Smith, Hurwitz, and Avery, 2017). Given this and other potential benefits, students who take the AP course and are predicted to pass the AP exam are likely leaving something on the table if they do not take the exam. Using data from four metropolitan Atlanta public school districts, we find that 15 percent of AP courses do not result in an AP exam. There are substantial disparities in AP exam-taking rates. White, Asian, and non-Hispanic students' AP courses do not result in an AP exam 13, 10, and 15 percent of the time, respectively, while Black and Hispanic students' AP courses do not result in an AP exam 23 and 18 percent of the time, respectively. We estimate that up to 32 percent of the AP courses that do not result in an AP exam would result in a score of 3 or higher, which generally commands college credit at colleges and universities across the country. We then turn to several explanations as to why students seemingly leave college credit on the table. We find evidence that students are more likely to take the AP exam when school districts offer more generous funding. In fact, free and reduced price lunch (FRL) students in the districts that more generously subsidize FRL students are 3 percentage points more likely to take an AP exam than their non-FRL counterparts. We find little evidence that a female student paired with a female AP course teacher takes the AP exam at a higher rate as compared to being paired with a male teacher, even in courses where female students are underrepresented.

The third chapter estimates the effect of the major components of ACA (Medicaid expansion, subsidized Marketplace coverage, the individual mandate, and changes in the insurance market regulations) on various types of insurance coverage during the first two years of the Trump administration (2017 and 2018). During this time, the implementation of different components of the ACA that would impose a burden have been tabled, budgets for outreach have been cut, and a vote to repeal the ACA narrowly failed. All of these things may have affected the impact of the ACA on insurnace coverage during 2017 and 2018. We use a difference-in-difference-in-differences model, developed in the recent ACA literature, to separately identify the effects of the national and Medicaid expansion portions of the law. We use data from the 2011–2018 waves of the American Community Survey (ACS), with the sample restricted to nonelderly adults. In non-expansion states, which should be impacted by the national components of the ACA alone, we see statistically significantly smaller coverage increases in 2017 and 2018 (3.8 percentage points in each year) as compared to 2016, where we estimate a 5 percentage point increase. In expansion states, which should be impacted both by the national components of the ACA and the Medicaid expansion, we estimate similar gains in coverage in 2016 through 2018 (about 11 percentage points in each year). This difference between expansion and non-expansion states is due smaller year over year increases in coverage due to the national components of the ACA and larger year over year increases due to the Medicaid expansion.

One connecting thread in all three chapters in this dissertation is their policy relevance. Each of the three chapters addresses a pressing policy issue. Worldwide 149 million under-five children are stunted, and 49 million are wasted.[1] In Bangladesh, a South-Asian country, 36 percent of all under-five children are stunted, and 14.3 percent are wasted.[2] WHO's Global Targets 2025 includes goals to reduce under-five stunting by 40 percent and wasting to less than five percent.[3] Given the results of the first chapter, increasing mother's education can be one of the policy levers, along with the traditional public health measures, through which these goals can be achieved.

Earning college credits while in high school is related to numerous positive collegiate outcomes, including performance in college and college graduation (Dougherty et al., 2006; Morgan and Klaric, 2007; An, 2012; Allen and Dadgar, 2012; Patterson and Ewing, 2013) and higher probability of graduating college in four years (Smith et al., 2017). These positive effects of successful completion of AP exam likely motivates states such as Georgia, to subsidize at least one AP exam for all free and reduced price lunch (FRL) students since 2003. Some Georgia school districts go above and beyond this state subsidy, in some cases making all AP exams free for all students. By documenting the prevalence of AP exam skipping, as well as possible inequalities by race and socioeconomic conditions and exploring possible determinants of this behavior, this study helps policymakers understand the nature and extent of the problem. This chapter also shows the effectiveness of AP subsidy policy in increasing AP exam taking and thus validates a policy tool used by some school districts in Georgia.

[1] UNICEF/WHO/World Bank Joint Child Malnutrition Estimates, March 2019 edition.
[2] National Institute of Population Research and Training [NIPORT] et al. 2016
[3] Reduction of Stunting and wasting is also part of multiple Sustainable Development Goal (SDG) targets set by United Nations in 2015 (SDG 1.1; 1.2; 1.3; 2.1; 2.2; 10.4) (World Health Organization, 2017).

Finally, the debate about the future of the ACA continues to be a dominant theme in all levels of politics and will feature prominently in the upcoming presidential election. Many states are debating an expansion of their Medicaid program under the ACA, especially given the current pandemic. All this makes our investigation into the effect of ACA in the first two years of the Trump administration policy relevant.

The remainder of this dissertation is organized around the three chapters.

1. Inter-generational Effects of Education on Health: Evidence from a Secondary School Stipend Program

1.1. Introduction

In this paper, I explore the inter-generational effects of education on health, exploiting the introduction of a secondary school stipend program in Bangladesh. There are multiple possible channels through which maternal education can affect a child's health. For example, maternal education may affect child health through income. There is ample evidence of the positive relationship between education and income (e.g., Angrist and Keueger, 1991; Ashenfelter and Krueger, 1994). Increased income due to more years of education can lead to better nutrition during pregnancy. The fetal origins hypothesis literature shows that better nutrition during gestation leads to improved child health (Barker, 1990; Almond and Mazumder, 2005; Almond, 2006). Thus more education for mothers can lead to healthier children through increased household income.

Mother's education may also affect child health through maternal health. Grossman's model of health production predicts education to have a positive effect on health by improving productive and allocative efficiency (Grossman, 1972). Empirical literature validates this prediction showing a positive causal relationship between education and own health (Lleras-Muney, 2005; Oreopoulos, 2006; Silles, 2009; Braakman, 2011; Clark and Royer, 2013). Also, maternal health at birth affects children's birth weight (Currie and Moretti, 2007; Black, Devereux, and Salvanes, 2007; Royer, 2009). Combining the effects of education on a mother's own health and the effect of maternal health on child health, we can argue that mothers' education may have a positive effect on child health.

Thus, there are separate strands of literature focusing on the relationship between maternal education and income, maternal education and maternal health, household income and child health, and maternal health and child health. However, there is a dearth of causal evidence quantifying the direct relationship between programs that increase mothers' education and child health. In other words, though there is ample evidence on pieces of the story, there is little evidence on the broader question of the inter-generational effect of education on health. In this paper, I estimate the inter-generational effect of education on child health, namely stunting, wasting, and the prevalence of being underweight.

The health outcomes that I consider are stunting, wasting, the prevalence of being underweight and the corresponding height for age z score (HAZ) weight for age z score (WAZ), and weight for height z score (WHZ). A child is considered stunted, underweight, or wasted if her height for age, weight for age, or weight for height are more than two standard deviations below the WHO Child Growth Standard Median. Both stunting and wasting indicate poor nutrition or repeated infections. Stunting is associated with lower cognition, educational performance, wage, and productivity. Wasting leads to an increased risk of death. Acknowledging the severe long-term adverse effects of stunting and wasting, WHO's Global Targets 2025 includes goals to reduce under-five stunting and wasting.[4] Early childhood malnutrition is one of the biggest healthcare challenges the world currently faces. According to the WHO, there were 149 million under-five children stunted and 49 million wasted in 2018.[5] These numbers make up 22 percent and 7.3 percent of the world's under five-year-old population in the world. Under-five stunting and

[4] "40 percent reduction in the number of children under-five who are stunted" and "…reduce and maintain childhood wasting to less than 5 percent". Reduction of Stunting and wasting is also part of multiple Sustainable Development Goal (SDG) targets set by United Nations in 2015 (SDG 1.1; 1.2; 1.3; 2.1; 2.2; 10.4) (World Health Organization, 2017).
[5] UNICEF/WHO/World Bank Joint Child Malnutrition Estimates, March 2019 edition.

wasting are the most prevalent in South Asia, where 32.7 percent of all under-five are stunted, and 14.6 percent of all under-five are wasted.[6] In Bangladesh, a South-Asian country, 36 percent of all under-five children are stunted, 14.3 percent are wasted, and the under-five mortality rate is 3.24 percent. [7,8]

The purpose of this paper is to estimate the causal effect of an educational intervention designed to increase years of schooling for girls on their children's health, namely stunting, wasting, and prevalence of being underweight. The relationship between a mother's years of schooling and her child's health is endogenous. Mother's years of schooling may be correlated with factors such as mother's innate abilities and how career-oriented she is, all of which may be correlated with child health. This would lead to omitted variable induced endogeneity problem and lead to biased estimates. To solve this problem, I exploit the introduction of a female secondary school stipend program in rural Bangladesh. I use mother's birth cohort and geographic variation in the program roll-out in a difference in differences framework. I also estimate the effect of mother's years of schooling on child health using the mother's years of eligibility for the stipend program as an instrument for mother's years of schooling. My data source is the Bangladesh Demographic Health Survey (BDHS) from the years 2000, 2004, 2007, 2011, and 2014.

While there have been a number of studies estimating the causal effect of parent's education on birth-weight (Currie and Moretti, 2003; McCrary and Royer, 2006), infant mortality (Breierova and Duflo, 2004; Chou et al., 2010; Grépin and Bharadwaj, 2015), age of marriage or pregnancy (Alam, Baez, and Carpio, 2011; Hong and Sarr, 2013; Hahn, et al., 2018) and Fertility

6 UNICEF/WHO/World Bank Joint Child Malnutrition Estimates, March 2019 edition.
7 National Institute of Population Research and Training [NIPORT] et al. 2016
8 UN Inter-agency Group for Child Mortality Estimation. Available in: https://childmortality.org/data and https://data.worldbank.org/indicator/SH.DYN.MORT?locations=BDanddisplay=graph

(Angeles, Guilkey, and Mroz, 2005; Long and Osili, 2008), to my knowledge there has been none estimating the effect of mother's education on under-five stunting, wasting and prevalence of being underweight.

Chou et al. (2010) and Breierova and Duflo (2004) study the effect of parent's education on infant mortality in Taiwan and Indonesia, respectively. Both of these countries had higher rates of female education and lower rates of under-five malnutrition at the beginning of the program compared to Bangladesh. My study is in a context where the issue of child malnutrition is the most pressing. One paper that is set in a country with similarly high malnutrition is Grépin and Bharadwaj (2015). It uses the post-liberation change in school attendance rule in Zimbabwe to identify the effect of maternal secondary education on child survival. But the paper also notes that there were several concurrent social and institutional changes in the healthcare sector in the then just-liberated country. This may have biased the results and over-estimated the effect of mother' schooling on child survival. Moreover, their outcome variable is child mortality and not measures of child health. Studies that look at broad conditional cash transfer (CCT) programs on child health also fail to isolate the effect of mother's education as the programs have multiple conditions such as school attendance, vaccination, and nutrition; some of which may affect child health positively. The Female School Stipend Program is a unique CCT in the sense that the primary condition of receiving the stipend is to remain in school and does not include other health and nutrition-related conditions.

Hahn et al. (2018) is the only paper that estimates the effect of the Female Secondary School Stipend Program (FSSSP), the same program that I evaluate, on second-generation child health. It, however, differs from my study in four important ways. First, they only have HAZ, WAZ, and WHZ and not stunting, wasting, and the prevalence of being underweight as outcomes.

Unlike HAZ, WAZ, and WHZ, the binary outcomes stunting, and wasting are unambiguously bad outcomes. Also, these binary measures are especially relevant margins for the under-five children of Bangladesh. Second, they use three survey rounds of BDHS for the years 2007, 2011 and 2014, whereas my study uses five rounds of BDHS data from 1999 and 2014. Third, because the effect of the program on second-generation child health is not a primary consideration of their paper, they don't test the robustness of these results.

This paper advances the literature in three ways. First, it estimates the causal effect of a program targeted to increase female education on the second generation's health. Second, it estimates the causal effect of mother' education on relatively understudied, yet important measures of child health: stunting, wasting, and prevalence of being underweight. Third, it is the first to provide credible causal estimates of the effect of mother's education on under-five child health in a country with a very low level of female education and a high level of under-five child malnutrition.

I find that one more year of mother's eligibility for the stipend program decreases the probability of her child being stunted by 1 percentage point (2.3 percent). The positive and statistically significant effect is stable across different specifications, including the addition of district by child's year of birth fixed effect and restricting sample to first births only. One more year of eligibility for the program also leads to 0.031 and 0.016 standard deviation increase in height for age and weight for age respectively. Using mother's years of eligibility as an instrument for her years of schooling, I present evidence that for those who were induced by the stipend program to acquire more education, one more year of mother's schooling leads to a 3.6 percentage point (8.4 percent) decrease in the prevalence of stunting. These effects are more pronounced for first births. These results are robust to multiple alternative specifications and falsification tests.

My results show that investment in mother's education leads to positive health effects in the next generation. Not including these positive inter-generational effects in the benefit-cost calculations of female education programs may lead to an underestimation of the program's effectiveness. The results have implications for policies that encourage female education in developing countries and inter-generational effect of female schooling in general.

1.2. Literature Review

The effect of education on health has been studied extensively in the literature. The Grossman model postulates that people produce health by using medical care as an input while education, health status, and income affect the production of health by affecting its shadow price (Grossman, 1972). Education can improve health by prompting an individual to increase inputs in the form of medical care and exercise (allocative efficiency), or more educated people may be better and more efficient producers of health using the same amount of inputs (productive efficiency). Several causal studies show a convincing positive relationship between education and own health (Lleras-Muney, 2005; Oreopoulos, 2006; Silles, 2009). Fewer studies investigate the relationship between the parent's education and child health, and the results are mixed. Some of the studies in developed countries include Currie and Moretti (2003), McCrary and Royer (2006), Lindeboom et al. (2009), and Silles (2009) while few such as Lucia Breierova and Esther Duflo (2004), Chou et al. (2010), and Grepin and Bharadwaj (2015) explore developing countries. These studies are discussed in more detail below.

1.2.1. Causal Impact of Education on Own-Health

Several studies exploit minimum compulsory schooling laws as a source of exogenous variation in years of schooling to identify the effect of schooling on own health. Most of these studies focus on developed countries. Lleras-Muney (2005) uses changes in state compulsory schooling laws from 1915 to 1939 as a source of exogenous variation to identify the impact on mortality rates using instrumental variable and regression discontinuity designs. Both Oreopoulos (2006) and Silles (2009) use the mandatory years of schooling law change in the UK from 14 to 15 in 1963 and then to 16 in 1973 as sources of exogenous variation in years of education. They use eligibility of birth year cohorts as an instrument for mother's years of schooling and find education to improve self-reported health. Braakmann (2011) uses the same schooling requirement laws in the UK but finds no effect of education on self-assessed health. Clark and Royer (2013) also uses the two British compulsory schooling laws and uses an RD design and also finds no evidence of education affecting self-assessed health.

1.2.2. Causal Impact of Parents' Education on Child Health

Chou et al. (2010) exploits a Taiwanese law extending compulsory education from 6 to 9 years in 1968 and the opening of 150 new junior high schools from 1968 to 1973 as an exogenous shock to education to identify its effect on infant mortality. The study uses the interaction of compulsory years of schooling increase and county-specific new school openings as an instrument for years of schooling. They find that an increase in parent's schooling decreases infant mortality. A similar methodology of exploiting the interaction of cohort and intensity of intervention as an instrument for years of schooling was used by Breierova and Duflo (2004) to estimate the effect of parent's education on child mortality in Indonesia. They exploit two rounds of a primary school construction program during 1973-1974 and 1978-1979. In this case, again, program intensity is defined by the number of schools built in each district. They find that an increase in parent's education reduces infant mortality.

Osili and Long (2008) apply a similar method to Breierova and Duflo (2004) in Nigeria, where they estimated the effect of female schooling on fertility. In 1976, Nigeria introduced a nationwide program that provided tuition-free primary education and increased the number of primary school classrooms at differential rates in the 19 states of the country. The paper used the interaction between the year of birth and program intensity, measured by the per capita amount of federal funds given to each state for classroom construction, as an IV for mother's years of schooling. The authors found that an increase in mother's education by one year reduces fertility by 0.26 births. This is relevant to child health because it is hypothesized in the literature that a lower number of births is associated with healthier children (Becker and Lewis, 1973).

Angeles, Guilkey, and Mroz (2005) used a maximum likelihood procedure to estimate the effect of mother's education and the existence of health centers providing family planning on fertility in Indonesia. The paper uses a discrete-time hazard model for age at marriage and education level of spouse and a logit model for fertility experience starting at age ten while allowing the availability of family planning services in the place of residence when a girl is seven and later in her life to have separate effects. They find that the presence of a family planning program in a young woman's area leads to higher educational attainment and lower fertility. The authors also find that the effect of higher education on reduced fertility is overestimated if the endogeneity of education and marriage is not controlled for. This shows that endogeneity of education is crucial to control for if we want to estimate the effect of it on fertility and, by extension, child health.

Most recently, Grepin and Bharadwaj (2015) use variation from age-specific exposure to a series of reforms in Zimbabwe (used as an IV) that increased access to education for females to identify the effect of maternal secondary education on child survival in Zimbabwe. They find that an extra year of maternal education was associated with a 1.7 percentage point decrease in the probability that a child was reported to have died before the survey. They also attempt to identify channels through which years of mother's schooling affect child survival by using the same IV. Some of the channels they find are a delay in cohabitation, decreased number of births, and increased the age at first birth.

There have been studies investigating the causal effect of mother's education on child health in developed countries as well. Currie and Moretti (2003) examine the relationship between maternal education and birthweight among US white women with data from individual birth certificates from the Vital Statistics Natality files from 1970 to 2000. They used the information on college openings between 1940 and 1990 to construct an availability measure of college in a woman's seventeenth year as an instrument for schooling. McCrary and Royer (2011) use a combination of school entry policies and compulsory years of schooling laws as an IV for mother's education to estimate the effect of mother's education on the probability of low birth weight and infant mortality in Texas and California in the US. They do not find any effect of years of mother's education on childbirth outcomes. Grytten et al. (2014) use the Norwegian education reform in compulsory years of schooling as a source of exogenous variation to estimate the causal effect of mother's education on childbirth weight. They find that an increase in mother's years of education leads to lower probability of low birth weight. Gunes (2015) uses the increase in compulsory years of schooling in Turkey and finds that mother's primary school completion decreases the likelihood of very low birth weight births by 17 percentage points, increases the height for age and weight for age by 1.1, and 1 standard deviation. He instruments mother's years of education with the exposure to compulsory schooling law across cohorts.

Most of the studies reviewed here estimate the relationship between parent's (in some cases, only mother's) education and child survival. This outcome is surely the first and one of the most immediate inter-generational effects of mother's education to consider. But survival is a narrow definition of health. To my knowledge, Currie and Moretti (2003), McCrary and Royer (2006), Grytten et al. (2014), and Gunes (2015) are the only papers that look at the effect of mother's education on the health of the children who do survive, specifically the prevalence of low birthweight. My paper takes the next step by estimating the effect of mother's education on stunting, wasting and prevalence of being underweight of under-five children.

There have been a few studies evaluating the impact of the Female Secondary School Stipend Program (FSSSP) in Bangladesh and another similar program in Pakistan. Most of these papers look at the short-run goals of the programs, i.e., school enrollment of females, while some look at fertility and age at marriage as outcomes. A number of studies find a positive effect of FSSSP on female secondary schooling (Fuwa, 2001; Khandker, Pitt and Fuwa, 2003; Schurmann, 2009; Hong and Sarr, 2013) increased age of marriage (Hong and Sarr, 2013; Hahn et al., 2018) and labor force participation in the formal sector (Hahn et al., 2018). In Pakistan, a similar CCT program shows increased female enrollment (Chaudhury and Parajuli, 2006) delayed marriage, fewer births, and higher matriculation rates (Alam, Baez, and Carpio, 2011).

1.3. Institutional Background

Primary schooling in Bangladesh is grades 1 through 5, and secondary school is 6 through 10. In the 1990s, there was a large disparity between years of schooling of men and women in Bangladesh. In 1991, 75 percent of the primary school-aged girls (between 6 and 10 years old) were enrolled in a school while 85 percent of boys of the same age were enrolled in a school. The disparity continued in secondary school, where 14 percent of girls between the age of 11 and 16 were in secondary school, while for boys, it was 25 percent (World Bank, 2003). To address the inequality in secondary education in Bangladesh, the Female Secondary School Stipend Program (FSSSP) started in 1994 with three stated goals. First, to increase female enrollment and retention rates in secondary school. Second, to increase the age at which girls marry. Third, to enhance female employment opportunities. The program targeted girls from rural schools only.

The eligibility criteria for receiving the stipend are that the girl has to (a) have a minimum of 75 percent attendance rate at school, (b) have at least a 45 percent score in the annual exam, and (c) remain unmarried. The program was rolled out in all rural schools of Bangladesh at the same time after a brief pilot period. The annual stipends were equivalent to US$12 in Grade 6, US$13.50 in Grade 7, US$15 in Grade 8, US$30.25 in Grade 9, and US$36.25 in Grade 10 (Hahn et al., 2018). These amounts are equivalent to 4 percent to 12.4 percent of the 1994 per capita GDP.[9] Also, a book allowance in grade 9 and an examination fee in grade 10 was available. Tuition fees were covered as well, which was paid to the school directly. The cash stipend was paid directly to the girls in two annual installments in the form of deposits into savings accounts in the nearest bank branch. Not all grades received a stipend from 1994. In 1994 girls in grades 7 and 9 received stipends; in 1995, girls in grades 6,7,9 and 10 received stipends, and from 1996 onwards, girls in all grades from 6 to 10 received stipends. As a result, mothers born before 1979 (those who were supposed to be in grade 10 or higher in 1994) should not be eligible for the stipend program while rural mothers born including and after the year 1979 were eligible for 2 to 5 years of stipend. Table 1.1 shows the years of eligibility for different cohorts of mothers.

[9] All allowance are simple currency conversions and not PPP.

Though direct eligibility to the stipend program can be for up to 5 years (grades 6-10), indirectly, the duration of exposure may be longer. One can argue that the availability of the stipend program in secondary school can positively influence a girl's family's decision to send her to primary school in anticipation of future stipend. Thus, considering this anticipatory exposure and treating it the same way as direct exposure, exposure to the FSSSP can be modeled to be up to 11 years (grades 1-10). In this paper, I define exposure to be direct exposure, i.e., only the girls in rural areas old enough to go to secondary school in 1994 or younger are considered eligible for the stipend program and thus are considered exposed to the program. Therefore, when I report the effect of one more year of mother's eligibility to the stipend program, it is the effect of direct exposure to the program. This strategy will attribute any differential increase in the number of women in secondary school in rural areas compared to urban areas conditional on relevant covariates after the intervention to the years of eligibility and will include the effect of any anticipatory exposure in the direct exposure estimates.

The FSSSP supports over 2 million girls each year and consists of 60 percent of the secondary school development budget and 13 percent of the total education budget of Bangladesh (World Bank, 1997). This makes the stipend program a major undertaking for the country. Just above half of the program cost is born by the government of Bangladesh (GOB) and the rest by a number of donor organizations. Even though the program has seen a lot of success in increasing female schooling and reducing child marriage, there is a push from the donors to change the program and make it more targeted by restricting the stipend to only girls whose family is below an income threshold. The government of Bangladesh has historically been very much opposed to this idea.

1.4. Data

The data source for this study is five rounds of Bangladesh Demographic Health Survey (BDHS) collected in 1999-2000, 2004, 2007, 2011, and 2014. BDHS is a nationally representative household survey that collects detailed health-related information for mothers and children. It includes information on women aged 15 to 49 and men aged 15 to 59. The surveys include individual and household level information. One unique feature of the BDHS survey is that the surveyors weigh and measures all children in the household under the age of five. This information is used to calculate the weight for height, height for age, and weight for age deviation from reference median, i.e., HAZ, WAZ, and WHZ. BDHS data also include the binary variables stunting, underweight, and wasting that are calculated from the continuous z scores. Calculation of the z scores, as well as the binary variables, is explained in the following paragraphs.

Table 1.2 shows the means of the variables of interest for rural and urban areas of Bangladesh for the pre and post 1979 birth-year cohort of mothers. Of particular interest are the outcomes stunting, underweight, and wasting. The literature on under-five child health uses WHO-approved z-scores for height for age (HAZ), weight for age (WAZ), and weight for height (WHZ). These scores are comparable across age and sex. The scores are calculated using the following formula:

$$HAZ_i = (h_{ij} - h_j)/\sigma_j$$

where hij is the observed height of child i in group j, where a group is defined according to a child's sex and birth month. hj and σj are the median and standard deviation of the height in group j, using American children as the reference population. HAZi is child i's height for age standard deviation from a reference median. Similarly, WAZi and WHZi are weight for age and weight for height standard deviation from a reference median for child i. Figure 1.1 shows the distribution of HAZ, WAZ, and WHZ in the BDHS survey sample. All three have a bell-shaped distribution with mean smaller than zero.

According to WHO, moderate to severe stunting, underweight and wasting is defined as height for age, weight for age and weight for height being two standard deviations below the median of the reference population, i.e., HAZ, WAZ and WHZ being less than -2 respectively. In this paper, the binary variables for stunting, wasting, and underweight have been constructed using this definition. According to table 1.2, 52.1 percent of under-five rural children in the pre-1979 cohort of mothers are stunted, while 44.5 percent are underweight, and 14.7 percent are wasted. Prevalence of stunting, wasting, and underweight among under-five children is higher in rural areas compared to urban areas for the pre-1979 mother cohort. This trend persists in the post-1979 mother cohort as well, but the difference between the health measures in rural and urban areas decreases for the post-1979 mother-cohort. Figure 1.2 suggests a parallel trend in mean years of education for the mother in rural and urban areas for the pre-1979 mother cohorts and a shrinking gap between rural and urban mothers' years of schooling in the post-1979 cohorts. Figure 1.3 suggests a parallel trend in the proportion of children stunted in rural and urban areas for pre-1979 mother cohorts and a shrinking gap between rural and urban areas for post-1979 mother cohorts. This parallel pre-trend is present for the corresponding z score, HAZ. Figure 1.3 suggests less of a parallel pre-trend for the proportion of children underweight and WAZ between rural and urban populations. Figures 1.4 do not suggest any parallel trends in the proportion of children wasted and WHZ between rural and urban mothers among the pre-1979 mother cohort. It is important to remember that these graphs do not control for any characteristic of the mother, child, or the geographic area and as such, can only be considered as suggestive evidence regarding the relationship between the FSSSP program and second-generation child health. I do a more formal test of pre-trend in the form of an event study (see table 1.6) which is discussed in detail in the results section.

The female secondary school stipend program in Bangladesh started in 1994 with girls from grades 7 and 9, and by 1996 it included all the girls in grades 6 to 10. The program was rolled out in all rural schools at the same time. The geographically non-staggered nature of implementation within rural areas is problematic for identification. There are, however, some variations that can be exploited such as (i) rural girls born on or after 1979 received stipend for at least a year whereas those who were born before 1979 were in 10th grade or higher in 1994 and did not receive any stipend and (ii) the stipend program was rolled out only in rural schools. I exploit these two sources of variation in a difference in differences design to estimate the effect of mother's years of eligibility for the stipend program on child health. I also examine the effect of mother's education on child health using mother's years of eligibility for the program as an instrument for mother's years of schooling.

One shortcoming of the BDHS dataset used in this study is that the mother's childhood place of residence is not available in the data. Due to the lack of availability of this variable in the survey data, I will use the urban/ rural classification of the current area of residence instead. Justification of this decision is that migration in Bangladesh is overwhelmingly from rural to urban areas (Richard Marshall and Rahman, 2013) and consequently, most mothers who currently live in rural areas grew up in rural areas as well. Given that the stipend program is implemented in rural areas only, the migration from rural to urban areas will lead to an underestimation of any positive effect of the program on mother's years of education and child health.

1.5. Methodology

I estimate the effect of a program aimed at increasing female schooling on their children's health. Specifically, I estimate the effect of the female secondary school stipend program (FSSSP) in Bangladesh on the recipients' children's stunting, underweight, and wasting prevalence as well as their height for age, weight for age, and weight for height deviations from reference median (z scores). The relationship of interest is represented in equation (i).

$$y_{ijdt} = \beta_o + \beta_1 rural_j + \beta_2 MyobFE_j + \beta_3 Myoe_j + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (i)$$

Here, y_{ijdt} represents health outcome (stunting, wasting, underweight, HAZ, WAZ or WHZ) for child i with birth year t to mother j in district d. $rural_j$ is a dummy variable that takes the value 1 if the mother j of child i lives in a rural area. $Myoe_j$ is the mother j's years of schooling, the independent variable of interest. $MyobFE_j$, μ_d, and λ_t are vectors of mother's year of birth fixed effects, district fixed effects, and child's year of birth fixed effects. Bangladesh is divided into 64 districts, and each district has both urban and rural areas. In equation (i) the coefficient of interest is β_3.

However, due to the possibility of an omitted variable induced endogeneity in the relationship between mother's education and child health, there is no causal interpretation of β_3. Mother's years of education may be correlated with an unobservable ability, which also affects mother's nurturing of children and child health. This gives rise to an endogeneity problem stemming from omitted variable bias in estimating the effect of mother's education on child health. I address the endogeneity issue by exploiting geographic and birth cohort variation in the eligibility for the stipend program in a difference in differences framework.

1.5.1. Difference in Differences Analysis

A girl's eligibility for the stipend program depends on two factors, being born after 1978 and living in a rural area. This gives rise to a difference in differences identification strategy where the treatment group is the rural mothers, and the control group is the urban mothers while the pre-period is the mother's birth cohort before 1979, and the post-period is the birth cohort including and after 1979.[10] Equation (ii) lays out the DID specification for the reduced form relationship.

$$y_{ijdt} = \beta_o + \beta_1 rural_j + \beta_2 MyobFE_j + \beta_3 rural_j * elig\_years_j + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (ii)$$

Like equation (i), y_{ijdt} represents a health outcome of child i with birth year t to mother j in district d. $rural_j$ is a dummy variable that takes the value 1 if the mother j of child i lives in a rural area. $MyobFE_j$, μ_d, and λ_t are vectors of mother's year of birth, district, and child's year of birth fixed effects. $elig\_years_j$ is the years of eligibility for the stipend program of mother j of child i. $elig\_years_j$ is determined by the mother's year of birth. As shown in Table 1.1, mothers born in rural areas before 1979 were in grade 10 or higher in 1994 when the program was introduced, so they did not receive stipend under FSSSP[11]. Due to the staggered nature of the implementation of the stipend program, rural girls born in or after 1983 received all five years of stipend. Girls born between 1979 and 1982 (inclusive) receive varying durations of stipend ranging from 2 to 4 years. The variable $elig\_years_j$ will take the value 0 if the mother was born in or before 1978, 5 if the mother was born in or after 1983 and 2 to 4 for mothers born between 1979 and 1984 (inclusive). The coefficient of interest in equation (ii) is β_3, which represents the percentage point change in a child health outcome due to one more year of mother's eligibility for the stipend program for

[10] See table 1 for a breakdown of mother's years of eligibility by birth cohorts
[11] In 1994 girls from only grades 7 and 9 received stipend.

binary outcome variables (stunting, wasting, and prevalence of being underweight). β_3 represents the standard deviation change in a child health outcome due to one more year of mother's eligibility for the stipend program for continuous outcome variables (HAZ, WAZ, and WHZ). The model controls for district and child's year of birth fixed effect. The district fixed effects (μ_d) control for persistent differences between districts in medical technology, availability of family planning program etc. while the child year of birth fixed effect (λ_t) controls for differences in medical technology and family planning programs over the years in all districts. Following Currie and Moretti (2003), I also estimate a model with a district by child year of birth fixed effect which controls for district-specific trends in medical technology, family planning programs and other characteristics that may affect child health which does not change my results by much (See appendix tables A1.5, A1.6, and A1.7).

Here identification relies on the fact that the program only affected mothers from rural areas and that it only affected mothers in the post-1978 birth cohorts. The identifying assumption in this reduced form analysis is that conditional on the controls, if the FSSSP did not take place in 1994 affecting post-1978 mother birth cohorts, then the health of children born to rural and urban mothers would not be trending differently. This means that for the results to be causally interpretable, it has to be the case that in the absence of FSSSP, children born to rural and urban mothers would have parallel trends in health outcomes, conditional on relevant controls. Though there is no way of directly verifying this assumption, we can check if the trend in health outcomes of the children of rural and urban mothers born before 1979 are the same. The pre-trend in outcome variables, i.e., percentage stunting, percentage wasting, and percentage underweight as well as HAZ, WAZ, and WHZ in rural and urban areas in Bangladesh can be seen in Figure 1.3, 1.4 and 1.5. Figure 1.3 shows parallel pre-trends for stunting and HAZ. A more formal examination of the pre-trend is presented in table 1.6 in the form of an event study. The events study results are discussed in detail in the results section.

This identification strategy is not without shortcomings. If there was a rural area specific program to improve child health after 1994 that differentially affected rural and urban child health and that differential changed over time, then my estimates may be biased. If such a program improved child health, my strategy would attribute the effect of that program to FSSSP. In a different context, one solution would be to include rural-by-mother's year of birth fixed effect to control for any time varying characteristics of rural areas with respect to urban areas that affect child health. But in this context, such a control variable would absorb all of the identifying variation. Nevertheless, any time invariant characteristic that affects rural child health, any district specific time invariant characteristic that affects child health, and any time varying national component of health care that affects child health are controlled for using rural, district and year fixed effects.

1.5.2. Two-Stage Least Squares

The difference in differences analysis unveils the effect of mother's eligibility for FSSSP on their children's health but not the effect of mother's education on child health in general. To estimate the effect of mother's education on child health, I use a mother's years of eligibility for the stipend program as an instrument for her years of education. Here, I follow Currie and Moretti (2003) who use college availability on the 17th year of a girl as an instrument for schooling. Equations (iii) and (iv) are the first and second stages of a 2SLS design where a mother's years of schooling is instrumented by her years of eligibility for the stipend program.

First Stage: $Myoe_{ijdt} = \alpha_o + \alpha_1 rural_j + \alpha_2 MyobFE_j + \alpha_3 rural_j * elig\_years_j + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (iii)$

Second Stage: $y_{ijdt} = \beta_o + \beta_1 rural_j + \beta_3 MyobFE_j + \beta_3 \widehat{\mathrm{Myoe}}_{\mathrm{j}} + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (iv)$

Here, all the notations from equation (ii) are preserved. The coefficient of interest in this model is β_3. For binary outcome variables (stunting, wasting, or prevalence of being underweight), β_3 represents the percentage point change in outcome due to one more year of mother's schooling. For continuous outcome variables (height for age, weight for age or weight for height), β_3 represents the standard deviation change in outcome due to one more year of mother's schooling.

Just like the difference in differences specification, the IV specification includes district and child's year of birth fixed effect. I also run a version of the model with a district by child year of birth fixed effect which controls for district-specific trends in medical technology, family planning programs and other characteristics that may affect child health which does not change my results by much (see appendix tables A1.5, A1.6, and A1.7). Just like the difference in differences case, this more demanding specification generates very similar results to the preferred specification with separate district and child year of birth fixed effect.

For the exclusion restriction to hold, eligibility, which is an interaction between the dummy variable $rural_j$ and the continuous variable years of eligibility ($elig\_years_j$), should influence child health only through the mother's education. But it can be argued that the mother being in a rural area may affect child health due to inferior medical service availability compared to an urban area, which in turn affects child health. This problem is ameliorated by including a $rural_i$ control, which absorbs the time-invariant effects of a rural versus urban address of the mother on child health.

However, as in the DID model, any time varying differential change in rural health care compared to urban health care may lead to biased estimates in the 2SLS strategy as well.

1.6. Results

Table 1.3 shows the result of naïve OLS regression of child health outcomes (stunting, underweight, wasting and the z scores) on mother's years of education, controlling for an indicator for mother living in a rural or urban area, mother's year of birth fixed effects, child year of birth fixed effects and district fixed effects. The naïve OLS shows a negative and significant relationship between mother's years of education and the probability of being stunted, underweight, and wasted. An increase in mother's schooling by one year is associated with a decrease in the probability of being stunted, underweight and wasted by 2.4, 2.2, and 0.05 percentage points, respectively. A mother's years of schooling is positively associated with all three z scores. A one-year increase in mother's schooling is associated with increases in height for age, weight for age, and weight for height of the child by 0.078, 0.068, and 0.031 standard deviations, respectively. These relationships, however, may not be causal. An omitted third variable like mother's ability may affect both mother's years of education and child health. This endogeneity problem will render these estimates biased. To solve the endogeneity problem, I use a difference in differences strategy to estimate the effect of eligibility for the stipend program on the second generation's early childhood health.

1.6.1. Difference in Differences

Table 1.4 shows the results from DID model specified in equation (ii). One extra year of a mother's eligibility for the stipend program decreases the probability of the child being stunted by 1 percentage point, which is equivalent to a 2.32 percent decrease in the probability of being stunted relative to the mean prevalence of stunting of 43.1 percent. One more year of mother's eligibility also increases height for age and weight for age by 0.031 and 0.016 standard deviation, respectively. No statistically significant effect of the program was detected on wasting, prevalence of being underweight and weight for height z score (WHZ). The fact that mother's years of schooling leads to an increase in WAZ but does not decrease the probability of being underweight shows that mother's schooling influenced child's weight for age at a margin that does not push underweight children above the threshold. Two possible reasons include mother's years of schooling increases weight-for-age of children who are not underweight or increases weight for age for underweight children but does not push them above the threshold. More investigation is required to check where in the child weight distribution mother's education has an effect.

The analysis in table 1.4 does not differentiate between first birth and subsequent births. However, first births may be different from subsequent births. First time mothers may be more information seeking and may benefit more from education. In later births, lack of schooling or information may be substituted by more experience. Running the same model restricting the sample to first births only leads to a more pronounced effect of mother's education on child health. Table 1.5 shows that considering only first births, one more year of mother's eligibility for the stipend program leads to 2.3 percentage points decrease in the probability of stunting as well as a 1.1 percentage point decrease in the probability of being underweight. Also, one more year of mother's schooling leads to 0.075 and 0.052 standard deviation increase in height for age and weight for age. So, first births benefit more from a mother being eligible for the program.

Table 1.6 shows the prep period coefficients from an events study model. This is a formal test for the differences in child health trends between rural and urban mothers for mother birth cohorts before 1978. In the event study, I run a version of the DID regression model specified in equation (ii) where the "*rural*elig_years*" variable is replaced with a vector of interactions of mother's birth cohort dummies (from 1965 till 1997) and the "rural" dummy. 1978 is the last birth year for mothers before treatment started. I treat 1978 as the base year in this event study by omitting it from the regression equation. Since the eligible cohort for FSSSP are mothers who were born in 1979 and later, ideally, we want to see no more of the pre-1979 coefficients significant than we should expect by chance. As can be seen from the first three columns of Table 1.6, for stunting, underweight, and wasting, there are only three significant coefficients out of 39 (7.69 percent) in the pre-1979 mother cohorts. Moreover, an F-test for the joint significance of all the pre-period coefficients for the interactions shows insignificant results for stunting and underweight. This indicates that there was no significant difference in childhood stunting and underweight trends between the rural and urban samples before the intervention, giving credence to the causal relationship between mother's eligibility to the stipend program and childhood stunting and underweight found in tables 1.4 and 1.5. Figure 1.6 shows the result of the event study for stunting.

Looking at columns 4, 5, and 6 of Table 1.6, we see 12 out of 39 (31 percent) pre-1979 coefficients are significant. There are two factors of note here. First, almost all the significant coefficients of the interactions are negative, which is the opposite direction of the anticipated treatment effect. So there was no upward pre-trend in health outcomes for children of pre-1979 mother's birth cohorts. Second, most of the significant coefficients are in the years 1971, 1974, and 1975. Bangladesh became independent after a nine-month-long war in December of 1971. 1974 was the year of the most catastrophic famine in Bangladesh, which disproportionately affected the rural farmers more than the urban population (Mia, 1976; Sen, 1981). Fetal origin hypothesis argues that better nutrition during the gestational period affects child health (Barker, 1990; Almond and Mazumder, 2005; Almond, 2006). There is also emerging literature that argues that health and nutrition inputs to mothers during gestation may affect the birth outcome of their children's children (East et al., 2017). While the famine of 1974 can explain the negative and statistically significant coefficients for HAZ, WAZ, and WHZ for the children of mothers born in 1974 and 1975, the liberation war can explain the negative and statistically significant coefficients for HAZ, WAZ, and WHZ for the children of mothers born in 1971. An F-test for the joint significance of the pre-period coefficients for all the interactions other than the years of war (1971) and famine (1974, 1975) shows insignificant result for HAZ but significant results for WAZ and WHZ. This means we can be confident in our results for the effect of mother's eligibility on a child's height for age z scores.

The fact that children of rural mothers born in 1971, 1974, and 1975 had significantly worse health compared to children of urban mothers born in same years gives rise to the possibility that the treatment effect I find in the primary DID specification (table 1.4) is picking up a reversion to the mean after the bad child health outcomes from mothers who were born during the war and famine. Appendix table A1.8 shows the DID results with the 1971, 1974, and 1975 mother cohorts excluded. The results are almost the same as my primary DID results in table 1.4. So a reversion to mean does not explain my primary DID results.

1.6.2. Two-Stage Least Squares

The IV strategy answers a related but different question to the DID strategy. DID estimates give us the effect of mother's years of eligibility for the FSSSP program on child health, which is the average treatment effect. The IV estimates, on the other hand, give us the effect of mother's years of education on child health for mothers who were induced by the program to acquire more years of education. The validity of IV depends on the strength of the first stage and the exclusion restriction. The strong positive and significant relationship between years of mother's education and years of eligibility conditional on other covariates specified in equation (iii) can be seen in table 1.7. Moreover, figure 1.2 shows that the treated (rural) mothers have a steeper trend in years of schooling in the post-period (mother birth cohorts after 1979) compared to the pre-periods (mother birth cohorts before and including 1978). Table 1.8 presents the result of the 2SLS IV design specified in equations (iii) and (iv). One extra year of schooling leads to a 3.6 percentage point decrease in the probability of stunting. This is equivalent to an 8.4 percent decrease in the probability of stunting relative to the mean prevalence of stunting in the sample of 43.1 percent. One more year of mother's education also leads to 0.11 and 0.06 standard deviation increase in height for age and weight for age, respectively. Like the DID specification, I find no statistically significant effect of mother's education on the prevalence of being underweight but a positive and significant effect on the weight for age z score (WAZ).

Estimating the same model with first births only (table 1.9) generates a more pronounced effect of mother's education on child health. Considering only 1st births, one more year of mother's education leads to a 5.1 percentage point decrease in the probability of stunting as well as a 2.4 percentage point decrease in the probability of being underweight. Also, one more year of mother's schooling leads to 0.16 and 0.11 standard deviation increase in height for age and weight for age. So, first births benefit more from an increase in mother's years of education.

1.7. Robustness Checks

I estimate several variants of the DID and 2SLS models as robustness checks. The first robustness check I do is a placebo treatment time assignment. Instead of 1994, I assume 1988 to be the year the FSSSP program started. The “before” and “after” birth cohorts are pushed back six years as well. I treat the birth cohorts from 1973 till 1978 as “after” sample and birth cohorts before 1973 as “before” sample. In the primary analysis, birth cohorts including and after 1979 were considered “after” cohorts and birth cohorts before 1979 were considered “before” cohort. I exclude birth cohorts, including and after 1979, which is the true “after” sample. Table 1.11 shows that eligibility for the placebo program does not have any statistically significant effect on stunting, wasting, the prevalence of being underweight, or the corresponding HAZ, WAZ, and HWZ scores. Also, Table 1.12 shows that using eligibility for the placebo program as an instrument for mother's years of education, mother’s education does not have any effect on the health of their children. These results show that the identification strategy used in this paper is robust to placebo treatment time and give credence to the primary results presented in tables 1.4 and 1.8.

Mother’s age at childbirth is a potential mechanism through which a mother’s education can affect child health. There is evidence that more education leads to lower fertility (Osili and Long, 2008). Moreover, a lower number of children is associated with healthier children (Becker and Lewis, 1973). Having both mother’s year of birth fixed effect and child year of birth fixed effect in the regression models essentially controls for mother’s age at childbirth as mother’s age at childbirth is a linear combination of mother’s year of birth and child’s year of birth. To avoid controlling for this possible mechanism, I use two different strategies. First, I run alternative DID and 2SLS regressions where I substitute the child’s year of birth fixed effect with a survey year fixed effect. A survey year fixed effect still controls for existing medical technology, access to health care, and other time-variant aspects that may affect child health. However, by avoiding controlling for the child’s year of birth fixed effect, I no longer control for mother’s age at childbirth. Appendix table A1.4 and A1.5 show that this alternative specification yields similar results to my primary DID and IV specifications (see tables 1.4 and 1.8).

A second way I avoid controlling for mother's age at childbirth is by using mother's years of potential eligibility (elig_years) based on her birth cohort instead of the mother's year of birth fixed effect in the primary specification. This alternate specification is laid out in equation (v) for a DID and equations (vi) and (vii) as 2SLS.

$$y_{ijdt} = \beta_o + \beta_1 rural_j + \beta_2 elig\_years_j + \beta_3 \mathrm{rural_j} * \mathrm{elig\_years_j} + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (v)$$

First Stage: $$Myoe_{ijdt} = \alpha_o + \alpha_1 rural_j + \alpha_2 elig\_years_j + \alpha_3 rural_j * elig\_years_j + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (vi)$$

Second Stage: $$y_{ijdt} = \beta_o + \beta_1 rural_j + \beta_3 elig\_years_j + \beta_3 \widehat{\mathrm{Myoe}}_{\mathrm{j}} + \mu_d + \lambda_t + \varepsilon_{ijdt} \qquad (vii)$$

Here, Here, *elig_years*$_j$ is years of potential eligibility for the mother. This can take values from 0 to 5 years, depending on a mother's year of birth. Appendix tables A1.6 and A1.7 show that results from these alternate specifications are similar to my primary DID and IV specifications, as presented in tables 1.4 and 1.8.

1.8. Discussion and Conclusion

Child malnutrition and subsequent irreversible effects on cognition, low performance at school, and low wages in the labor market is a cycle that may lead to inter-generational poverty. Countries have been designing various supply-side (better healthcare facilities, increasing the number of antenatal care facilities, increasing the number of trained nurses for childbirth, providing family planning services, etc.) and demand-side (information campaign for family planning, hygiene campaign, etc.) interventions to reduce child malnutrition. In this paper, I present quasi-experimental evidence on the inter-generational health effect of mother's education. I find that one more year of mother's eligibility for a secondary school stipend program leads to a statistically significant 1 percentage point decrease in the child's probability of stunting, which represents a 2.32 percent decrease in the probability of stunting compared to the mean of 43.1 percent. One more year of mother's eligibility for the stipend program also leads to 0.031 and 0.016 standard deviation increase in height for age and weight for age z scores for the child. Using an IV strategy, I find that one more year of schooling for the mother leads to a 3.6 percentage point decrease in the child's probability of stunting. This represents an 8.4 percent decrease in the probability of stunting compared to the mean probability of stunting in the sample. Furthermore, one more year of mother's education increases the child's height for age and weight for age by 0.11 and 0.06 standard deviation, respectively. The results are robust to a falsification test using a placebo treatment time as well as a host of alternative specifications.

My first stage result, i.e., the effect of the stipend program on girls' years of schooling, is consistent with existing literature. I find that one more year of eligibility for FSSSP leads to 0.27 more years of schooling. For the full five years of eligibility for the program, a mother's education would increase by 1.35 years. Hahn et al. (2018) found the same program to increase mother's years of education by 1.21 years. Shamsuddin (2015) finds eligibility for FSSSP led to a 1-year increase in education level. Yasmin (2016) finds the same program to increase mother's years of schooling by 1.14 years. My first stage results are comparable to some other educational interventions as well. My first stage estimate of 1.35 extra years of schooling for five years of eligibility to the stipend program is slightly smaller than the effect of being treated in the school building project in Zimbabwe (led to 1.68 more years of maternal education) found by Grepin and

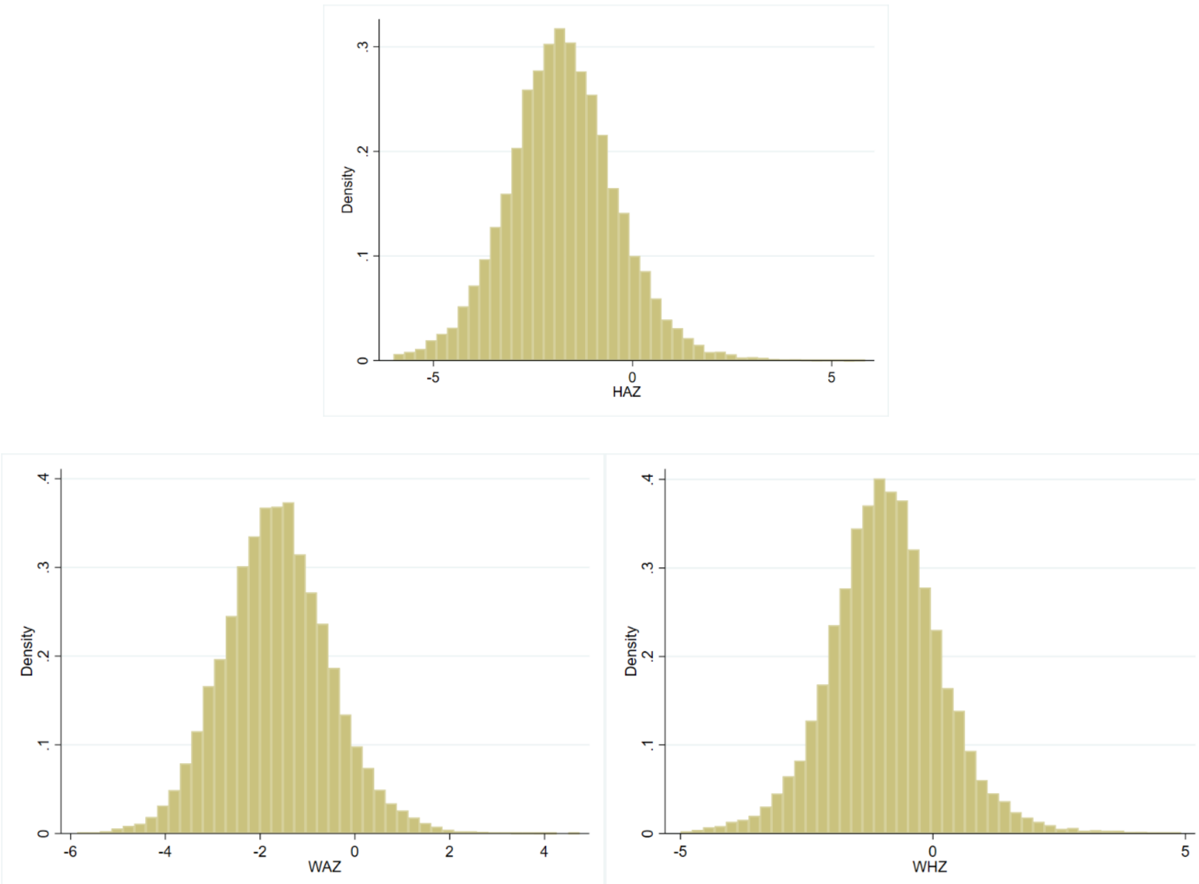
Bharadwaj (2015) and larger than the effect of one new 4-year college opening per 1000 18-22-year-olds (led to 1 more year of maternal education by 1 full year) found by Currie and Moretti (2003).

This is the first paper to estimate the intergenerational effects of an educational intervention on the recipients' children's stunting, wasting, and prevalence of being underweight. I find that the full five years of mother's eligibility for the stipend program leads to 0.155 and 0.08 standard deviation increase in a child's HAZ and WAZ (0.031*5 and 0.016*5). My results are comparable to the 0.143 and 0.106 standard deviation increase in HAZ and WAZ estimated by Hahn et al. (2018) which investigates the same stipend program in Bangladesh. It, however, use fewer years of survey rounds and don't examine stunting, wasting, and prevalence of underweight. Moreover, Hahn et al. (2018) do not provide any robustness check for their intergenerational effects results, thus their results regarding child health are open to alternate explanations. My estimates are higher than the effect of mother completing primary education on a child's HAZ and WAZ (1.1 and 1 standard deviation for HAZ and WAZ) estimated by Gunes et al. (2015).

Given the effectiveness of mother's education in improving child health, the cost-benefit analysis of the FSSSP in specific and education-related CCTs, in general, should include the long run inter-generational health effect of such programs. These results show that mother's education may be an integral part of reaching the WHO Global Targets 2025 and UN Sustainable Development Goals (SDG), both of which include goals of lowering childhood stunting world-wide.

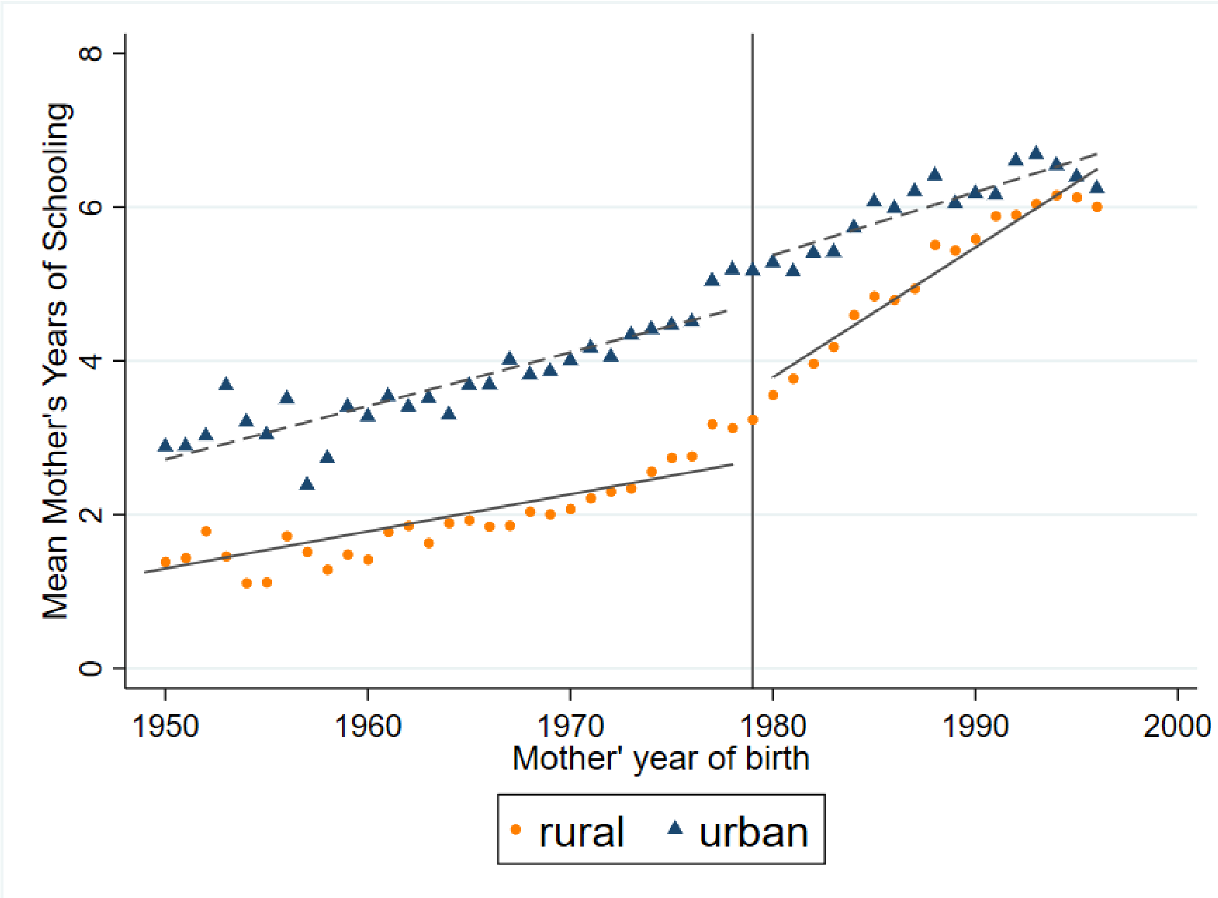
1.9. Figures and Tables

Figure 1.1: Z Score Distributions



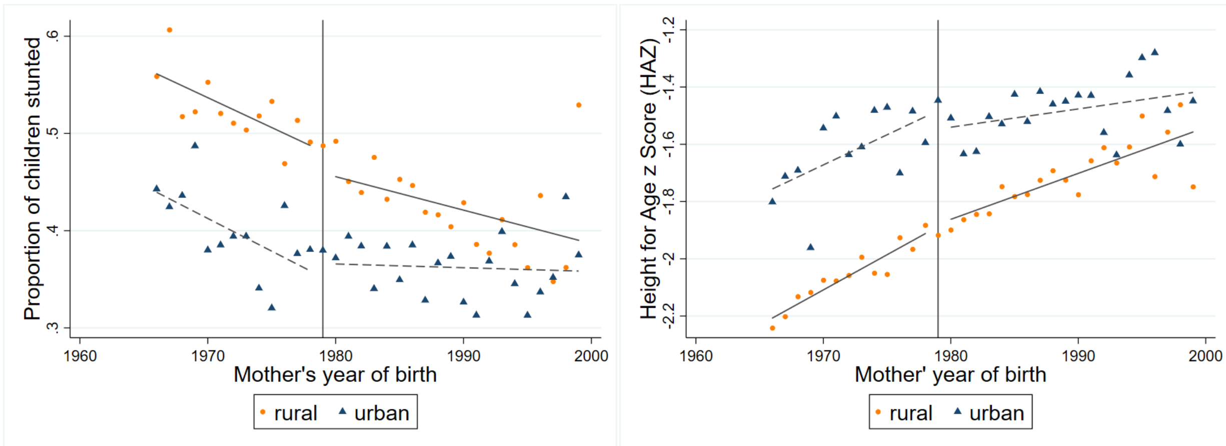
Note: Height for age, weight for age, and weight for height deviation from the reference median. Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before survey.

Figure 1.2: Mother's Years of Education Trend



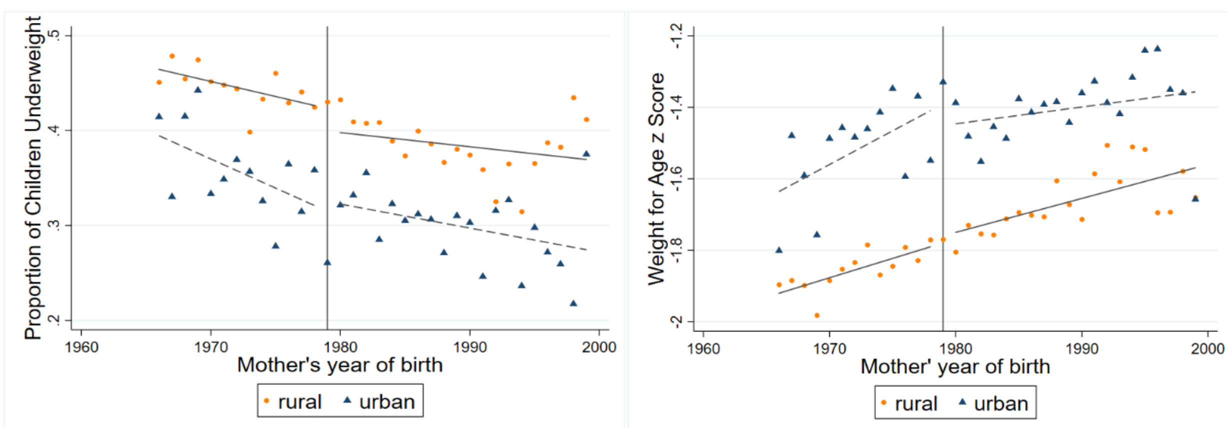
Note: Mother's years of education trend by year of birth. Data are from five rounds of Bangladesh Demographic Health Surveys (BDHS) from 1999 to 2014.

Figure 1.3: Stunting and HAZ Trends



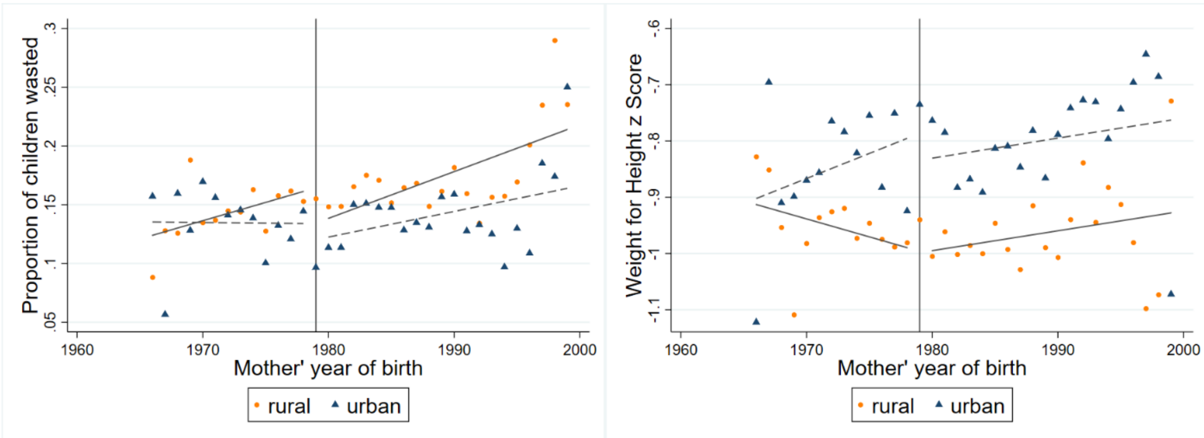
Note: Child stunting and height for age z-score (HAZ) trend. Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey.

Figure 1.4: Underweight and WAZ Trends



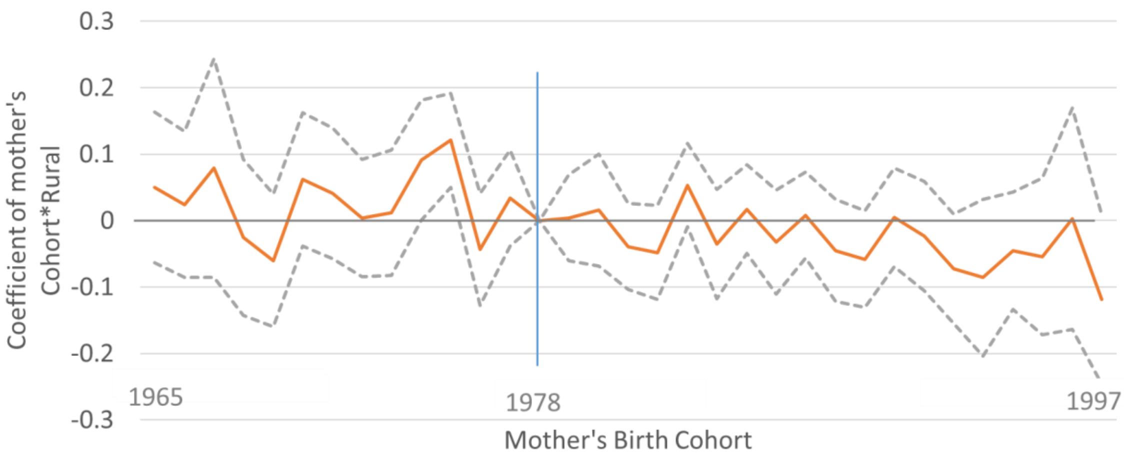
Note: Child prevalence of underweight and weight for age z-score (WAZ) trend. Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey.

Figure 1.5: Wasting and WHZ Trends



Note: Child wasting and weight for height z-score (WHZ) trend. Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey.

Figure 1.6: Event Study



Note: Events study for the outcome stunting. The treatment variable "rural" is interacted with the full set of year dummies from 1965 to 1997 omitting 1978. The base year is 1978, the last birth year for mothers before treatment started.

Table 1.1: Defining Treatment and Control Groups

Grade in 1994	Age in 1994	Year of birth for each age in 1994	Years for which rural girl received stipend if at school	Treatment status
≤ 1	≤ 7	≥ 1987	5	Treated
2	8	1986	5	
3	9	1985	5	
4	10	1984	5	
5	11	1983	5	
6	12	1982	4	Partially Treated
7	13	1981	3	
8	14	1980	2	
9	15	1979	2	
10	16	1978	0	Control
11	17	1977	0	
12	18	1976	0	
13	19	1975	0	
≥ 14	≥ 20	≤ 1974	0	

Note: This table assumes that a child goes to grade 1 at the age of 7 as is the norm in South Asian countries.

Table 1.2: Summary Statistics

Variable	Full Sample	Rural (Treatment) Pre-1979 mother cohort	Urban (Control) Pre-1979 mother cohort	Rural (Treatment) Post 1979 mother cohort	Urban (Control) Post 1979 mother cohort
Outcomes					
HAZ: Height for age	-1.742	-2.043	-1.612	-1.733	-1.459
SD	(1.392)	(1.388)	(1.444)	(1.349)	(1.389)
WAZ: Weight for age	-1.637	-1.825	-1.505	-1.664	-1.383
SD	(1.139)	(1.082)	(1.252)	(1.098)	(1.177)
WHZ: Weight for	-0.915	-0.961	-0.837	-0.961	-0.796
height SD	(1.131)	(1.058)	(1.160)	(1.138)	(1.177)
Stunted[12]	0.431	0.521	0.395	0.425	0.354
Underweight [13]	0.376	0.445	0.351	0.380	0.296
Wasted[14]	0.149	0.147	0.134	0.162	0.132
Other Variables					
Mother's years of	4.738	2.591	5.141	5.164	6.499
education	(4.131)	(3.475)	(5.032)	(3.671)	(4.125)
Mother's age at	26.63	31.22	31.24	23.73	24.06
interview	(5.984)	(5.980)	(5.489)	(4.024)	(3.996)
Total children of the	2.894	4.287	3.519	2.260	2.019
mother	(1.830)	(2.150)	(1.879)	(1.202)	(1.065)
Mother's age at first	18.07	18.23	29.38	17.56	18.22
child birth	(3.192)	(3.512)	(4.256)	(2.551)	(3.012)
Mother's age at child	23.74	28.00	27.96	21.06	21.34
birth	(5.724)	(5.944)	(5.481)	(3.799)	(3.791)
Marriage to first birth	28.30	35.99	33.11	23.85	
interval	(25.200)	(32.18)	(30.07)	(3.799)	24.76
					(19.99)
Birth order number	2.661	3.990	3.300	2.036	1.844
	(1.776)	(2.128)	(1.845)	(1.140)	(1.003)
N	29,553	7,368	3,060	13,123	6,002

Note: Standard deviation in parenthesis. Data are from 5 Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. Sample includes children from mothers that gave birth at ages 16 to 46 and had a child less than 5 years old on the day of survey.

[12] Equals 1 if HAZ<-2 i.e. height is 2 SD lower than standard set by WHO
[13] Equals 1 if WAZ<-2 i.e. weight is 2 SD lower than standard set by WHO
[14] Equals 1 if WHZ<-2 i.e. weight is 2 SD lower than standard set by WHO

Table 1.3: Naive OLS

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.024***	-0.022***	-0.005***	0.078***	0.068***	0.031***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
Rural	0.044***	0.042***	0.015**	-0.169***	-0.161***	-0.082***
	(0.008)	(0.008)	(0.006)	(0.027)	(0.019)	(0.022)
Observations	31,170	31,170	31,170	31,170	31,170	31,170
R-squared	0.106	0.072	0.017	0.145	0.121	0.036

Note: Data are from 5 Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within 5 years before survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.4: DID

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's years of eligibility	-0.010***	-0.003	0.001	0.031***	0.016**	-0.003
	(0.003)	(0.003)	(0.002)	(0.009)	(0.006)	(0.006)
Rural	0.106***	0.081***	0.018**	-0.368***	-0.302***	-0.115***
	(0.013)	(0.010)	(0.007)	(0.040)	(0.026)	(0.029)
Observations	31,200	31,200	31,200	31,200	31,200	31,200
R-squared	0.074	0.044	0.014	0.103	0.073	0.026

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.5: DID: First Births Only

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's years of eligibility	-0.023***	-0.011*	-0.001	0.075***	0.052***	0.009
	(0.006)	(0.006)	(0.005)	(0.015)	(0.013)	(0.013)
Rural	0.165***	0.107***	0.036*	-0.564***	-0.470***	-0.183***
	(0.027)	(0.026)	(0.021)	(0.065)	(0.059)	(0.060)
Observations	10,348	10,348	10,348	10,348	10,348	10,348
R-squared	0.092	0.058	0.023	0.132	0.107	0.039

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The table includes only the first child born during the five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.6: Events Study: Pre-Trend

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
1965 Cohort × rural	0.050	0.033	-0.059	-0.489**	-0.220	0.095
	(0.058)	(0.062)	(0.051)	(0.238)	(0.153)	(0.157)
1966 Cohort × rural	0.024	-0.017	-0.077	-0.229	0.078	0.322*
	(0.056)	(0.066)	(0.047)	(0.183)	(0.143)	(0.173)
1967 Cohort × rural	0.079	0.088	0.059	-0.242	-0.208	-0.096
	(0.084)	(0.068)	(0.038)	(0.236)	(0.176)	(0.136)
1968 Cohort × rural	-0.025	-0.021	-0.042	-0.182	-0.107	0.007
	(0.060)	(0.064)	(0.053)	(0.199)	(0.180)	(0.166)
1969 Cohort × rural	-0.060	-0.030	0.048	0.078	-0.019	-0.148
	(0.051)	(0.052)	(0.034)	(0.129)	(0.129)	(0.135)
1970 Cohort × rural	0.062	0.051	-0.048	-0.244*	-0.172	-0.049
	(0.051)	(0.054)	(0.040)	(0.134)	(0.125)	(0.149)
1971 Cohort × rural	0.041	0.041	-0.031	-0.341***	-0.206*	-0.035
	(0.050)	(0.052)	(0.035)	(0.116)	(0.109)	(0.116)
1972 Cohort × rural	0.004	0.007	-0.007	-0.130	-0.129	-0.112
	(0.045)	(0.038)	(0.037)	(0.128)	(0.087)	(0.100)
1973 Cohort × rural	0.012	-0.017	-0.014	-0.138	-0.130	-0.077
	(0.048)	(0.051)	(0.033)	(0.137)	(0.101)	(0.109)
1974 Cohort × rural	0.091*	0.063	0.017	-0.355**	-0.291**	-0.123
	(0.046)	(0.047)	(0.028)	(0.144)	(0.118)	(0.098)
1975 Cohort × rural	0.121***	0.127***	0.014	-0.363***	-0.307**	-0.122
	(0.036)	(0.043)	(0.029)	(0.114)	(0.118)	(0.112)
1976 Cohort × rural	-0.043	0.023	0.022	-0.019	-0.046	-0.070
	(0.043)	(0.041)	(0.031)	(0.135)	(0.105)	(0.096)
1977 Cohort × rural	0.034	0.069	0.032	-0.222*	-0.267***	-0.204**
	(0.037)	(0.046)	(0.034)	(0.122)	(0.098)	(0.102)
F-Test w/o ’71, ’74, ’75	1.21	1.13	1.9**	1.68	2.10**	1.78*
Observations	30,232	30,232	30,232	30,232	30,232	30,232

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Mother’s year of birth restricted from 1965 to 1997. Coefficients for the interaction of mother’s year of birth and the binary variable “rural” reported in the table. Standard errors clustered at the district level in parentheses. Includes mother’s year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.7: First Stage

VARIABLES	Mother's years of education
Years of Eligibility	0.274***
	(0.045)
Rural	-2.186***
	(0.198)
Mean Mother's years of education	4.738
Observations	31,170
F-stat	2593.52
Prob>F	0
R-squared	0.188

Note: Data are from 5 Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. Robust standard errors in parentheses and clustered at the district level. Includes mother's year of birth, child year of birth, and district fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table 1.8: 2SLS

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Stunting	Underweight	Wasting	HAZ	WAZ	WHZ
Education in single years	-0.036***	-0.011	0.004	0.113***	0.060***	-0.011
	(0.008)	(0.009)	(0.006)	(0.022)	(0.019)	(0.020)
Rural	0.028**	0.057***	0.028**	-0.121***	-0.172***	-0.141***
	(0.014)	(0.016)	(0.011)	(0.044)	(0.034)	(0.031)
Observations	31,170	31,170	31,170	31,170	31,170	31,170
R-squared	0.099	0.065	0.007	0.136	0.120	0.017

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.9: 2SLS: First Births Only

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.051***	-0.024*	-0.001	0.163***	0.114***	0.021
	(0.011)	(0.013)	(0.010)	(0.029)	(0.026)	(0.028)
Rural	0.020	0.038**	0.032***	-0.098**	-0.145***	-0.123***
	(0.014)	(0.018)	(0.012)	(0.049)	(0.031)	(0.037)
Observations	10,337	10,337	10,337	10,337	10,337	10,337
R-squared	0.085	0.079	0.024	0.115	0.129	0.048

Note: Data are from 5 Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The table includes only the first child born during the 5 years before survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.10: Naive OLS with Placebo Treatment Time

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.026***	-0.025***	-0.006***	0.082***	0.071***	0.033***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
Rural	0.051***	0.024**	0.002	-0.181***	-0.132***	-0.031
	(0.011)	(0.010)	(0.009)	(0.032)	(0.032)	(0.038)
Observations	10,428	10,428	10,428	10,428	10,428	10,428
R-squared	0.117	0.080	0.027	0.153	0.123	0.046

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes children born within five years before the survey. Instead of the actual year when the treatment started (1994) I consider a placebo treatment year, 1988. In this falsification test, mothers born before 1973 are considered "before" sample and mother born from 1973 till 1978 are considered "after" sample. For this exercise, I get rid of the observations for mother birth cohorts after 1978 which is the true "after" sample. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.11: DID with Placebo Treatment Time

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's years of eligibility	0.001	0.004	0.005	0.013	-0.003	-0.012
	(0.007)	(0.007)	(0.003)	(0.019)	(0.015)	(0.013)
Rural	0.105***	0.067***	0.002	-0.383***	-0.278***	-0.074
	(0.022)	(0.020)	(0.012)	(0.061)	(0.047)	(0.049)
Observations	10,433	10,433	10,433	10,433	10,433	10,433
R-squared	0.079	0.043	0.023	0.104	0.065	0.033

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes children born within five years before the survey. Instead of the actual year when the treatment started (1994) I consider a placebo treatment year, 1988. In this falsification test, mothers born before 1973 are considered "before" sample and mother born from 1973 till 1978 are considered "after" sample. For this exercise, I get rid of the observations for mother birth cohorts after 1978 which is the true "after" sample. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table 1.12: 2SLS with Placebo Treatment Time

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	0.013	0.117	0.151	0.375	-0.079	-0.353
	(0.191)	(0.303)	(0.253)	(0.616)	(0.507)	(0.808)
Rural	0.134	0.326	0.337	0.447	-0.454	-0.855
	(0.405)	(0.639)	(0.539)	(1.331)	(1.081)	(1.724)
Observations	10,428	10,428	10,428	10,428	10,428	10,428
R-squared	0.030					

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes children born within five years before the survey. Instead of the actual year when the treatment started (1994) I consider a placebo treatment year, 1988. In this falsification test, mothers born before 1973 are considered "before" sample and mother born from 1973 till 1978 are considered "after" sample. For this exercise, I get rid of the observations for mother birth cohorts after 1978 which is the true "after" sample. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

1.10. Appendix

Table A1.1: Naive OLS with Child Year of Birth - District Fixed Effect

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.025***	-0.022***	-0.005***	0.079***	0.068***	0.030***
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
Rural	0.041***	0.042***	0.018***	-0.158***	-0.159***	-0.089***
	(0.008)	(0.008)	(0.006)	(0.028)	(0.019)	(0.022)
Observations	31,170	31,170	31,170	31,170	31,170	31,170
R-squared	0.142	0.111	0.056	0.179	0.157	0.081

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The table includes all children born within five years before the survey. Robust standard errors in parentheses and clustered at the district level. Includes mother's year of birth, child year of birth fixed effect, district fixed effect and child year of birth by district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.2: DID with Child Year of Birth - District Fixed Effect

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's years of eligibility	-0.009***	-0.004	0.000	0.028***	0.019**	0.003
	(0.003)	(0.003)	(0.002)	(0.010)	(0.008)	(0.006)
Rural	0.103***	0.086***	0.023***	-0.351***	-0.309***	-0.139***
	(0.015)	(0.011)	(0.008)	(0.043)	(0.028)	(0.031)
Observations	31,200	31,200	31,200	31,200	31,200	31,200
R-squared	0.110	0.084	0.054	0.137	0.111	0.072

Note: Data are from Mother's years of eligibility Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The table includes all children born within five years before the survey. Robust standard errors in parentheses and clustered at the district level. Includes mother's year of birth, child year of birth fixed effect, district fixed effect, and child year of birth by district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.3: 2SLS_cont_elig with Child Year of Birth - District Fixed Effect

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.031***	-0.014*	0.001	0.094***	0.063***	0.010
	(0.008)	(0.008)	(0.007)	(0.024)	(0.021)	(0.021)
rural	0.032**	0.053***	0.026**	-0.137***	-0.166***	-0.117***
	(0.013)	(0.015)	(0.011)	(0.044)	(0.036)	(0.031)
Observations	31,170	31,170	31,170	31,170	31,170	31,170
R-squared	0.140	0.108	0.052	0.177	0.157	0.077

Note: Data are from 5 Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The table includes all children born within 5 years before survey. Robust standard errors in parentheses and clustered at the district level. Includes mother's year of birth, child year of birth fixed effect, district fixed effect and child year of birth by district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.4: DID with Survey Year Instead of Child's Year of Birth

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's Years of Eligibility	-0.010***	-0.003	0.001	0.031***	0.016**	-0.004
	(0.003)	(0.003)	(0.002)	(0.009)	(0.006)	(0.005)
rural	0.105***	0.081***	0.018**	-0.366***	-0.303***	-0.115***
	(0.014)	(0.010)	(0.008)	(0.043)	(0.026)	(0.029)
Observations	31,200	31,200	31,200	31,200	31,200	31,200
R-squared	0.038	0.030	0.010	0.050	0.048	0.019

Note: Data are from Mother's years of eligibility Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, survey year, and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.5: 2SLS with Survey Year Instead of Child's Year of Birth Fixed Effect

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.035***	-0.011	0.004	0.112***	0.058***	-0.013
	(0.009)	(0.009)	(0.007)	(0.023)	(0.019)	(0.020)
Rural	0.028**	0.058***	0.028**	-0.122***	-0.176***	-0.144***
	(0.014)	(0.016)	(0.011)	(0.042)	(0.034)	(0.031)
Observations	31,170	31,170	31,170	31,170	31,170	31,170
R-squared	0.066	0.051	0.003	0.089	0.098	0.009

Note: Data are from Mother's years of eligibility Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, survey year, and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.6: DID with Years of Potential Eligibility

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's years of eligibility	-0.010***	-0.003	0.001	0.030***	0.016**	-0.003
	(0.003)	(0.003)	(0.002)	(0.009)	(0.006)	(0.006)
Rural	0.105***	0.081***	0.019**	-0.365***	-0.302***	-0.116***
	(0.014)	(0.010)	(0.008)	(0.042)	(0.027)	(0.028)
Observations	31,200	31,200	31,200	31,200	31,200	31,200
R-squared	0.067	0.039	0.012	0.093	0.067	0.024

Note: Data are from Mother's years of eligibility Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes child's year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.7: 2SLS with Years of Potential Eligibility

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.034***	-0.010	0.004	0.107***	0.057***	-0.011
	(0.008)	(0.009)	(0.006)	(0.022)	(0.019)	(0.020)
Rural	0.030**	0.058***	0.028**	-0.127***	-0.176***	-0.140***
	(0.014)	(0.016)	(0.011)	(0.043)	(0.034)	(0.031)
Observations	31,170	31,170	31,170	31,170	31,170	31,170
R-squared	0.094	0.060	0.006	0.130	0.115	0.015

Note: Data are from Mother's years of eligibility Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes child's year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1

Table A1.8: DID Excluding Mother Birth Cohorts 1971, 1974 and 1975

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Mother's years of eligibility	-0.009***	-0.003	0.001	0.030***	0.016**	-0.004
	(0.003)	(0.003)	(0.002)	(0.009)	(0.006)	(0.005)
Rural	0.104***	0.080***	0.019**	-0.363***	-0.299***	-0.114***
	(0.014)	(0.010)	(0.007)	(0.041)	(0.026)	(0.029)
Observations	31,200	31,200	31,200	31,200	31,200	31,200
R-squared	0.069	0.042	0.014	0.096	0.070	0.026

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. Excludes mother birth cohorts 1971, 1974 and 1975 to test for mean reversion from the bad outcomes of mother birth cohorts from the year of liberation war (1971) and the years of famine (1974, 1975). The sample includes all children born within five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1.

Table A1.9: Naive OLS: First Births Only

VARIABLES	(1) Stunting	(2) Underweight	(3) Wasting	(4) HAZ	(5) WAZ	(6) WHZ
Education in single years	-0.024***	-0.021***	-0.005***	0.074***	0.069***	0.035***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
Rural	0.044***	0.041***	0.028***	-0.180***	-0.187***	-0.110***
	(0.011)	(0.011)	(0.008)	(0.039)	(0.026)	(0.027)
Observations	10,337	10,337	10,337	10,337	10,337	10,337
R-squared	0.119	0.079	0.026	0.165	0.146	0.049

Note: Data are from five rounds of Bangladesh Demographic Health Survey (BDHS) from 1999 to 2014. The table includes only the first child born during the five years before the survey. Standard errors clustered at the district level in parentheses. Includes mother's year of birth, child year of birth and district fixed effect. *** p<0.01, ** p<0.05, * p<0.1.

2. College Credit on the Table? Advanced Placement Course and Exam Taking[15]

2.1. Introduction

Taking advanced college-level coursework in high school is pervasive across the United States. Among the public high school graduates in the class of 2019 nationally, 1,245,527 (38.9 percent) took at least one Advanced Placement (AP) exam, which is more than a 50 percent increase over the last ten years.[16] Between the 2002-03 and 2010-11 academic years, the number of students taking college-level courses in a dual-enrollment program increased 80 percent to 1.2 million.[17] While there are many arguments in favor of (and in opposition to) such advanced course work, one potential benefit is the ability to earn college credit while still in high school. Earning college credit while in high school is related to numerous positive collegiate outcomes, including performance in college and college graduation (Dougherty et al., 2006; Morgan and Klaric, 2007; An, 2012; Allen and Dadgar, 2012; Patterson and Ewing, 2013).

This paper addresses the simple but novel research question—do students who take Advanced Placement (AP) courses take the corresponding AP exams or are they leaving college credit on the table? For more than 30 AP courses, students can obtain college credit by performing well on the corresponding AP exam at the end of the school year. College credit through AP exams causally results in a higher probability of graduating college in four years (Smith et al., 2017). This is especially important since many college students struggle to graduate and graduate on time (Bound et al., 2010; Denning, et al., 2019). Scoring high enough on an AP exam to earn college

---

[15] This chapter is co-authored with Todd Jones (toddrandyjones@gmail.com), Department of Economics, Mississippi State University, and Jonathan Smith (jsmith500@gsu.edu), Department of Economics, Georgia State University

[16] Source: College Board website. Link: https://reports.collegeboard.org/ap-program-results/class-2019-data

[17] Source: National Center for Education Statistics. Link: https://nces.ed.gov/pubs2013/2013002/tables/table_01.asp

credit also induces more advanced coursework in high school (Smith et al., 2017), more females taking upper level STEM courses in college (Gurantz, 2019), and increases the probability the student's college major will be in the AP subject (Avery, et al., 2018). Collectively, taking an AP exam and performing well substantially impacts students' trajectories through college.

There are a number of reasons students who take an AP course may not take the corresponding AP exam, which motivate our analyses. First and foremost, the costs of taking the AP exam may outweigh the benefits. The exam fees constitute a direct financial cost, which can be as high as $91. Moreover, there may be indirect costs such as time spent studying and sitting for the exam as well as psychological costs of a high stakes exam (e.g., Banks and Smyth, 2015). All these costs, especially the exam fee, likely motivate states like Georgia to subsidize at least one AP exam for students eligible for free and reduced-priced lunch (FRL). Many Georgia school districts, including all four we explore, go above and beyond the state subsidy by making some (or all) AP exams free for some (or all) students. On the benefit side, students may know they are likely to perform poorly on the exam, resulting in few if any benefits (i.e., college credit) and so even small costs may outweigh the benefits.

The second set of reasons students who take an AP course may not take the AP exam stems from a mountain of evidence on the behavioral and informational constraints that lead to undesirable educational outcomes on the path to and through college.[18] On the behavioral side, some students are dissuaded by small costs and procedures in the college application process, even if the benefits actually outweigh the costs, such as taking the SAT or ACT (Klasik, 2013; Hurwitz et al., 2015; Goodman, 2016; Hyman, 2017) or paying a small college application fee and writing an admission essay (Smith, et al., 2015). In this context, an exam fee or a three-hour exam may be

[18] See Page and Scott-Clayton (2016) for a thorough review of the evidence.

sufficient to overlook the benefits of performing well. On the informational side, disadvantaged students may not have the same resources, peers, and schools as relatively advantaged students to help them navigate educational procedures and decisions (e.g., Bettinger, et al., 2012; Hoxby and Turner, 2013; Dillon and Smith, 2017). Here, some AP course enrollees may not know about college credit policies or how they can benefit from college credit, should they score high enough.

Using data from four metro-Atlanta public school districts from school year (SY) 2014-15 to SY 2016-17, we find that 15 percent of students' AP course enrollments do not result in an AP exam. We also estimate that up to 32 percent of the AP courses that do not lead to exams would receive scores of 3 or higher, generally corresponding to college credit. To our knowledge, we are the first paper to document these simple results on AP exam-taking rates. This also represents a potential actionable policy lever—incentives for students to take the AP exam—that corresponds to the massive growth in advanced coursework enrollment over the last decade.

Motivated by the previously mentioned literature, we also examine AP exam-taking rates by different student subgroups and find substantial disparities in AP exam-taking rates between traditionally disadvantaged populations and relatively advantaged students. Eighteen percent of the courses taken by FRL students do not lead to an exam compared to 15 percent for non-FRL students. Black students take an AP course but do not take the AP exam 23 percent of the time, compared to 10 and 13 percent for Asian students and White students, respectively. The rates are 18 and 15 percent for Hispanic and non-Hispanic students, respectively.[19]

Next, we estimate multiple linear regressions to assess whether the above unconditional statistics on AP exam taking are driven by factors correlated with FRL status and race/ethnicity or

[19] Ethnicity and race are not mutually exclusive categories in these data.

by those measures in and of themselves. We also investigate potential policy levers to increase AP exam-taking rates. We find that after adding a rich set of controls—including AP course grades—courses taken by FRL students are 2 percentage points more likely to result in an AP exam than those taken by non-FRL students. FRL students typically have worse educational outcomes than non-FRL students, on average (e.g., Papay et al., 2015), making this an uncommon result in the literature. The districts that provide a higher AP exam subsidy for FRL students than non-FRL students drive the positive relationship, with courses taken by FRL students being 3 percentage points more likely to result in an AP exam. In the districts with high AP exam subsidies but not differentially so by FRL status, we see no difference in exam-taking rates by FRL status. Taken together and with a series of other analyses, the evidence is consistent with these AP exam subsidy policies leading to an increase in AP exam taking for FRL students. These are the very students who are targeted by the subsidies and are often subject to financial and informational constraints.

After adding a rich set controls to our regressions, we still see troubling racial disparities in AP exam-taking rates. Courses taken by Black students are almost 4 percentage points less likely to take an AP exam compared to their White peers, even when accounting for differences in AP course grades, FRL status, and high school enrolled. We also find that courses taken by Hispanic students are almost 1.5 percentage points less likely to take the exam compared to non-Hispanic students. Our analyses do not provide evidence as to why the Black-White disparity persists in this context and find this to be a compelling place for further research and school intervention.

Next, we turn to analyses on gender; it is possible that AP courses and exams can both highlight potentially new measures of gender disparity and serve to exacerbate or reduce downstream gender disparities, such as those disparities documented and explored in college major and occupation (e.g., Blau and Kahn, 2000; Speer, 2017). We first note that there are gender gaps in AP course enrollment in our four school districts, despite gender parity in composition of the student body. Females comprise of only 22 and 24 percent of the two AP computer science courses and 22, 31, and 39 percent of the three AP physics courses. Females are substantially overrepresented in AP Psychology and the two AP English courses.

With these AP course enrollment statistics in mind, we investigate whether there is a lower AP exam-taking rate by females, especially in AP courses where female students are under-represented. We also explore if a female student being paired with female teachers leads to more AP exam-taking. Importantly, we do not have random assignment of students into AP courses nor is there random assignment of instructors to classrooms, which is how the most compelling literature on teacher and student gender match ("role model effects") proceed, albeit with mixed results.[20] We find no evidence that females are less likely to take an AP exam, even in AP courses where they are the minority. We also find a null effect of a female student being paired with a female teacher on the probability of taking the AP exam; we also fail to detect an effect even in AP subjects in which females are underrepresented. Although our estimates are not causal, we suspect that selection issues for the gender match of the student and instructor would bias our results upward, and so a lack of statistically positive coefficients is suggestive evidence against a role model effect in AP exam taking.

Finally, we examine two other factors that potentially affect AP exam taking. First, although twelfth-grade students are the least likely grade to take an AP exam, the impact of exam

[20] While the effect of having a female teacher on female student's performance in middle or high school (e.g., Ehrenberg et al., 1995; Nixon and Robinson 1999; Dee 2005, 2007; Winters et al., 2013; Gon et al., 2018) and college (e.g., Canes and Rosen 1995; Rothstein 1995; Neumark and Gardecki 1996; Bettinger and Long 2005; Hoffmann and Oreopoulos 2009; Carrell et al., 2010) is either undetectable or positive, the evidence for elementary school is mixed. Most of the studies estimating the effect of female students having female teachers in elementary school find a positive effect on test scores (e.g., Winters et al., 2013) with one notable exception (Antecol et al., 2015). Antecol et al. (2015) finds that having a female teacher in primary school lowers math test scores of female students in disadvantaged neighborhoods.

subsidies appear to be largest for these students. Second, we find some evidence that taking relatively more AP courses reduces the probability of taking all AP exams. Combined, these results suggest that the timing and number of exam subsidies enter the calculus of whether to take an AP exam; policymakers and educators can take this into account if they are looking to increase taking rates.

Overall, our results highlight the fact that not all students who take AP courses take the AP exam, including some of whom we predict would score a 3 or higher and earn college credit. Although we only focus on a few potential ways to induce AP exam taking, we provide evidence consistent with a positive effect of exam subsidies on AP exams taking, thereby suggesting that students can be incentivized to take AP exams. This may be particularly important given the underlying AP exam-taking disparities, especially for Black students with similar academic credentials as their White peers. These results build on a rich literature about AP that typically focuses on AP course enrollment, spanning inequality in access and enrollment (e.g., Solórzano and Ornelas, 2002; Klopfenstein, 2004), impacts on learning (e.g., Conger et al., 2019), and impacts on college enrollment (e.g., Jackson, 2010). In an era where school districts (and researchers) are fully immersed in advanced coursework, getting students over the last hurdle to take an AP exam may be a relatively straightforward policy lever.

2.2. Data and Setting

2.2.1. Advanced Placement

Since 1955, Advanced Placement (AP) gives high school students an opportunity to take college level courses while in high school and potentially earn college credit. There are currently 38 different AP courses in seven different subject categories: History and Social Sciences, English, Science, Math and Computer Science, World Language and Culture, Arts, and AP Capstone. Each AP course is designed to cover the concepts from the corresponding introductory college level course. Schools must be accredited by the AP Course Audit to offer AP courses.

At the end of an AP course in May, students can take the corresponding AP exam, which typically lasts between two to three hours. The exam is designed by college faculty and high school teachers to test a mastery of the subject matter at a college level. Exam scores range from 1 to 5 in integer values. The exams are only offered once a year, so retaking is rare.

Most universities provide credit for an AP exam score of 3 or above, but there is considerable variation in credit granting across colleges and AP subjects. For example, Appendix Table A2.1 shows that credit granting AP exam scores for the ten most popular AP courses in Georgia at some of the largest public universities in the state varied between 3 to 5; however, the modal score is a 3 and the most academically selective public institution in the state accepts mostly 4s. In lieu of college credit, some colleges use AP scores for placement in and out of certain courses.

2.2.2. Sample Overview

Our analyses consider four large metro-Atlanta districts from SY 2014-15 to SY 2016-17 using administrative, student-level data from the Metro Atlanta Policy Lab for Education (MAPLE), a collaboration between academic researchers and several Atlanta-area public school districts. We refer to these districts as Districts 1, 2, 3, and 4.[21] The data include information on student and teacher demographics; AP course-taking and course grades; and AP exam taking and

[21] We do not reveal the names of the districts for confidentiality reasons.

scores.

To construct the sample, we begin with the high school student-course level data, which contains all courses students take in a school year after eighth grade. We restrict the course data to only AP courses as identified by course codes from the Georgia Department of Education. We do not consider AP Research or AP Seminar courses, classes that were almost never taken across our sample (e.g., AP Japanese), or AP Art courses that require a student to submit a portfolio in lieu of an exam.

Most AP courses are two semesters long and most students take both semesters, so we primarily make use of the second semester course as a marker for having taken the AP course.[22] For our main sample, we consider only students' AP courses that are most likely the terminal course in a sequence of AP courses (e.g., the first semester of a one-semester sequence or the second semester of a two-semester sequence). In particular, we include courses from students who took more than one term of a course in a year,[23] those who took a course in the final term of the year (e.g., in the second semester), and those who took a course in which most students in the broad sample take only one semester.[24] We view the resulting sample as a good proxy for terminal-course takers. We run robustness checks with an alternate definition of AP course-taking where we consider a student to have taken an AP course if she took any course in a sequence.[25]

We determine that a student took the AP exam if an AP exam record exists for the student's

[22] It can also happen that, e.g., an AP subject that is typically two semesters is only offered for one semester.
[23] We also include those enrolled in a year-long term.
[24] These courses are Physics C: Electricity and Magnetism; Government and Politics: Comparative, Economics: Microeconomics, and Economics: Macroeconomics. These courses appear to rarely be offered as a two-term sequence; in these cases, if a student took only the first term of this sequence, they would be included in the main sample. Other courses also appear to in some cases be offered as a one-semester course; if a student took one of these one-term courses in a term before the final term (e.g., in the first semester), they would not be included in the main sample.
[25] We also include students who did not earn high school course credit in the course.

corresponding course. We do not consider the small set of students who took an AP exam but did not take the corresponding course in that same year.[26] We also drop the few observations for which no student in a district in a year took the corresponding exam (minimum 10 observations).[27]

The school data include information on student race and ethnicity. Race is not mutually exclusive. Because most students select a mutually exclusive option, we categorize students as White, Black, Asian, or other race. The other race category includes those who select multiple races as well as American Indian and Pacific Islander. Hispanic/non-Hispanic is its own variable indicating if a student self-identifies as having Hispanic origin.[28] Additionally, we consider a student in a year as being FRL if they are considered eligible for either free or reduced-price lunch.

We make use of two primary samples with the above data. The first sample is at the student-year-course level (which we refer to as student-course hereafter) and includes 194,778 observations. This sample allows for multiple observations per student when they take multiple AP courses. The second sample is at the student-year level and includes 95,074 observations. It has only one observation per student per year to examine student decisions at the aggregate level. This sample includes only observations in which all of a students' courses are what we consider to be terminal courses in the year.

The Appendix contains additional details on the data preparation, such as how we clean the raw data and how we treat students who transfer districts but have conflicting demographic

---

[26] This can happen for a number of reasons such as if the student took a similar course that was not AP and wanted to try their luck on the AP exam or if the student has accumulated exam-specific knowledge in another way such as by learning a language outside of school.

[27] In particular, we drop (observations in a district-year) twice for Physics 1 and once for Physics 2 (all the same district), and twice for Government & Politics: Comparative (both for the same district). These represent approximately 1 percent of observations. In some cases, 0 students took the exam in a school in a district, potentially due to missing data; we keep these observations as long as there was at least one who took it in the district.

[28] In the summary statistics, we group and define underrepresented minorities (URM) by combining students who are at least one of Hispanic, American Indian, Black, and Pacific Islander.

information. All these decisions are about small subsets of students and will not impact our results.

2.2.3. Summary Statistics

Table 2.1 shows the summary statistics of high school students in general (column 1) and AP course takers (column 2). The unit of observation is the student, regardless of how many AP exams they took. The first two columns show that there are disproportionately lower percentages of FRL, Black, and Hispanic students taking AP courses compared to their shares in the student body. The first column of Table 2.1 shows that around half of the high school students in our four metro-Atlanta districts are FRL eligible, are female, or are Black. Column 2 shows that students who enroll in an AP course are less likely to be FRL eligible, Black, or Hispanic (33 percent are FRL eligible, 33 percent are Black, and 13 percent are Hispanic) than the entire student body (50 percent are FRL eligible, 49 percent are Black, and 18 percent are Hispanic). Most of the AP students are from grades 11 and 12 (32 and 27 percent, respectively). In addition, most of the AP students are from districts 3 and 4 (62 percent) rather than 1 and 2 (38 percent).

The remaining columns of Table 2.1 are at the student-course level such that there can be multiple observations per student. It shows all AP course enrollments (column 3), AP course enrollments that do not lead to an exam (column 4), and AP course enrollments that do lead to an exam (column 5). The third column shows that 30 percent of all AP courses are taken by FRL students, 54 percent by female students, and 28 percent by Black students.

Similarly, there are stark differences between the statistics in the fourth and fifth columns, which highlight the differences between AP course enrollees who take an AP exam and those who do not. Specifically, the fourth column shows that 34 percent of the AP courses that do not lead to an exam are taken by FRL students, 54 percent by female students, and 41 percent by Black students. The fifth column shows that 29 percent of the AP courses that lead to an exam are taken by FRL students, 55 percent by female students, and 25 percent by Black students. Twelfth graders take the highest share of total AP courses (41 percent) and the highest share of courses that do not lead to an AP exam (59 percent). Importantly, the average numeric grade in AP courses is higher for exam takers compared to non-exam takers. The average numeric grade for all AP course takers is 91.97, exam non takers is 86.11, and exam takers is 93.04. This is the first indication of positive selection into AP exam taking.

AP exam-taking behavior also varies by AP course. Table 2.2 presents the number of AP course enrollments, the percent of AP courses leading to AP exam, predicted percent of AP courses leading to AP exam (adjusted for academic performance and student demographics), average AP exam score for exam takers, and percent of females enrolled in AP courses for different AP courses. The AP courses are grouped by the broad subjects and sorted by the number of course takers within each category. From column 1 we see that some of the most popular AP courses in our sample are World History and US History and they have a high unconditional probability of leading to an AP exam (92 percent and 89 percent, respectively, compared to 85 percent in the full sample). Some of the least popular courses are foreign languages such as Chinese, Spanish and German.

From column 2, we find that the highest unconditional probability of taking an exam is for Calculus BC and Chinese (94.3 percent and 93.4 percent, respectively). Column 3 shows the probability of taking an exam conditional on academic performance and students demographic is highest for Calculus BC and Spanish Language (92.3 percent and 90.7 percent, respectively). From column 4, we see that some of the highest average AP exam scores are in Spanish Language and Calculus BC (3.8 and 3.7, respectively), which can explain their higher exam-taking rates. Column

5 shows that some of the highest percentages of female student enrollments are in French and Spanish Literature (68.0 percent and 67.3 percent female, respectively) while some of the lowest female participation is in Computer Science A and Physics C - Electricity & Magnetism (22 percent and 22.3 percent female, respectively).

2.2.4. AP Exam Prices and Subsidies

The full price of each AP exam is approximately $90 but varies by year. "Low-income" students pay considerably less but it varies by school district. The College Board pays about $30 of all AP exams for low-income students and the state of Georgia pays the remaining balance of one AP exam for all low-income public school students.[29] The College Board defines "low-income" as a family with income below 185 percent of the national poverty level or qualified as an "identified student."[30] In practice, school counselors and school administrators help find and validate subsidy eligible students.

School districts also offer varying subsidies for AP exams. During the sample period, two sample districts offered unlimited free AP exams for all students.[31] In contrast, the policies of the two other districts varied by the FRL status of the student: one offered two free AP exams for FRL eligible students and one for non-FRL students and another offered unlimited free AP exams for FRL-eligible students and one free AP exam for non-FRL-eligible students.[32]

[29] Details as of 2020 found here: https://apcentral.collegeboard.org/ap-coordinators/exam-ordering-fees/exam-fees/federal-state-assistance

[30] Students at a Community Eligibility Provision participant schools do not automatically qualify but need further validation such as being an "identified student." An identified student is defined by College Board as a student in foster care, in Head Start, experiencing homelessness or migrancy, or living in households that receive SNAP/Food Stamps, TANF cash assistance, or who receives the Food Distribution on Indian Reservations benefits. See details here: https://apcentral.collegeboard.org/ap-coordinators/exam-ordering-fees/exam-fees/reductions

[31] One of these districts requires all students to pay a $10 fee.

[32] See Appendix Table A2.3 for details.

2.3. Methodology

2.3.1. Predicting AP Exam Scores

We start by predicting AP exam scores for students who do not take the exams and hence have no score. As noted, we find that many AP courses do not lead to AP exams, yet this fact alone is not necessarily concerning. If a student is unlikely to perform well on the AP exam and will likely not earn college credit, then nothing is lost and, arguably, something is gained (e.g., time) by not taking the exam. However, if they are likely to score well on the exam but do not take it, they may be "leaving college credit on the table."

To predict AP exam scores, we use the student-course level sample and regress the AP exam scores of AP exam takers on a number of predictor variables, as seen below:

$$Score_{iscgt} = \alpha_0 + X_{it}\alpha_1 + \tau_t + \lambda_g + \eta_{sc} + \varepsilon_{iscgt} \quad (1)$$

where individual i in school s took AP course c in grade level g in year t. Each student may appear in multiple observations if they take multiple AP courses. Score is the integer score obtained on the AP exam. We include a vector of observed covariates X which includes indicators for gender, race/ethnicity, FRL status, and course grade. We also include a set of fixed effects, including school-by-course fixed effects (η_{sc}) – because there are different propensities to take the exam across different courses (see Table 2.2) and across schools—as well as year (τ_t) and grade level (λ_g) fixed effects.

We obtain coefficient estimates from equation (1) using Ordinary Least Squares and apply them to all observations to obtain predicted AP exam scores. These predictions are not integers because of our linear probability model but the continuous score can be thought of as a weighted probability of two integer scores. We also test the robustness of the result by estimating a logit model where the outcome is whether the student scores a 3 or above. This allows us to sum the predicted probabilities and determine how many students would get a 3 or higher. We obtain similar results, so we only present the above equation and results.

Regardless of our estimation strategy, our predictions are based only on observable student characteristics. It is likely that students do not take AP exams because they believe they would perform poorly for reasons unobserved to the researchers. For example, “bad exam takers” may choose not to take the exam and this could be correlated with some of our observable characteristics. We believe that this scenario and most other scenarios are likely to bias our predicted scores of non-takers upward and thereby inflate the fraction of non-exam takers who would receive a 3 or higher. As such, our estimates should be considered an upper bound.

2.3.2. Determinants of AP Exam Taking

We estimate predictors and correlates of AP exam taking among all the AP courses by using variations of the following equation:

$$TookExam_{iscgt} = \beta_0 + X_{it}\beta_1 + \beta_2 FRL_i + \theta_c + \tau_t + \delta_{sg} + \epsilon_{iscgt} \quad (2)$$

This equation is similar to equation (1) and preserves its notation, but the outcome variable is now an indicator for whether the student took the exam corresponding to the course.[33] The sample now includes all courses, not only the courses that resulted in an exam. We estimate several variations of this equation, but in our preferred specification, X includes gender and race/ethnicity indicators, and linear course grade. In all the variations of this model we include year (τ_t) and course (θ_c)

[33] The main differences are that in equation 1, to improve predictive power, we include bins of course grade and school-by-subject fixed effects, while in equation 2, we include a linear course grade, course subject fixed effects, and up to school-grade level (ninth, tenth, etc.) fixed effects.

fixed effects and the preferred model includes school-by-grade (δ_{sg}) fixed effect as well. We vary the sets of fixed effects to include different combinations of district, grade level, district-by-grade level, and school fixed effects to test the stability of our estimates on the other coefficients. We cluster the standard errors at the school level. We are particularly interested in β_2, which represents the percentage point difference in the probability of taking an AP exam, conditional on enrolling in the course, between FRL students and non-FRL students.

In addition to the student-course level analyses above, we also define the unit of observation to be at the student-year level in the following equation:

$$NumExam_{isgt} = \gamma_0 + X_{it}\gamma_1 + \gamma_2 FRL_i + \gamma_3 NCourse_i + \tau_t + \delta_{sg} + \varepsilon_{isgt} \quad (3)$$

Here, *NumExam* is a count variable denoting how many AP exams a student took. *NCourse* is a count variable for the number of AP courses the student took. We use this model to analyze how FRL status, race, ethnicity, gender, and the number of courses relate to the number of AP exams taken, conditional on the number of AP courses. The other variables are as previously defined; course grade is now the average grade of each of the courses. The coefficients of primary interest are γ_2 and γ_3, which show the difference in the number of exams taken by FRL students and the conversion rate to exam for each additional course, respectively.

In one specification that relies on equation (3), we analyze AP exam subsidy policies by making use of the number of exams a student can take for free. We determine this based on the student's FRL status and district. In districts 1 and 2, the number of free exams is equal to the number of courses. For one district among districts 3 and 4, the number of free exams is one for non-FRL students and two for FRL students (who take at least two courses). For the other district, it is one for non-FRL students, and is equal to the number of courses for FRL students.

We also consider an alternative dependent variable in equation (3), an indicator of whether the student took the same number of exams as courses (*AllExam*). The intent of this analysis is to assess whether taking relatively more courses reduces the probability of taking all exams. Students may be short on time or money and have to choose from their set of AP courses. In fact, Pope and Fillmore (2015) find that the timing and order of AP exams impacts performance.

2.4. Results

2.4.1. Basic Statistics

Overall, 15 percent of the AP courses do not lead to an AP exam (Figure 2.1). We also find substantial disparities between traditionally disadvantaged populations and relatively-advantaged students. For example, 18 percent of the courses taken by FRL students do not lead to an exam but the statistic is 15 percent for non-FRL students. For Black students, 23 percent of courses do not lead to exams while for White and Asian students it is 13 percent and 10 percent, respectively. Eighteen percent of the courses taken by Hispanic students do not lead to an exam compared to 15 percent for non-Hispanic students.

We also see differential exam-taking rates by grade and district. Twenty-three percent of twelfth graders' AP courses do not lead to an AP exam, while lower grades are close to 10 percent. Students from districts 3 and 4 are more likely to be enrolled in an AP course that do not lead to exams (18 percent) compared to those from districts 1 and 2 (10 percent).

Table 2.3 explores the relationship between the number of AP courses taken by a student in a year and the number of AP exams taken. More than half of students take only one course in a year (column 5), and few take more than five. The second column shows that the percentage of students who take zero AP exams decreases monotonically (but not linearly) from 16.6 percent to

2 percent as the number of AP courses increases from one to five. As the number of AP courses increase from one to five, the percentage of students taking all the AP exams in the courses enrolled in decreases from 83.4 percent to 66.2 percent (third column): the more AP courses, the lower the probability of taking all the corresponding AP exams. Also, the mean number of exams taken increases monotonically as the number of AP courses increases (fourth column).

Combined, Table 2.3 shows the intuitive result that students who take more AP courses take more AP exams. However, it also shows that students taking more AP courses are less likely to take all exams. This is despite the fact that we may expect the students taking more courses to be positively selected. These unconditional statistics motivate later analyses on how the number of courses relates to the number of exams, controlling for student characteristics.

2.4.2. Predicting AP Exam Scores

On average, we find that students predicted to earn higher AP grades do earn higher grades in reality, albeit with a fair amount of dispersion. Figure 2.2 presents box and whisker plots of predicted AP exam scores for each actual AP exam score, with the sample limited to courses that resulted in AP exams. The box represents the interquartile range, or the 25th through 75th percentiles, with the median in the middle. The whiskers are the top (bottom) percentile plus (minus) 1.5 times the interquartile range. We obtain a relatively high correlation coefficient between predicted and actual score of 0.78 and an R-squared of 0.61 from the prediction equation. The prediction performs relatively poorly in the tails but, fortunately, students predicted to score at the tails are less likely to be misclassified as scoring above/below a 3, which is the point of the exercise.

AP exam non-takers are predicted to perform less well than AP exam takers. Figure 2.3 plots the kernel density of predicted AP exam scores for AP exam takers and non-takers. AP exam takers have a higher predicted score (solid line) than AP exam non takers (dashed line). This is an expected result since it may not make sense to take an exam if likely to not perform well. However, this analysis is based entirely on observable characteristics, so it is possible that the true densities are further apart, depending on the role of unobservables in the decision to not take an exam.

Table 2.4 catalogues the results of the prediction exercise and leads us to the result that up to 32 percent of non-exam takers could score a 3 or higher. We group the continuous predicted scores into the discrete bands of length 0.5,indicated in the first column. The second column is the count of AP courses that did not lead to exams in a given band, and the third columns shows the corresponding percent. Summing the third column for predicted AP grades of 2.5 and above, we find that 32 percent of the courses that do not lead to an AP exam would receive a 3 or higher, a score than generally corresponds to college credit. This totals 9,495 AP courses in the four school districts in three years. The more conservative approach, which is to include only students predicted to receive a 3 or higher without rounding, yields a substantially smaller estimate of 16 percent.[34] The fourth column is the total number of students with predicted (not actual) AP exam scores, regardless of whether they took the exam or not. The fifth column is the ratio of the second and fourth columns and represents the share of courses that had no exam among all courses in each predicted grade band. The decreasing nature of this ratio is consistent with students choosing not to take the exam based on their probability of not scoring well. Yet, there are still a substantial number of students at high scores who do not take the exams and potentially leave college credit on the table.

[34] The logit model described in section 2.3.1 corresponds to the conservative model.

2.4.3. Determinants of AP Exam-taking

Table 2.5 explores the determinants of AP exam taking to investigate whether disparities by subgroup exist (after controlling for other factors) and whether any potential policies to increase AP exam taking and alleviate disparities present themselves. We estimate equation (2) with only course and year fixed effects and find that FRL-eligible students are 3.6 percentage points less likely to take an AP exam after taking the AP course. This negative relationship is common between measures of income and educational outcomes but in this context, it was not a foregone conclusion given that the FRL students were already enrolled in the course.

FRL status is likely correlated with race/ethnicity and other variables, so we add in a set of controls to test the stability of our initial FRL coefficient. Column 2 adds controls for sex and race/ethnicity in addition to the course and year fixed effects in column 1. The coefficient on FRL status is now statistically indistinguishable from zero but the coefficients on race and ethnicity highlight some disparities. Black and Hispanic students are 11 and 4.3 percentage points less likely to take an AP exam than White and non-Hispanic students, respectively.

We next provide evidence that students with lower AP course numeric grades are less likely to take the AP exam. Figure 2.4 shows the positive relationship between numeric AP course grade and the probability of taking the AP exam. Not only does this reinforce the previous analyses that some students with a low probability of performing well (as measured by course grades) are not taking the exam but it also suggests that AP course grade is an important determinant that could be confounded with our demographic variables; thus, it is an important control variable. Another notable insight from Figure 2.4 is that there are more circles above the diamonds when the numeric score is lower than 70. This shows that FRL eligible students are more likely to take the AP exam even for lower AP course grades.

Adding course grade to the regressions, as in column 3 of Table 2.5, confirms the positive relationship between course grades and AP exam taking, even conditional on other student characteristics. Interestingly, the coefficient on FRL status is now a positive 2.1 percentage points, showing that FRL students with similar course performances as non-FRL students are more likely to take the AP exam. This motivates analyses below as to whether the subsidies that target FRL students are playing a role. The coefficients on race and ethnicity are somewhat muted relative to the previous column but the disparities remain.

We add various combinations of district, school, and grade-level fixed effects in columns 4 through 7 of Table 2.5, and the coefficients are relatively insensitive to our choice of fixed effects.[35] Column 7 is our preferred specification because the school-by-grade fixed effects mean we are comparing students with different characteristics (e.g., race or FRL status) but who are in the same high school and grade, which could conceivably explain the coefficients (but does not in practice). We find that FRL eligible students have a 2 percentage point higher probability of taking the AP exam than non-FRL students. Given the baseline prevalence of taking an AP exam is 85 percent, this amounts to a 2.35 percent increase in the probability of taking the AP exam. This is not an especially large coefficient or implied percent; however, the fact that FRL students tend to lag behind in most educational outcomes makes this coefficient stand out relative to the existing literature. Also, this positive relationship between FRL status and exam taking is consistent with higher AP exam subsidies for FRL students than for non-FRL students in two of our four partner school districts. We explore this further in section 2.4.4.

[35] We obtain nearly identical results for the model with school fixed effects when we instead use school random effects.

In column 7, we see small differences in AP exam-taking rates by gender (females are 0.8 percentage points lower than males), but Black students are 3.8 percentage points less likely to take an AP exam than their White peers (in the same school and grade). While we cannot explain this result, it is worthy of future investigation. Finally, other race students have a 1.8 percentage point lower probability of taking an exam compared to their White peers, and Hispanic students have a 1.3 percentage point lower probability of taking an AP exam compared to their non-Hispanic peers.[36]

2.4.4. AP Subsidies

To explore if the positive relationship between FRL eligibility and AP exam taking (found in column 7 of Table 2.5) might be driven by the two school districts that provide higher AP exam subsidies for FRL students than for non-FRL students, we perform subsample analysis in Table 2.6. Panel A includes only districts 1 and 2—those that offer all AP exams for free to all students, regardless of their FRL status. In column 1, we obtain a large, negative, and significant coefficient on the FRL variable. However, as we add controls, especially fixed effects, we find no statistical relationship with FRL status and AP exam taking in these districts. This null effect is largely a desirable result for districts because FRL students tend to lag behind in educational outcomes.

Panel B of Table 2.6 only includes districts 3 and 4, the districts in which FRL students get more AP exam subsidies than non-FRL students. With a full set of controls, we show that being FRL eligible is associated with a 2.9 percentage point increase in the probability of taking the AP exam compared to non-FRL students. Taken together, Panels A and B of Table 2.6 show that the

[36] We do a robustness check for these results in table with an alternate definition of AP course taking. In Appendix Table A2.3, we define a student to have taken an AP course if she took any course in that AP course sequence in that year. The results are qualitatively the same.

results in Table 2.5 are driven by districts 3 and 4—the very districts that provide extra financial incentives for FRL students to take the AP exams.

We next look deeper into the impact of AP exam subsidies by analyzing the data at the student level and looking across districts and FRL status. Table 2.7 shows the results of equation 3 where the outcome is the number of AP exams taken. Similar to Table 2.5, we see that the coefficient on FRL is positive and statistically significant, implying that conditional on a host of variables, particularly numeric grade, FRL students take 0.025 more AP exams than non-FRL students. The coefficient is relatively small in magnitude but surprisingly non-negative.

We then explore the relationships of various demographic and academic variables with the number of AP exams taken by a student for different subsamples. To that end, we restrict the sample to various district and FRL status combinations in columns 2 through 7. We first consider districts 1 and 2, where the amount of the AP exam subsidy does not depend on FRL status and all students get all exams for free. In column 2, we cannot detect a relationship between FRL status and the number of AP exams taken. Column 5 restricts to districts 3 and 4, where the AP exam subsidy varies by FRL status. In contrast to the result for districts 1 and 2, we find that FRL students take 0.040 more AP exams compared to non-FRL students, which is consistent with the exam subsidy influencing exam taking. Further evidence in support of the positive role of subsidies is shown by comparing the coefficients on the number of courses in columns 2 and 5. We find that taking one more AP course results in 0.939 more AP exams in the districts that provide all exams for free (districts 1 and 2), while in the districts that do not, taking one more AP course results in 0.880 more AP exams.

Next, we next split column 2 (districts 1 and 2) into FRL students (column 3) and non-FRL students (column 4) to further examine the role of subsidies by FRL status. Non-FRL students have a higher conversion rate of course to exam (one more course leads to 0.945 more AP exams) compared to FRL students (one more course leads to 0.899 more AP exams). Absent higher subsidies compared to non-FRL students, FRL students in districts 1 and 2 are less likely to take an AP exam, which is consistent with most literature and the negative relationship between FRL status and educational outcomes. However, we find the opposite in districts 3 and 4 in which FRL students take more exams than non-FRL students. Specifically, taking one more course leads to 0.920 more exams for FRL students, while this number is 0.854 for non-FRL students. These results are consistent with a "subsidy effect" that drives the course to exam conversion rate up, even more so than any "FRL effect," which tends to drive the course to exam conversion rate down.

Column 8 adds the number of subsidized exams to the previous analysis to estimate the relationship between the number of subsidized AP exams and the number of AP exams taken. This number depends on the district a student is in, her/his FRL status, and the number of courses to which s/he is enrolled. We find that after controlling for FRL status and the number of AP courses enrolled in, one more subsidized exam leads to 0.075 more AP exams. Again, these findings are consistent with subsidies positively affecting AP exam taking.[37]

We also test for non-linearities in the number of AP courses but results, shown in Appendix Table A2.5, are qualitatively similar to our linear specification.

2.4.5. Gender and Role Models

[37] We do a robustness check for the results in Appendix Table A2.4 with an alternate definition of AP course taking. We define a student to have taken an AP course if she took any course in that AP course sequence in that year. The results are qualitatively the same.

Next, we follow the literature on gender matching of student and instructors serving to improve educational outcomes. Specifically, we explore whether female students take fewer AP exams than male students in AP courses that are underrepresented by females and whether female instructors have the potential to influence female students' exam-taking rates. We start by splitting the AP subjects into those that are underrepresented and overrepresented by females across our four school districts: subjects where less than 40 percent of the students are female, subjects where 40-60 percent of the students are female, and subjects where more than 60 percent of the students are female.[38] Females are underrepresented in subjects like computer science and physics but overrepresented in subjects like psychology and English.

We have two predictions about female student exam-taking rates based on previous literature regarding gender differences in education and subject and career tracts (e.g., Buser et al., 2014). First, female students are less likely to take AP exams in subjects where they are underrepresented. Second, female students with female instructors are more likely to take an AP exam than their male counterparts in subjects where females are underrepresented.

Neither of our predictions are confirmed by the analysis. Table 2.8 shows the regression results where the unit of observation is again a student-course (equation 2) and the outcome is whether an exam is taken, but now we include teacher gender and the interaction with student gender.[39] In the full sample (column 1), we estimate a small and negative relationship between student gender and exam taking (-1.1 percentage point). However, the estimates between subjects that are underrepresented by females (column 2) and those well represented by females (column 3) are not meaningfully different from one another. We also find no statistical relationship of the

[38] See Table 2.2 for percentage female by subject.
[39] The few observations with missing teacher gender are excluded from these analyses.

interaction of female teacher with female students on exam taking, even in underrepresented subjects (column 2). Overall, we find no evidence that matching female students with female instructors has any influence on the probability of taking an AP exam.

We note that these estimates are not based on random assignment of students and teachers, so our estimates are likely biased. The coefficient on female is likely biased upward if female students in underrepresented subjects are substantially different than the typical female student (on observables and unobservables). Similarly, if female students sort into classrooms with female instructors, they may be positively selected and bias our estimates upward. But in this case, an upward bias has some meaning because we found no statistical relationship between AP exam taking and females being paired with female teachers. An upward bias implies that the unbiased estimate is bounded by zero, which is a rejection of the "role model effect" of female teachers on female students on AP exam taking in underrepresented subjects.

2.4.6. Twelfth Graders

High school twelfth graders face different incentives to take an AP exam than students in lower grades. First, twelfth graders may know with more certainty if and where they will go to college and the college credit offered (or not) for an AP exam score. Second, time in high school is ending, so students only have one more opportunity to earn college credit while in high school. Our data shows that twelfth graders are both the most common AP course and exam takers but they also have the highest prevalence of not taking an exam after taking the course (23 percent compared to 10, 8, and 12 percent for grades 9, 10 and 11, respectively).

To further explore the relationship between high school grade of the student on AP exam taking, we split the sample into twelfth graders and those in other grades and re-estimate equation

3 (analogous to Table 2.7). In panel A of Table 2.9 we see how FRL eligibility, number of courses enrolled in, and number of subsidized tests relate to the number of AP exams for non-twelfth graders. In panel B we see the same relationships for twelfth graders.

Comparing the coefficients for FRL indicator in the two panels of Table 2.9 reveals that the earlier results of Table 2.7 were driven by twelfth graders. That is, being an FRL student in twelfth grade is associated with taking more AP exams and this is driven by districts 3 and 4 where exam subsidies are higher for FRL students than for non-FRL students. If students are not to receive credit at the college they plan to attend, then there is little incentive to take the AP exams, even with a subsidy.

Next, evaluating the coefficients on the number of courses by panel reveals that in districts 1 and 2, each additional course enrollment leads to a higher number of additional exams compared to districts 3 and 4 (compare column 2 and 5). However, the difference is much larger for twelfth graders compared to non-twelfth graders. So, the higher subsidy generosity is related to more exam taking for twelfth graders more so than to non-twelfth graders. Comparing columns 3 and 4 in the two panels we see that in districts that have the same subsidy policy for all students (districts 1 and 2), the non-FRL students have a higher "conversion rate" from course to exam compared to the FRL students for both twelfth graders and non-twelfth graders. However, this relationship is more pronounced for twelfth graders (0.854 for FRL students versus 0.920 for non-FRL students) compared to non-twelfth graders (0.961 for FRL students versus 0.976 for non-FRL students). Also, comparing columns 6 and 7 in the two panels we see that for the districts where the AP exam subsidy generosity depends on FRL status (districts 3 and 4), FRL eligible students have a higher conversion rate from course to exam for each additional course compared to non FRL students. Again, this relationship is more pronounced for twelfth graders (0.866 for FRL students versus

0.783 for non-FRL students) compared to non-twelfth graders (0.964 for FRL students versus 0.916 for non-FRL students). Taken together, this means that while AP subsidy is positively associated with AP exam taking for the full sample, the association is primarily driven by twelfth graders. It also leaves open the question as to why exam subsidies are seemingly less effective in earlier graders.

2.4.7. Number of Courses

In this section, we further investigate how taking different numbers of courses relates to the probability of taking all the AP exams corresponding to those courses. As discussed previously, the number of courses in which a student enrolls is related to the number of AP exams. We also learned that the probability of taking all exams monotonically decreases with the number of courses in which a student enrolls (Table 2.3). This analysis uses a variation of equation 3 but the dependent variable is a binary indicator for whether the student took all the AP exams corresponding to the set of courses in which the student enrolled.

The first column of Table 2.10 shows that FRL students have a 1.3 percentage point higher probability of taking all the exams among her/his courses compared to non-FRL students. From columns 2 and 3, we see that the full sample results are driven by districts 3 and 4 where the AP subsidy generosity is based on FRL status. This is consistent with what we see in Table 2.7 and shows that an AP subsidy is associated with a higher probability of taking all the exams. Furthermore, from column 1 we see that being Black or other race is associated with 3.4 and 2.3 percentage point lower probability of taking all the AP exams compared to being White. From columns 2 and 3 we see that these race specific results are also driven by districts 3 and 4, the districts that have higher AP subsidies for FRL students compared to non-FRL students.

We also see that the probability of taking all the exams monotonically decreases with the number of courses even conditional on a rich set of controls. From columns 2 and 3, we see a similar pattern for both groups of districts. However, a negative and statistically significant association of taking all the exams with additional courses show up from the fourth course onward for districts 1 and 2, whereas a more pronounced negative significant association shows up from second course onwards in districts 3 and 4.[40]

These results speak to district exam subsidy policy and AP course enrollment policy. Students will not take all AP exams as they take more courses. The results also show that some districts (1 and 2) do not see this until students take relatively many AP courses. But this last fact is only true for the districts that provide all AP exams are free for all students. On the other hand, the districts that do not do so (districts 3 and 4) show a much higher level of non-exam taking at higher number of AP course loads.

2.5. Discussion and Conclusion

We find that the practice of AP course-taking without exam taking is fairly prevalent in the four metro-Atlanta school districts that comprise our sample, with 15 percent of the AP courses not leading to an exam. We estimate that up to 32 percent of the courses that do not lead to an AP exam would receive a score of 3 or higher if the exam was taken. Thus, high school students seem to be leaving credit on the table by not taking the AP exams.

In our sample of four school districts over three years, this amounts to an upper bound of 9,495 AP courses that could have turned to college credits if the AP exams were taken. In SY

[40] In Table 2.9, we split the sample by ninth-eleventh grade and twelfth grade. We find that the FRL results are driven by twelfth graders. We also find roughly similar results for additional courses, with larger magnitudes for twelfth graders.

2017-18, the tuition and fees faced by a four-year in-state Georgia public college student was $7,206.[41] Assuming a typical college student enrolls for 30 credits per year (10 three credit courses), one three credit college course costs $721. 9,495 successful AP exams would save students in the four districts up to $6.8 million over three years.[42] This is roughly eight times the cost of these AP exams ($0.9 million)[43] for the students or the state and school districts,[44] though it is a smaller multiple if unsuccessful exams are factored in.

AP course offerings and exam subsidies are well within the control of school districts, though budgets constrain the ability to pay for these courses and subsidies. In exploring the possible determinants of AP exam taking, we show strong evidence that districts' policies on exam subsidies seem to improve exam-taking rates. The existing policies make it no less likely that FRL students take exams than non-FRL students in the same high school with the same course grades. We also show some evidence that the timing and number of AP courses relates to the probability of taking exams, which can inform which students are at risk of not taking the exam.

We also show no statistical relationship between females having female versus male instructors in classes where females are underrepresented. This does not leave us with a direct policy lever relating to gender, but it does add to the growing literature on "instructors like me." Unlike most research in the area, our null effect highlights that not all contexts yield positive results.

Our results quantify the potential issue of not taking AP exams and highlight that this may be a relatively straightforward policy lever for schools and districts, especially because students

---

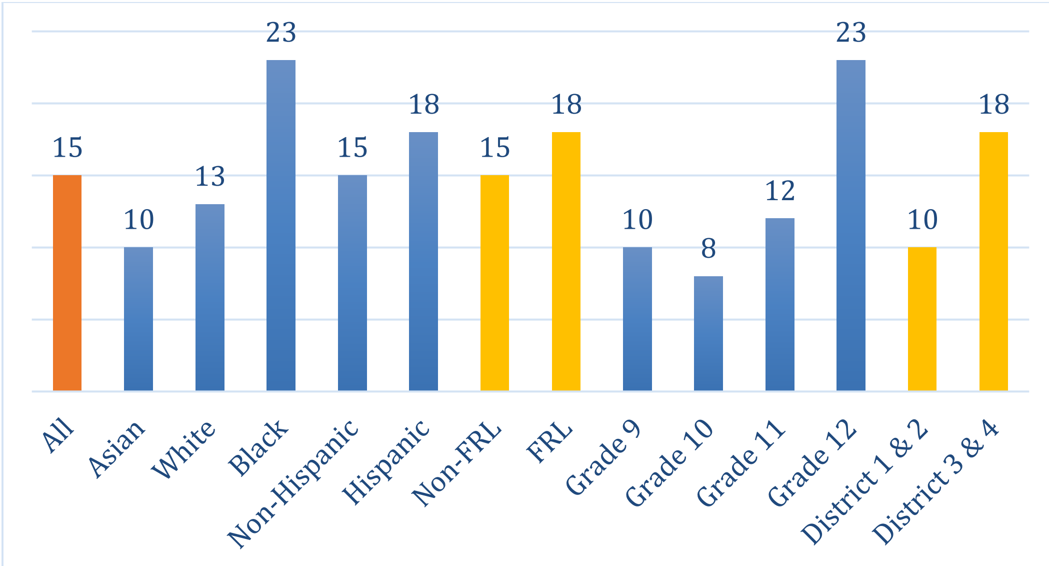
[41] Source: National Center for Education Statistics, https://nces.ed.gov/programs/digest/d18/tables/dt18_330.20.asp. Table 330.20.
[42] $721*9,495 = $6,845,895.
[43] $91*9,495 = $864,045. This assumes the previous cost of $91, which is now $94.
[44] The calculation depends on whether the student, district, state, or College Board pays the fees.

are already taking the AP course. However, we cannot and do not quantify the benefits and costs of AP course enrollment, which is front and center of many administrators' minds. The benefits (and costs) of the course enrollment may far outweigh the benefits (and costs) of taking the exam.

Our work also only examines policies and practices that are observable in the data. Schools and districts likely push students to take AP exams in ways unobserved to the researchers. Policies, initiatives, and intervention around AP exam taking are a fruitful area for future research.

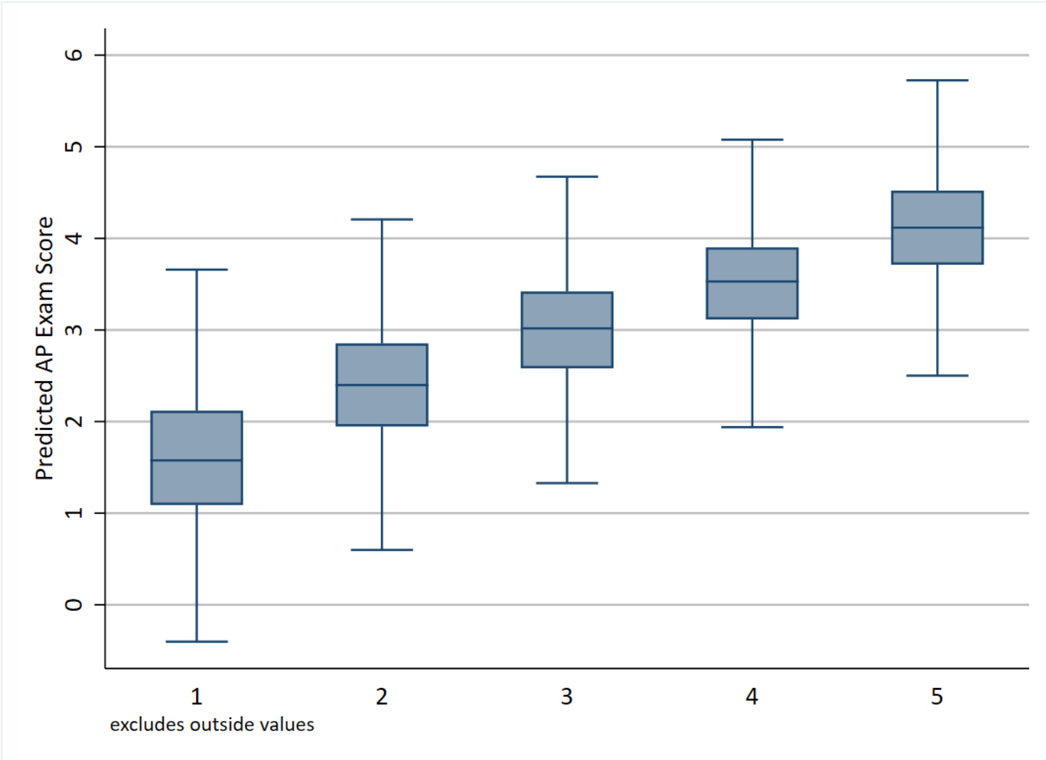
2.6. Figures and Tables

Figure 2.1: Percent of AP Courses Without a Corresponding AP Exam



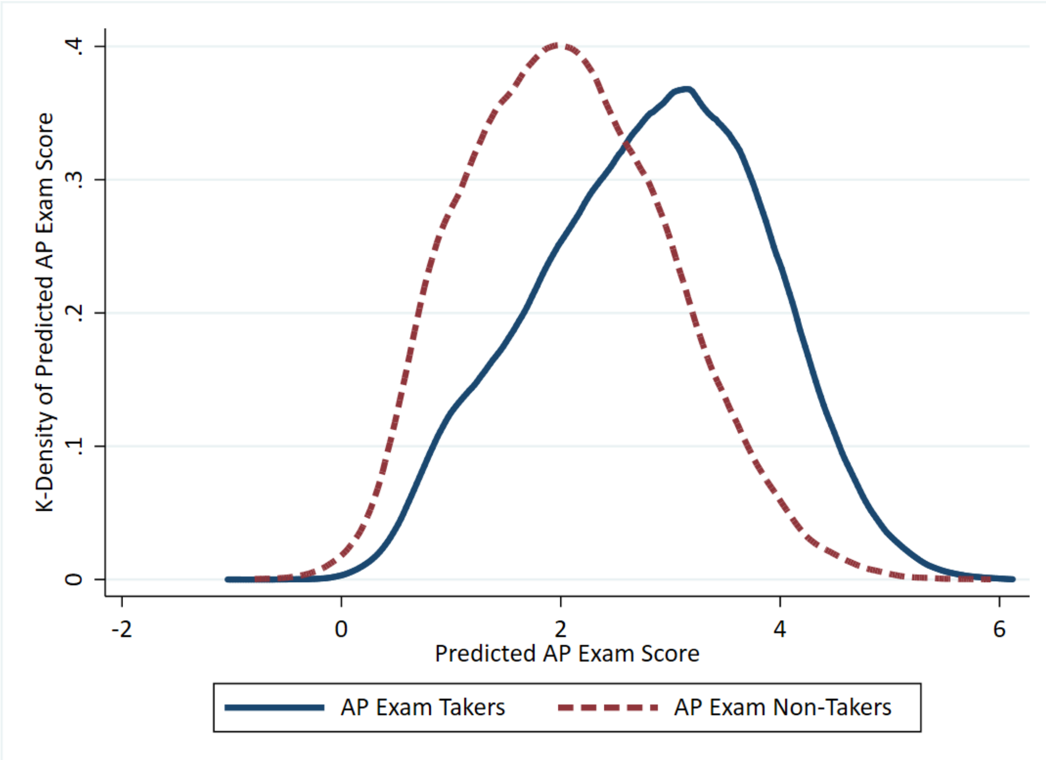
Note: This figure shows the percent of AP courses that do not lead to an AP exam in the full sample and in subsamples by race, ethnicity and FRL status.

Figure 2.2: Actual and Predicted AP Exam Grade of AP Exam-Takers



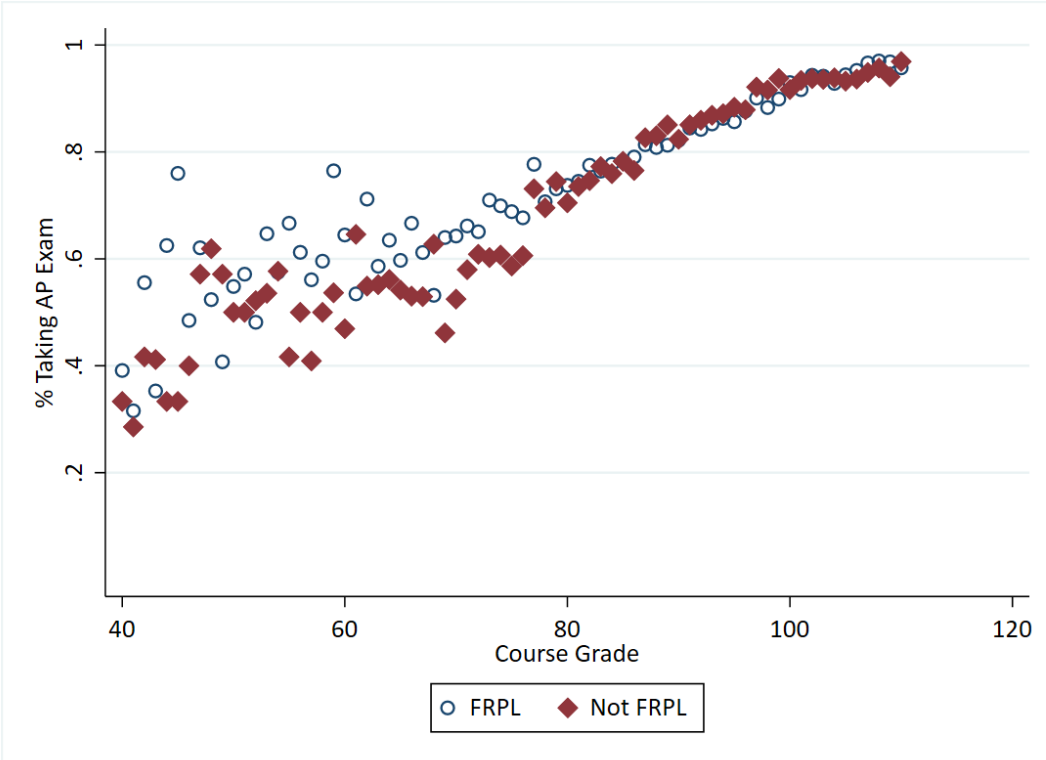
Notes: This figure shows a box and whisker plot of predicted AP exam score vs. actual AP exam score for AP exam-takers (where the unit observation is a student-year-course). The middle of the box is the median, and the edges are the interquartile range, or the 25th and 75th percentiles. The top line, or whisker, is the adjacent value, which is the 75th percentile plus the interquartile range multiplied by 1.5. The bottom line, or whisker, is the 25th percentile minus the interquartile range multiplied by 1.5 Observations beyond the whiskers, or outside values, are not displayed.

Figure 2.3: Predicted AP Exam Grade Distribution for Exam Takers and Non-Takers



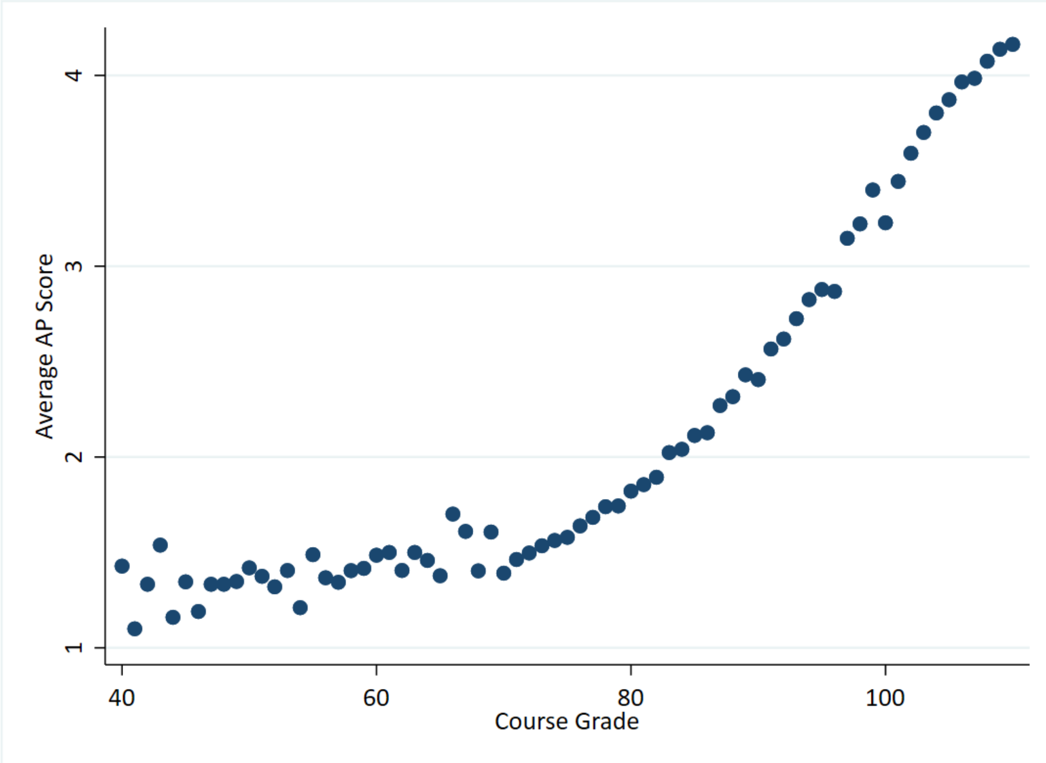
Notes: This figure shows kernel densities of predicted AP exam grade for AP exam-takers (where the unit observation is a student-year-course) in blue and for non-AP Exam-takers in red.

Figure 2.4: Relationship Between Course Grade and Exam-Taking



Notes: This figure shows the relation between course grade (x-axis) and the percentage of student with a given course grade of at least 40 who take the corresponding AP exam (y-axis). Calculations are performed at the student-course-year level.

Figure 2.5: Relationship between Exam Score and Course Grade



Notes: This figure shows the relation between course grade (x-axis) and the average AP exam score (y-axis) for the corresponding AP exam for students with a given course grade of at least 40. Calculations are performed at the student-course-year level, and are conditional on the student taking the exam.

Table 2.1: Summary Statistics

	Student-level		Course-level		
	Full Sample	AP Course-taking Students	All	No Exam	Exam
FRL	0.50	0.33	0.30	0.34	0.29
Female	0.49	0.56	0.54	0.54	0.55
Black	0.49	0.33	0.28	0.41	0.25
Asian	0.09	0.17	0.22	0.14	0.23
White	0.31	0.41	0.42	0.35	0.44
Other	0.11	0.09	0.09	0.10	0.08
Hispanic	0.18	0.13	0.13	0.15	0.12
Grade 9	0.29	0.16	0.09	0.05	0.09
Grade 10	0.26	0.25	0.17	0.09	0.18
Grade 11	0.22	0.32	0.34	0.26	0.36
Grade 12	0.23	0.27	0.41	0.59	0.37
District 1 and 2	0.35	0.38	0.35	0.23	0.37
District 3 and 4	0.65	0.62	0.65	0.77	0.63
Numeric Grade		90.54	91.97	86.11	93.04
Observations	387,698	95,074	194,778	30,125	164,653

Notes: This table shows means for several populations: 1. Full Sample, which includes all students—regardless of if they took an AP exam—in grades 9-12; this column is at the student-year level. 2. The AP course-taking sample, which is at the student-year level. 3. The AP course-taking sample, which is at the student-year-course level. The final two columns split column 3 by whether or not the course led to an exam.

Table 2.2: AP Course Subjects

Course	N	% Took Exam	Predicted % Took Exam	Exam Grade	Female
History and Social Sciences					
World History	20,209	92.0	86.2	2.7	56.2
US History	19,594	89.4	88.3	2.6	55.5
Economics: Macroeconomics	14,528	73.9	78.4	2.8	50.6
Psychology	12,438	85.9	87.7	3.1	63.3
Human Geography	12,184	90.6	86.5	2.7	55.3
Government & Politics: US	9,019	83.0	81.8	2.6	53.1
Economics: Microeconomics	4,343	76.0	80	3.2	44.0
European History	1,105	82.7	81.6	3.3	47.6
Government & Politics: Comparative	1,046	82.5	81.8	3.4	49.4
English					
English Language & Composition	17,635	91.1	90.1	2.8	62.6
English Literature & Composition	12,390	75.5	81.5	2.6	63.6
Sciences					
Environmental Science	10,572	75.5	77.2	2.6	54.3
Biology	8,572	88.2	87.5	2.8	58.2
Physics 1	7,597	80.9	79.1	2.3	45.7
Chemistry	4,828	87.7	85.9	2.8	51.6
Physics C – Mechanics	1,097	83.9	85.2	3.9	30.9
Physics C - Electricity & Magnetism	993	86.2	86.7	3.6	22.3
Physics 2	633	56.4	55.7	2.6	39.0
Math and Computer Science					
Calculus AB	9,771	85.0	87.6	2.8	52.4
Statistics	9,717	79.5	82.5	2.7	52.5
Calculus BC	4,513	94.3	92.3	3.7	44.3
Computer Science A	3,624	71.1	69.6	2.8	22.0
Computer Science Principles	1,053	78.1	76.3	3.2	24.2

World Languages and Cultures					
Spanish Language	3,365	88.9	90.7	3.8	63.2
French	1,067	85.4	88.5	3.1	68.0
Latin	429	88.1	87.9	2.6	49.0
German	338	88.5	87.5	3.6	50.9
Spanish Literature	278	83.8	86.5	2.5	67.3
Chinese	151	93.4	90.3	4.0	44.4
Arts					
History of Art	923	92.4	89.7	3.0	62.1
Music Theory	766	71.7	73.1	3.1	47.1

Notes: This table shows summary statistics broken by AP subject where the unit of observation is a student-year-course. The first column shows the number of courses, followed by the percentage of these courses that resulted in an AP exam. The third column adjusts for FRL, female, race, Hispanic, and grade, and is the predicted exam-taking percentage for the average student along each of these characteristics. The fourth column is the average AP exam grade received conditional on taking the exam. The final column is the percentage female.

Table 2.3: AP Exam-taking Behavior

#AP Courses	% Take 0 Exams	% Take All Exams	Mean # Exams	N
1	16.6	83.4	0.8	53,679
2	8.1	79.2	1.7	22,194
3	4.4	75.5	2.6	10,939
4	3.3	71.2	3.5	5,507
5	2.0	66.2	4.4	2,130
6-9	1.4	57.3	5.3	625

Notes: This student-year level table shows AP exam-taking behavior for students enrolled in different number of courses (denoted in column 1). Students taking 6 through 9 courses are combined. The second column shows the percentage of students in the row who take 0 exams, while the third column shows the percentage of students in the row who take all of their exams. The fourth column shows the mean number of exams taken.

Table 2.4: Predicted AP Exam Scores

Predicted AP grade	Courses that did not lead to AP exam in this grade band	Courses that did not lead to AP exam in this grade band as percent of total number of courses that did not lead to an AP exam	# Total courses in this grade band irrespective of the AP exam being taken or not	Courses that did not lead to AP exam as percent of all AP courses in this grade band
<1	4122	13.7	12251	33.6
[1, 1.5)	4883	16.2	17373	28.1
[1.5, 2)	5842	19.4	23586	24.8
[2, 2.5)	5783	19.2	29212	19.8
[2.5, 3)	4553	15.1	32743	13.9
[3, 3.5)	2831	9.4	32296	8.8
[3.5, 4)	1421	4.7	25759	5.5
[4, 4.5)	511	1.7	14427	3.5
>=4.5	179	0.6	7131	2.5
Total	30125	100.0	194778	

Notes: This is a student-course-year level table where the counts are number of individual courses. Actual AP scores are in integer values from 1 to 5. Our predicted AP scores are continuous and unbounded. In this table we present the predicted AP scores in buckets of 0.5 increments.

Table 2.5: Determinants of AP Exam Taking

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FRL	-0.036**	0.004	0.021*	0.031***	0.031***	0.019***	0.020***
	(0.015)	(0.012)	(0.012)	(0.010)	(0.009)	(0.005)	(0.005)
Female		0.005*	-0.008***	-0.007***	-0.007***	-0.008***	-0.008***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Black		-0.110***	-0.067***	-0.038***	-0.036***	-0.040***	-0.038***
		(0.022)	(0.021)	(0.011)	(0.010)	(0.006)	(0.006)
Asian		0.020**	0.005	0.013*	0.013*	0.003	0.003
		(0.010)	(0.010)	(0.007)	(0.007)	(0.004)	(0.004)
Other		-0.042***	-0.027**	-0.004	-0.004	-0.019***	-0.018***
		(0.013)	(0.014)	(0.011)	(0.011)	(0.006)	(0.006)
Hispanic		-0.043***	-0.023**	-0.004	-0.003	-0.015**	-0.013**
		(0.010)	(0.010)	(0.010)	(0.009)	(0.006)	(0.005)
Grade Level = 10			0.020	0.047		0.034	
			(0.037)	(0.035)		(0.029)	
Grade Level = 11			-0.017	-0.008		-0.017	
			(0.035)	(0.033)		(0.027)	
Grade Level = 12			-0.083**	-0.068**		-0.077***	
			(0.036)	(0.034)		(0.028)	
Numeric Grade			0.008***	0.008***	0.008***	0.009***	0.009***
			(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.909***	0.927***	0.198***	0.278***	0.209***	0.195***	0.191***
	(0.017)	(0.016)	(0.072)	(0.062)	(0.068)	(0.056)	(0.056)
Observations	194,778	194,778	194,778	194,778	194,778	194,778	194,778
R-squared	0.038	0.057	0.111	0.137	0.146	0.196	0.223
Year FE	YES	YES	YES	YES	YES	YES	YES
Course FE	YES	YES	YES	YES	YES	YES	YES
District FE				YES			
District-Grade FE					YES		
School FE						YES	

School-Grade FE							YES
Mean of Dependent Variable	0.845	0.845	0.845	0.845	0.845	0.845	0.845

Notes: Observations are at the student-course-year level. The dependent variable is a binary variable for taking the AP exam. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.6: Determinants of AP Exam Taking

Panel A: Districts 1 and 2 (Same Subsidy for All Students)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FRL	-0.064***	0.019	0.029	-0.000	-0.001	-0.009	-0.009
	(0.019)	(0.024)	(0.023)	(0.011)	(0.011)	(0.006)	(0.006)
Observations	68,169	68,169	68,169	68,169	68,169	68,169	68,169
R-squared	0.026	0.066	0.104	0.164	0.168	0.247	0.290
Mean of Dependent Variable	0.897	0.897	0.897	0.897	0.897	0.897	0.897

Panel B: Districts 3 and 4 (Higher Subsidy for FRL Students than non-FRL Students)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FRL	-0.003	0.019	0.040***	0.040***	0.040***	0.028***	0.029***
	(0.017)	(0.013)	(0.011)	(0.012)	(0.012)	(0.006)	(0.006)
Observations	126,609	126,609	126,609	126,609	126,609	126,609	126,609
R-squared	0.057	0.069	0.135	0.135	0.137	0.184	0.200
Year FE	YES	YES	YES	YES	YES	YES	YES
Course FE	YES	YES	YES	YES	YES	YES	YES
Demographic Controls		YES	YES	YES	YES	YES	YES
District FE				YES			
District-Grade FE					YES		
School FE						YES	
School-Grade FE							YES
Mean of Dependent Variable	0.818	0.818	0.818	0.818	0.818	0.818	0.818

Notes: Observations at the student-year-course level. The dependent variable is an indicator for taking the AP exam. Demographic controls are Female, Asian, Black, Other, and Hispanic. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.7: Number of Exams Taken

VARIABLES	(1) Full Sample	(2) Districts 1 & 2 *Same subsidy for all*	(3) Districts 1 & 2: FRL Only	(4) Districts 1 & 2: Non-FRL Only	(5) Districts 3 & 4 *Different subsidy by FRL*	(6) Districts 3 & 4: FRL Only	(7) Districts 3 & 4: non-FRL Only	(8) Full Sample with Number of Subsidized Exams
FRL	0.025***	-0.012			0.040***			-0.012
	(0.009)	(0.008)			(0.011)			(0.009)
Female	0.001	0.004	0.012	0.003	-0.001	0.013	-0.010	0.001
	(0.005)	(0.006)	(0.011)	(0.007)	(0.006)	(0.008)	(0.007)	(0.005)
Black	-0.048***	-0.000	-0.006	0.001	-0.067***	-0.034**	-0.068***	-0.051***
	(0.010)	(0.013)	(0.024)	(0.014)	(0.012)	(0.016)	(0.016)	(0.011)
Asian	0.017**	0.020	0.060**	0.013	0.008	0.053***	0.001	0.012
	(0.008)	(0.012)	(0.029)	(0.011)	(0.009)	(0.015)	(0.012)	(0.007)
Other	-0.027***	0.010	0.019	0.008	-0.038***	-0.012	-0.054***	-0.031***
	(0.008)	(0.011)	(0.027)	(0.013)	(0.010)	(0.010)	(0.015)	(0.008)
Hispanic	-0.011	0.021*	0.016	0.019*	-0.022*	-0.008	-0.012	-0.013
	(0.010)	(0.012)	(0.024)	(0.011)	(0.012)	(0.016)	(0.016)	(0.010)
Avg. Course Grade	0.013***	0.008***	0.007***	0.009***	0.015***	0.013***	0.017***	0.013***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
# Courses	0.901***	0.939***	0.899***	0.945***	0.880***	0.920***	0.854***	0.859***
	(0.014)	(0.017)	(0.030)	(0.016)	(0.018)	(0.018)	(0.022)	(0.020)
# Subsidized Tests								0.075***
								(0.018)
Observations	95,074	35,834	7,655	28,179	59,240	23,393	35,847	95,074
R-squared	0.809	0.862	0.753	0.876	0.776	0.787	0.773	0.810
Mean of Dpt. Variable	1.506	1.584	1.264	1.671	1.459	1.381	1.510	1.506

Notes: Observations are at the student-year level. The dependent variable is the number of AP exams taken by the student in a year. All columns include School-Grade and year fixed effects. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.8: Determinants of AP Exam Taking: Gender

	(1) Full Sample	(2) AP Subjects with 0-40% Female Participation	(3) AP Subjects with 40-60% Female Participation	(4) AP Subjects with 60-100% Female Participation
Female	-0.011***	-0.016	-0.011***	-0.002
	(0.003)	(0.010)	(0.003)	(0.006)
Teacher Female	-0.001	-0.046	0.014**	-0.013
	(0.006)	(0.047)	(0.007)	(0.016)
Female-Teacher Female	0.005	0.017	-0.001	0.010
	(0.004)	(0.014)	(0.004)	(0.007)
Observations	192,842	7,131	138,000	47,711
R-squared	0.226	0.381	0.241	0.217
Mean of Dependent Variable	0.847	0.752	0.849	0.856

Notes: Observations are at the student-year-course level. The dependent variable is a binary variable for taking the AP exam. Observations with missing teacher gender are dropped. All columns include School-Grade, course, and year fixed effects, and control for Asian, Black, Other, and Hispanic. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.9: Number of Exams Taken, Ninth-Eleventh Grade and Twelfth Grade Samples

Panel A: Ninth-Eleventh Grade Sample

	(1) Full Sample	(2) Districts 1 & 2	(3) Districts 1 & 2: FRL Only	(4) Districts 1 & 2: Non-FRL Only	(5) Districts 3 & 4	(6) Districts 3 & 4: FRL Only	(7) Districts 3 & 4: non-FRL Only	(8) Full Sample with # Subsidized Tests
FRL	0.003	-0.004			0.007			-0.014**
	(0.006)	(0.007)			(0.008)			(0.007)
# Courses	0.948***	0.974***	0.961***	0.976***	0.935***	0.964***	0.916***	0.925***
	(0.009)	(0.008)	(0.019)	(0.008)	(0.013)	(0.013)	(0.014)	(0.013)
#Subsidized Tests								0.043***
								(0.011)
Observations	69,038	25,018	5,084	19,934	44,020	16,998	27,022	69,038
R-squared	0.844	0.886	0.778	0.900	0.822	0.817	0.826	0.844
Mean of Dpt. Variable	1.416	1.428	1.182	1.490	1.409	1.333	1.456	1.416

Panel B: Twelfth Grade Sample

	(1) Full Sample	(2) Districts 1 & 2	(3) Districts 1 & 2: FRL Only	(4) Districts 1 & 2: Non-FRL Only	(5) Districts 3 & 4	(6) Districts 3 & 4: FRL Only	(7) Districts 3 & 4: non-FRL Only	(8) Full Sample with # Subsidized Tests
FRL	0.079***	-0.031			0.130***			0.000
	(0.022)	(0.023)			(0.027)			(0.018)
# Courses	0.852***	0.910***	0.854***	0.920***	0.815***	0.866***	0.783***	0.788***
	(0.019)	(0.026)	(0.046)	(0.023)	(0.026)	(0.024)	(0.033)	(0.030)
#Subsidized Tests								0.109***
								(0.028)
Observations	26,036	10,816	2,571	8,245	15,220	6,395	8,825	26,036
R-squared	0.765	0.826	0.722	0.836	0.724	0.752	0.709	0.767
Mean of Dpt. Variable	1.746	1.945	1.427	2.106	1.605	1.508	1.675	1.746

Notes: Observations are at the student-year level. The dependent variable is the number of AP exams taken by the student in a year. Panel A restricts the sample to students in grades 9-11, while Panel B restricts to students in grade 12. All columns include School-Grade and year fixed effects, and control for Asian, Black, Other, and Hispanic. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table 2.10: Took All Exams

	(1) Full Sample	(2) Districts 1&2	(3) Districts 3&4
FRL	0.013**	-0.002	0.019***
	(0.005)	(0.007)	(0.007)
Female	-0.002	0.005	-0.006
	(0.004)	(0.004)	(0.005)
Black	-0.034***	0.004	-0.049***
	(0.007)	(0.010)	(0.007)
Asian	0.002	0.009	-0.004
	(0.005)	(0.007)	(0.006)
Other	-0.023***	0.002	-0.030***
	(0.006)	(0.010)	(0.007)
Hispanic	-0.008	0.015	-0.014*
	(0.006)	(0.009)	(0.007)
Took 2 Courses	-0.034***	-0.003	-0.052***
	(0.007)	(0.009)	(0.008)
Took 3 Courses	-0.070***	-0.017	-0.101***
	(0.015)	(0.014)	(0.021)
Took 4 Courses	-0.085***	-0.031*	-0.117***
	(0.017)	(0.016)	(0.026)
Took 5 Courses	-0.087***	-0.049*	-0.110***
	(0.019)	(0.025)	(0.028)
Took 6+ Courses	-0.150***	-0.103***	-0.178***
	(0.039)	(0.026)	(0.053)
Observations	95,074	35,834	59,240
R-squared	0.284	0.320	0.262
Mean of Dependent Variable	0.802	0.860	0.767

Notes: Observations are at the student-year level. The dependent variable is an indicator for the student taking the same number of exams as courses. All columns include average course grade, School-Grade and year fixed effects. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

2.7. Appendix

Table A2.1: Credit Granting AP Exam Scores

AP Course	Georgia State University	University of Georgia	Georgia Tech University	Kennesaw State University	Georgia Southern University	Perimeter College
World History	3	4	4	4	3	3
US History	3	4	4	4	3	3
English Language and Composition	3	5	4	5	3	3
Government and Politics: US	3	4	4	4	3	3
Macroeconomics	3	3	4	3	3	3
Psychology	3	3	4	3	3	3
English Literature and Composition	4	3	4	5	5	3
Human Geography	3	5	4	3	3	3
Environmental Science	4	3	4	3	4	4
Calculus AB	4	3	4	3	4	3

Notes: Shows credit granting AP exam scores in major universities and colleges in Georgia for the ten most popular AP courses in our sample. Source: College Board. Link: https://apstudents.collegeboard.org/getting-credit-placement/search-policies

Table A2.2: Subjects

District	Enrolled Non FRL	Enrolled FRL
State of GA[45]		State pays for one AP exam
One District among Districts 1 and 2	District pays for all exams that are not paid by the state or federal funds	District pays for all exams that are not paid by the state or federal funds
The Other District among Districts 1 and 2	District pays for all exams that are not paid by the state or federal funds	District pays for all exams that are not paid by the state or federal funds
One District among Districts 3 and 4	District pays for one exam	District pays for one exam. The state funds one additional exam for any course. Students who are eligible for AP exam fee reduction pay roughly \$53 (depending on the year) per exam if they take more than 2 exams
The Other District among Districts 3 and 4	District pays for one exam. All students pay a non-refundable \$10 fee.	District pays for all exams that are not paid by the state or federal funds. All students pay a non-refundable \$10 fee.

[45] See https://www.ajc.com/news/local-education/some-fear-change-exam-subsidy-slights-low-income-students/jMvPp7FznJQvZw936Jv6oM/.

Table A2.3: Determinants of AP Exam Taking, Taken Any Course in Sequence Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
FRL	-0.047***	-0.005	0.020*	0.033***	0.033***	0.019***	0.020***
	(0.016)	(0.013)	(0.012)	(0.010)	(0.010)	(0.005)	(0.005)
Female		0.008**	-0.012***	-0.012***	-0.012***	-0.011***	-0.012***
		(0.003)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)
Black		-0.116***	-0.057***	-0.033***	-0.031***	-0.041***	-0.039***
		(0.022)	(0.020)	(0.011)	(0.011)	(0.006)	(0.005)
Asian		0.031***	0.009	0.015**	0.016**	0.002	0.003
		(0.010)	(0.011)	(0.007)	(0.007)	(0.004)	(0.004)
Other		-0.041***	-0.021	0.001	0.001	-0.016**	-0.015**
		(0.014)	(0.015)	(0.013)	(0.013)	(0.007)	(0.007)
Hispanic		-0.046***	-0.018	0.001	0.002	-0.014**	-0.011**
		(0.011)	(0.011)	(0.011)	(0.011)	(0.006)	(0.005)
Grade Level = 10			0.016	0.046	0.024	0.039	-0.059***
			(0.034)	(0.032)	(0.039)	(0.027)	(0.014)
Grade Level = 11			-0.022	-0.011	0.023	-0.014	-0.046***
			(0.034)	(0.030)	(0.039)	(0.025)	(0.015)
Grade Level = 12			-0.098***	-0.079**	0.005	-0.084***	-0.091***
			(0.035)	(0.031)	(0.041)	(0.027)	(0.012)
Numeric Grade			0.011***	0.011***	0.011***	0.012***	0.012***
			(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.876***	0.892***	-0.102*	-0.040	-0.097*	-0.121***	-0.110***
	(0.019)	(0.019)	(0.057)	(0.049)	(0.057)	(0.042)	(0.040)
Observations	212,935	212,935	212,935	212,935	212,935	212,935	212,935
R-squared	0.030	0.049	0.143	0.164	0.171	0.219	0.243
Year FE	YES	YES	YES	YES	YES	YES	YES
Course FE	YES	YES	YES	YES	YES	YES	YES
District FE				YES			
District-Grade FE					YES		
School FE						YES	

School-Grade FE							YES
Mean of Dependent Variable	0.808	0.808	0.808	0.808	0.808	0.808	0.808

Notes: Observations are at the student-course-year level. The dependent variable is a binary variable for taking the AP exam. The sample includes students who took any course in the AP course sequence. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table A2.4: Number of Exams Taken, Taken Any Course in Sequence Sample

VARIABLES	(1) Full Sample	(2) Districts 1 & 2 *Same subsidy for all*	(3) Districts 1 & 2: FRL Only	(4) Districts 1 & 2: Non-FRL Only	(5) Districts 3 & 4 *Different subsidy by FRL*	(6) Districts 3 & 4: FRL Only	(7) Districts 3 & 4: non-FRL Only	(8) Full Sample with Number of Subsidized Tests
FRL	0.037***	-0.008			0.054***			-0.023**
	(0.010)	(0.009)			(0.012)			(0.011)
Female	-0.008	-0.007	0.000	-0.009	-0.008	0.009	-0.020**	-0.008
	(0.006)	(0.006)	(0.015)	(0.007)	(0.008)	(0.010)	(0.009)	(0.006)
Black	-0.053***	-0.019	-0.025	-0.015	-0.069***	-0.020	-0.075***	-0.057***
	(0.009)	(0.014)	(0.023)	(0.015)	(0.011)	(0.015)	(0.015)	(0.010)
Asian	0.019**	0.006	0.061*	-0.001	0.015	0.077***	0.002	0.011
	(0.009)	(0.013)	(0.031)	(0.012)	(0.011)	(0.015)	(0.015)	(0.008)
Other	-0.023**	-0.002	-0.004	-0.002	-0.031**	0.005	-0.053***	-0.028***
	(0.010)	(0.011)	(0.035)	(0.011)	(0.013)	(0.015)	(0.015)	(0.011)
Hispanic	-0.011	0.004	-0.002	0.010	-0.017	0.006	-0.009	-0.014
	(0.009)	(0.014)	(0.026)	(0.013)	(0.011)	(0.012)	(0.016)	(0.009)
Avg. Course Grade	0.019***	0.015***	0.013***	0.017***	0.021***	0.019***	0.023***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
#Courses	0.861***	0.924***	0.873***	0.929***	0.832***	0.876***	0.806***	0.811***
	(0.017)	(0.015)	(0.030)	(0.014)	(0.022)	(0.019)	(0.027)	(0.024)
#Subsidized Tests								0.094***
								(0.020)
Observations	111,115	39,312	8,351	30,961	71,803	28,163	43,640	111,115
R-squared	0.782	0.847	0.717	0.862	0.751	0.760	0.747	0.784

Mean of Dependent Variable	1.549	1.582	1.202	1.685	1.531	1.423	1.601	1.549

Notes: Observations are at the student-year level. The dependent variable is the number of AP exams taken by the student in a year. All columns include School-Grade and year fixed effect. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table A2.5: Number of Exams Taken, # Course Bins

VARIABLES	(1) Full Sample	(2) Districts 1 & 2 *Same subsidy for all*	(3) Districts 1 & 2: FRL Only	(4) Districts 1 & 2: Non-FRL Only	(5) Districts 3 & 4 *Different subsidy by FRL*	(6) Districts 3 & 4: FRL Only	(7) Districts 3 & 4: non-FRL Only	(8) Full Sample with Number of Subsidized Tests
FRL	0.025***	-0.012			0.040***			-0.013
	(0.009)	(0.008)			(0.011)			(0.009)
Female	0.001	0.004	0.010	0.003	-0.001	0.014*	-0.011	0.001
	(0.005)	(0.006)	(0.011)	(0.007)	(0.006)	(0.008)	(0.006)	(0.005)
Black	-0.049***	-0.000	-0.004	0.001	-0.069***	-0.036**	-0.069***	-0.052***
	(0.010)	(0.013)	(0.024)	(0.014)	(0.012)	(0.016)	(0.016)	(0.011)
Asian	0.015*	0.020	0.066**	0.013	0.007	0.050***	0.001	0.011
	(0.008)	(0.012)	(0.030)	(0.011)	(0.009)	(0.015)	(0.012)	(0.007)
Other	-0.028***	0.010	0.020	0.008	-0.038***	-0.011	-0.055***	-0.031***
	(0.007)	(0.011)	(0.027)	(0.013)	(0.010)	(0.010)	(0.016)	(0.008)
Hispanic	-0.012	0.021*	0.018	0.019	-0.023*	-0.010	-0.013	-0.014
	(0.010)	(0.012)	(0.024)	(0.011)	(0.012)	(0.016)	(0.016)	(0.010)
Avg. Course Grade	0.013***	0.008***	0.007***	0.009***	0.015***	0.013***	0.017***	0.013***
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Took 2 Courses	0.875***	0.931***	0.904***	0.941***	0.844***	0.883***	0.812***	0.828***
	(0.015)	(0.021)	(0.026)	(0.022)	(0.017)	(0.014)	(0.026)	(0.022)
Took 3 Courses	1.779***	1.885***	1.843***	1.896***	1.717***	1.760***	1.679***	1.691***
	(0.030)	(0.036)	(0.051)	(0.035)	(0.041)	(0.044)	(0.050)	(0.044)
Took 4 Courses	2.711***	2.828***	2.666***	2.849***	2.641***	2.727***	2.580***	2.577***
	(0.044)	(0.061)	(0.133)	(0.056)	(0.057)	(0.069)	(0.063)	(0.061)
Took 5 Courses	3.653***	3.758***	3.560***	3.780***	3.592***	3.778***	3.474***	3.475***

	(0.054)	(0.073)	(0.194)	(0.065)	(0.074)	(0.067)	(0.089)	(0.081)
Took 6+ Courses	4.675***	4.680***	4.223***	4.715***	4.651***	5.130***	4.429***	4.475***
	(0.131)	(0.096)	(0.511)	(0.079)	(0.174)	(0.151)	(0.193)	(0.158)
#Subsidized Tests								0.077***
								(0.018)
Observations	95,074	35,834	7,655	28,179	59,240	23,393	35,847	95,074
R-squared	0.808	0.863	0.755	0.876	0.775	0.786	0.771	0.809
Mean of Dpt. Variable	1.506	1.584	1.264	1.671	1.459	1.381	1.510	1.506

Notes: Observations at the student-year level. The dependent variable is the number of AP exams taken. Courses over are 6 are grouped into the 6 bins. Course = 1 is the omitted category. All columns include School-Grade and year fixed effects. Standard errors are clustered at the school level. *** p<0.01, ** p<0.05, * p<0.1.

Table A2.6: Took All Exams, Ninth-Eleventh Grade and Twelfth Grade Samples

Panel A: Ninth-Eleventh Grade Sample

	(1) Full Sample	(2) Districts 1&2	(3) Districts 3&4
FRL	0.001	-0.003	0.004
	(0.005)	(0.006)	(0.006)
Took 2 Courses	-0.020***	0.007	-0.035***
	(0.006)	(0.009)	(0.007)
Took 3 Courses	-0.047***	-0.007	-0.068***
	(0.014)	(0.010)	(0.020)
Took 4 Courses	-0.066***	-0.014	-0.093***
	(0.020)	(0.014)	(0.029)
Took 5 Courses	-0.077***	-0.014	-0.106***
	(0.023)	(0.016)	(0.032)
Took 6+ Courses	-0.080	0.021	-0.118
	(0.075)	(0.024)	(0.095)

Observations	69,038	25,018	44,020
R-squared	0.225	0.320	0.186
Mean of Dependent Variable	0.861	0.897	0.841

Panel B: Twelfth Grade Sample

	(1) Full Sample	(2) Districts 1&2	(3) Districts 3&4
FRL	0.041***	-0.002	0.061***
	(0.011)	(0.013)	(0.014)
Took 2 Courses	-0.065***	-0.020	-0.094***
	(0.013)	(0.017)	(0.017)
Took 3 Courses	-0.110***	-0.031	-0.166***
	(0.022)	(0.024)	(0.031)
Took 4 Courses	-0.114***	-0.046*	-0.159***
	(0.019)	(0.023)	(0.029)
Took 5 Courses	-0.112***	-0.065*	-0.142***
	(0.022)	(0.032)	(0.032)
Took 6+ Courses	-0.187***	-0.129***	-0.225***
	(0.036)	(0.028)	(0.048)
Observations	26,036	10,816	15,220
R-squared	0.266	0.286	0.204
Mean of Dependent Variable	0.646	0.775	0.554

Notes: Observations are at the student-year level. The dependent variable is an indicator for the student taking the same number of exams as courses. Panel A restricts the sample to students in grades 9-11, while Panel B restricts to students in grade 12. All columns include School-Grade and year fixed effects, and control for average course grade, Asian, Black, Other, and Hispanic. *** p<0.01, ** p<0.05, * p<0.1.

2.8. Appendix: Additional Details of the Data

AP Test Data

The format in which the raw AP score data arrived was not uniform across districts. Two districts provide one file per school year; such files contain test scores for students who took a test in the current year and in previous years. For these files, we only consider exams for the year of the file. We also drop the small number (far less than 1 percent) of cases for which a student-year appears multiple times in a file (which we think mainly occur due to imperfect fuzzy matching). The other two districts provided us a file that they had already cleaned; if we observe a student-year-test that appears twice, we keep the one with the higher score. For all data, if we are missing a unique identifier—which prevents us from matching to the course data—we drop these instances. Dropping these observations and the observations identified earlier in the paragraph will cause us to very slightly underestimate the percentage of students taking the corresponding AP exam.

In cases where we observe that the student took a test, but do not observe the test score (often because the test is still pending or the score has been canceled) or being coded as having a 0 score, we code the student as having taken the test, but record the score variable as missing. Thus, these individuals are not used in the first step of the prediction exercise.

Course Data

We first process the course data within district. In the very rare instances that a student has multiple observations for a district-year-course-term, we keep the observation that is not a transfer credit (if one is and one is not), and then break ties with the higher credits earned and then higher exam score. In some instances, students do not take semester courses (S1, S2) and take one of four 9-week terms (N1, N2, N3, N4) or year-long courses (Y1). In order to decide which observation is the terminal course (for the final sample), we create a hierarchy as follows, where we choose the observation appearing latest: N1, N2, S1, N3, N4, S2, Y1.

Not all subjects appear in all districts. Two AP courses map to the same English AP exam. We treat both courses as being the same. In some cases, the observation in the main data set is coded with a teacher denoting that it is a transfer course; they represent less than 0.1 percent of observations of the data (before restricting to the terminal course dataset), and we keep them in the data.

It is very rare that a student's course will be associated with multiple teachers. In these cases, we only consider the first-listed teacher and use his/her demographics. We construct the teacher gender variables within district-year, but using information from three separate files. In cases of disagreement between the files, we assign them female if any are female.

Combined Data

When we aggregate from the student-course-year level to the student-year level, there is a very small number of cases in which variables are not constant within student-year. If a student transferred schools or districts, their demographics could vary within year. If they do, we consider the student to be female if any observation if female, other race if not all observations have the same race, Hispanic if any observation is Hispanic, and FRL if any observation is FRL. We assign them the latest-occurring school and district if unique. In the extremely rare cases of ties, we go with the school/district in which the course was not a transfer course and in the other instance the school/district that had the highest number of courses. When constructing the sample for the Full Sample of Table 2.1 (Summary Statistics), we use all students who appear in both the demographic and course data files, keeping the observation with the highest grade in rare cases of a student appearing with multiple grades. We follow a similar process as above in cases of transfer students; if there was still a tie after the above process, we break it randomly.

3. ACA Coverage Impacts in the Trump Era[46]

3.1. Introduction

Almost halfway through the fourth year of the Trump administration, the Affordable Care Act (ACA) and its impact on insurance coverage remains a contentious political issue. The most recent example is the fact that while eleven states and the District of Columbia have opened enrollment in their state Marketplaces to allow workers laid-off as a result of the coronavirus, the Trump administration is not currently planning to open the federal Marketplace. Despite the fact that the president has stated Republicans are doing a great job managing the ACA, his administration is supporting the challenge to the ACA's constitutionality that will go before the Supreme Court this fall. Given this ongoing debate it is natural to ask how insurance coverage under the ACA has fared during the Trump administration.

The purpose of this paper is to estimate the causal effect of the ACA on insurance coverage by type during 2017 and 2018, the first two years of the Trump administration, using data from the American Community Survey (ACS). While a large literature has developed on the coverage impacts of the ACA (Courtemanche, Marton, and Yelowitz, 2016; Obama, 2016; Gruber and Sommers, 2019), to our knowledge only one prior publication (Courtemanche et al., 2020b) has included both 2017 and 2018 data. Courtemanche et al. (2020b) found continued gains in overall coverage in 2017 and 2018, but were unable to differentiate between different types of insurance

[46] This chapter is co-authored with the following people. Charles Courtemanche: U. of Kentucky, Dept. of Economics, Gatton College of Business and Economics, Lexington, KY, 40506-0034, USA; 859-323-7900; courtemanche@uky.edu. James Marton, Georgia State U., P.O. Box 3992, Atlanta, GA 30302-3992, USA; 404-413-0256; marton@gsu.edu. Benjamin Ukert: Texas A&M U., Dept. of Health Policy and Management, 212 Adriance Lab Road, 1266 TAMU, College Station, TX 77843-1266; 979-436-9056; bukert@tamu.edu. Aaron Yelowitz: U. of Kentucky, Dept. of Economics, Gatton School of Business and Economics, Lexington, KY, 40506-0034, USA; 859-257-7634; aaron@uky.edu. Daniela Zapata: Impaq International, 11325 G Street, NW, Suite 900, Washington, DC 20005, USA; 202-774-1981; dzapata@impaqint.com.

coverage given the way insurance information is collected in the Behavioral Risk Factor Surveillance System (BRFSS). Thus our paper contributes to the literature by providing estimates on the impact of the ACA on overall coverage in 2017 and 2018 from the ACS to compare to those from the BRFSS, as well as being the first to provide estimates by type of coverage.

There were several key events associated with the ACA occurring in 2017 and 2018 which might be expected to influence the effect of the ACA on insurance coverage. First, President Trump's first executive order in January of 2017 encouraged the federal government to waive or delay the implementation of any features of the ACA that would impose a burden, either financial or regulatory (White House, 2017). Second, funding for ACA outreach and education programs was reduced for open enrollment periods associated with 2017 and 2018 plans (GAO, 2018). Third, the administration discontinued cost sharing reduction (CSR) payments to insurers for silver Marketplace plans in October 2017 (HHS, 2017). Finally, political debate surrounding the ACA persisted, including the failed vote to repeal the ACA in July 2017 and the vote to pass the tax reform package that included a repeal of the ACA individual coverage mandate in December 2017 (Commonwealth Fund, 2018). The addition of 2017 and 2018 ACS data allows us to examine the initial causal impact of these events on insurance coverage by type, such as private vs. Medicaid coverage. Recent descriptive evidence suggests that coverage actually fell by about 0.5 percentage points between 2017 and 2018 (Berchick et al., 2019).

Our primary methodological approach, borrowing from the recent ACA literature, involves estimating difference-in-difference-in-differences (DDD) models with the differences coming from time, state Medicaid expansion status, and local area pretreatment uninsured rate in order to estimate the impact of the fully implemented ACA (Courtemanche et al., 2017; Courtemanche et al., 2018a; Courtemanche et al., 2018b). This approach differs from much of the initial ACA literature, which employed simpler difference-in-differences (DD) models comparing changes in expansion states to non-expansion states in order to identify the effect of the ACA Medicaid expansion alone. Identifying the impact of the national components of the ACA, such as the individual mandate and subsidized Marketplace coverage, requires a different approach because they were implemented in every state at the same time. The inclusion of a third difference handles the fact that the national components of the ACA were implemented in every state at the same time. Our third difference identifies the combined effect of these national components because they should provide the most intense "treatment" in local areas with the highest uninsured rates prior to the ACA.

In non-expansion states, which should be impacted by the national components of the ACA alone, we see statistically significantly smaller coverage increases in 2017 and 2018 (3.8 percentage points in each year relative to pre-ACA) as compared to 2016, where we estimate a 5 percentage point increase relative to pre-ACA. In expansion states, which should be impacted both by the national components of the ACA and the Medicaid expansion, we estimate similar gains in coverage in 2016 through 2018 (about 11 percentage points in each year relative to pre-ACA). This difference between expansion and non-expansion states is due smaller year over year increases in coverage due to the national components of the ACA and larger year over year increases due to the Medicaid expansion.

3.2. Data

Our analysis uses data from 2011 through 2018 waves of the American Community Survey (ACS). The ACS is the most comprehensive survey of the U.S. population across all 50 states and the District of Columbia, sampling roughly one percent annually or about three million individual respondents per year. Its mandatory nature reduces sample selection concerns and provides a high and stable response rate. For the purpose of our study we restrict the sample to respondents aged 19 to 64 because this was the ACAs target population. We only use data starting in 2011 to avoid measuring the ACAs effect of provisions enacted in 2010 and because the three pre-2014 ACA years provide us with a pre-treatment period which we use to evaluate the plausibility of the assumptions of our econometric model.

In terms of geographic identifiers, the ACS includes a state identifier for each respondent and an identifier for the respondents Public Use Microdata Area (PUMA), which represent an area within state of at least 100,000 people. The PUMA identifier is important for our study because it better represents local characteristics than a state identifier. Our empirical strategy relies on within-state variation in uninsured rates in 2013 to identify the causal effect of the ACA marketplace. Ideally, we would use the local PUMAs within a state to identify uninsured rates, however the PUMA definition changed during our sample period, and the new boundaries were applied to the 2013 ACS wave and later. As a result, we cannot continuously identify the respondents' geographic area from 2011 to 2018. We follow Courtemanche et al. (2017) and identifying core-based statistical areas (CBSAs) identifiable in all years using the old and new PUMA classification systems. A complicating factor with CBSAs is that they can span multiple states and we isolate the portion of the CBSA in each state as separate local areas. CBSAs also do not cover all areas within a state, to avoid dropping respondents in unassigned areas we create additional local areas for the non-CBSA portion in each state. These adjustments avoid dropping respondents. After making these aforementioned adjustments, our dataset consists of 630 local CBSA and non-CBSA areas that each contain between 356 and 78,781 respondents in 2013, with a median of 1,020 and a mean of 2,811 respondents.

The ACS includes a wide number of socio-demographic characteristics for each respondent. We take advantage of a question on health insurance coverage to estimate the 2013 local uninsured rates and also to evaluate the effect on insurance coverage after the 2014 ACA provisions took effect. Specifically, the ACS health insurance question asks at the time of the survey if the person is currently coved by any time of health insurance and provides a list of eight types of health insurance categories to choose from. These include "insurance though a current or former employer or union", "insurance purchased directly from an insurance company", "Medicare", "Medicaid, Medical Assistance, or any kind of government-assistance plan for those with low incomes or a disability", "TRICARE or other military health care", "VA (including those who have ever used or enrolled for VA health care)", "Indian Health Service", and "any other type of health insurance or health coverage plan." Answers are not mutually exclusive, and a respondent can choose to answer more than one type of coverage or choose to answer "no" to all eight categories as uninsured. Based on the health insurance question we create six indicator variables that are equal to one if an individuals responded "yes" to any insurance option, any private insurance (either employer sponsored or directly purchased), employer-sponsored insurance, directly purchased insurance, Medicaid, and any other coverage (defined as coverage as neither private nor Medicaid coverage).

Other socioeconomic variables in the ACS that we use as control variables in our models cover gender, race/ethnicity, origin, family structure, education, labor force participation, and household income. Specifically, we create binary variables for the age of the respondent (binary indicators for each year from 19 to 64), female, race/ethnicity (non-Hispanic white, non-Hispanic black, Hispanic, and other race/ethnicity), foreign born, U.S. citizenship status, married, and separate binary variables counting the number of children under the age of 18 living in the household (one, two, three, four, and five or more). We group educational attainment into indicator for the highest level of completed education (less than a high school degree, high school degree, some college, and college graduate), and we measure labor force participation with binary variables indicating whether the individual is a student or not, whether the individual is unemployed or not. We measure household income in our regression relative to the Federal Poverty Limit (FPL), and then create 50 dummy variables measuring the separate impact for each 10-point increment of income as a percentage of the FPL (with the highest category including everyone over 500 percent of the FPL). Lastly, we also include as a control variable the annual state unemployment rate collected from the Bureau of Labor Statistics.

Our second binary independent variable of interest, whether and when a state expanded Medicaid via the ACA, was collected from the Kaiser Family Foundation (KFF, 2020). The majority of states expanded Medicaid effective January 1, 2014, but several states expanded at a later time. At the end of 2018, 31 states and the District of Columbia expanded their Medicaid program and seven states expanded Medicaid after January 1, 2014. We assign the starting date of these states' Medicaid expansions in our expansion indicator accordingly. We also collect information from the Kaiser Family Foundation (KFF, 2020) and Kowalski (2014) regarding the implementation of the 2014 ACA insurance marketplaces. Several states struggled with the initial rollout of the marketplace enrollment platform. We include binary variables indicating whether a state set up their own state-run insurance exchange and whether the exchange experienced glitches in 2014. We include these variables to control for differential responses in take-up in the individual marketplace related to initial troubles in outreach and sign-up of uninsured individuals.

Summary statistics are displayed in table 3.1 for our six health insurance coverage outcomes in 2013 for the full sample as well as stratified by state Medicaid expansion decision and whether the uninsurance rate was above or below the median. In the full sample the average baseline insurance rate was 79 percent in 2013, with 67 percent having private coverage, 60 percent having employer sponsored coverage, 9.4 percent having individually purchased coverage, about 11 percent having Medicaid coverage, and 3 percent having other types of coverage. In Medicaid expansion states (columns 2 and 3) the percent of individuals reporting any coverage was slightly higher than in non-expansion states (columns 4 and 5). The somewhat higher average level of 2013 coverage stems from a higher proportion of individuals having Medicaid coverage in (future) expansion states. These baseline differences provide the first evidence that 2013 uninsured rates and Medicaid expansion status may be appropriate sources of variation for our analysis. To give a rough sense of how coverage changes over time, figure 3.1 plots changes in insurance coverage by type over time for our sample, also stratified by expansion status and 2013 uninusrance rate.

3.3. Methods

In order to uncover the causal impact of the ACA on coverage disparities after five years, we follow the recent ACA literature by estimating DDD models with the differences come from time, state Medicaid expansion decisions, and pre-ACA local area uninsured rates (Courtemanche et al., 2017; Courtemanche et al., 2018a; Courtemanche et al., 2018b). Our baseline DDD regression equation is given by equation (1) below.

$$\begin{aligned} y_{iast} &= \gamma_0 + \gamma_1(UNINSURED_{as} * POST_t) + \gamma_2(MEDICAID_{st} * POST_t) + \\ &\gamma_3(UNINSURED_{as} * MEDICAID_{st} * POST_t) + \gamma_4\boldsymbol{X_{iast}} + \theta_t + \alpha_{as} + \varepsilon_{iast} \end{aligned} \tag{1}$$

where

- y_{iast} is an indicator of insurance coverage for individual *i* in local area *a* in state *s* in year t,
- $POST_t$ is an indicator for whether period *t* is in the post-reform period of 2014 or later,

- X_{iast} is a vector of control variables previously described,
- $MEDICAID_{st}$ is an indicator for whether state *s* participated in the ACA's Medicaid expansion in year *t*,
- $UNINSURED_{as}$ is the 2013 (pre-ACA) uninsured rate in local area *a* within state *s*,
- θ_t denotes year fixed effects,
- α_{as} denotes local area fixed effects,
- and ε_{iast} is a standard error term.

The term $POST_t$ is not separately included in equation (1) since it is absorbed by the year fixed effects, while the terms $UNINSURED_{as} * MEDICAID_{st}$ are not separately included since they are absorbed by the local area fixed effects.

The effect of the ACA without the Medicaid expansion is given by $\gamma_1 * UNINSURED_{as}$, which means it is assumed to be zero in a (hypothetical) area with a 0 percent uninsured rate at baseline and to increase linearly as the pre-ACA uninsured rate rises (Courtemanche et al., 2017). Similarly, the effect of the Medicaid expansion alone is given by $\gamma_3 * UNINSURED_{as} * MEDICAID_{st}$, meaning it is zero in non-expansion states (where $MEDICAID_{st} = 0$) and $\gamma_3 * UNINSURED_{as}$ in expansion states (where $MEDICAID_{st} = 1$). We consider γ_2 to represent unobserved confounders rather than capturing part of the expansion's causal effect, since the Medicaid expansion should not causally affect coverage in an area with a 0 percent baseline uninsured rate. The effect of the "fully implemented" ACA, i.e. in Medicaid expansion states, combines the impacts of the Medicaid and non-Medicaid components: $\gamma_1 * UNINSURED_{as} + \gamma_3 * UNINSURED_{as}$. In our results we report the predicted effect of the ACA at the sample mean pretreatment uninsured rate. Formally, this predicted effect is given by $\gamma_1 * \overline{UNINSURED_{as}}$ in non-expansion states and $\gamma_1 * \overline{UNINSURED_{as}} + \gamma_3 * \overline{UNINSURED_{as}}$ in expansion states.

While estimates based on equation (1) provide average effects over the 2014-2018 time period, we are primarily interested in how the effects varied over time across these five years, especially in 2017 and 2018. In order to analyze changes over time, we estimate event-study models as our preferred set of specifications, where we replace $POST_t$ with a set of year dummies. The event study DDD model is

$$y_{iast} = \varphi + \sum_{t=1}^{T} \theta_t (UNINSURED_{as} * Y_t) + \sum_{t=1}^{T} \alpha_t (MEDICAID_s * Y_t) \\ + \sum_{t=1}^{T} \beta_t (UNINSURED_{as} * MEDICAID_s * Y_t) + \delta X_{iast} + \alpha_{as} + \varepsilon_{iast} \tag{2}$$

where Y_t, is an indicator for whether year t is 2011, 2012, …, 2018, respectively for t = 1, 2,…,7, with 2013 being the omitted reference year and the other terms being as described in equation (1). Here the effects of the ACA without the Medicaid expansion during 2014, 2015, …, 2018 are given by $\theta_3 * UNINSURED_{as}$, $\theta_4 * UNINSURED_{as}$, …, $\theta_7 * UNINSURED_{as}$ respectively, while the effects of the Medicaid expansion in 2014, 2015, …, 2018 are similarly given by $\beta_3 * UNINSURED_{as}$, $\beta_4 * UNINSURED_{as}$, … $\beta_7 * UNINSURED_{as}$.

This event study model also allows us to test the identifying assumptions from our main DDD specification (Courtemanche et al., 2017; Courtemanche et al., 2018a). The identifying assumption for the effect of ACA without Medicaid expansion is that, in the absence of the ACA, any changes in the outcomes that would have occurred in 2014–2018 would not have been systematically correlated with local area uninsured rates, conditional on the controls. The identifying assumption for the impact of Medicaid expansion is that, in the absence of ACA, the differential changes in the outcomes in 2014–2018 between Medicaid expansion and non-expansion states would not have been correlated with pre-reform uninsured rates. If the event study suggests evidence that changes in the outcomes from 2011-2013 are correlated with pre-ACA uninsured rates (i.e. θ_1 or θ_2 are significant) or the interaction of the local area uninsured rate with Medicaid expansion status (i.e. β_1 or β_2 are significant), this would suggest problems with these assumptions.

We also estimate a number of robustness checks. The first one is excluding 19-25-year olds, who were affected by the dependent coverage provision of the ACA that was implemented in 2010; thus, were partially affected prior to 2013. Robustness checks two to six consist of dropping states that expanded Medicaid at different points in time (KFF, 2020). In the second and third checks we drop early expansion states using two different specification - states that expanded between April 2010 and March 2012[47] and states that expanded prior to 2010[48]. In the fourth, fifth and sixth checks we drop all early expanders, all late expanders and early and late expanders respectively. In the seventh specification check we include *Medicaid*Uininsured* in the DDD analysis and in the eighth specification check we use state level uninsured rates with additional state controls.

3.4. Results

Table 3.2 provides our combined post-ACA results. Here we report implied effects of the ACA that multiply the coefficient estimates from equation (1) by the local average pre-treatment 2013 uninsured rate (20.3 percent). Each column represents results from a different regression that measures the effect of the ACA on a different coverage outcome. The first row displays the implied effect of the national portion of the ACA, including the health insurance marketplaces, the

[47] These states include California, Connecticut, Washington DC, Minnesota, New Jersey, and Washington.
[48] These 5 states are: Delaware, Washington DC, Massachusetts, New Your, and Vermont. These early expanders were also excluded by Kaestner et al., 2017.

individual mandate, etc., while the second row displays the effect for the Medicaid expansion alone. The last row displays the full effect of the ACA, which is the expected change for an area that was treated both by the national portion of the ACA as well as the Medicaid expansion.

Table 3.2 shows that the probability of having any coverage increased by 3.7 percentage points from the national portion of the ACA between 2014 and 2018 relative to pre-ACA period. The Medicaid expansion contributed a 5.3 percentage point increase in having any coverage during the same time period. The full effect of the ACA, which is a combination of these two effects, was a 9.1 percentage point increase in having any coverage. The fully implemented ACA also increased private coverage by 2.6 percentage points (column 2), which was driven by the national portion of the law (a 3.3 percentage point increase). Employer sponsored coverage increased by 1.8 percentage points (column 3) and individually purchased coverage increased by 0.9 percentage points due to the fully implemented ACA (column 4). Finally, Column 5 shows that the fully implemented ACA increased Medicaid coverage by 6.8 percentage points and that effect is driven by the Medicaid expansion (which represents a 6.2 percentage point increase).

The validity of the identifying assumptions of our baseline DDD model cannot be tested directly. We can, however, test their validity indirectly by examining the pre-ACA coefficients from an event study model where we interact the treatment variables with the full set of year fixed effects, treating 2013 to as the base year. We report the results from this event study analysis separately in tables 3.3 and 3.4. Table 3.3 reports only the estimated pre-treatment (2011 and 2012) coefficients from the event study, while table 3.4 reports only the post-treatment implied effects (for the years 2014, 2015, 2016, 2017, and 2018). This split allows us to focus on the indirect test of our identifying assumptions in table 3.3 and decomposing the year-by-year effects of the ACA in the post-period in table 3.4.

With that in mind, we first examine the results from table 3.3. The numbers reported in table 3.3 are not implied effects but estimated coefficients taken directly from the estimation of equation (2). Ideally, we want to see no more of the pre-2013 coefficients to be significant than we should expect by chance, which is around 5 percent. A substantially higher percentage would call our identification strategy into question by suggesting an "impact" of the ACA even before it was implemented. We see that 2 out of 24 coefficients (8.3 percent) are significant in table 3.3.[49] Moreover, results of an F test for the joint significance of all pre-ACA interactions show that we can reject the null that there is no effect in just one instance, the effect of ACA without the Medicaid expansion on individually purchased insurance.[50] In all other cases, we are unable to reject the null that there was no effect of the national components of ACA or the Medicaid expansion alone in 2012 and 2013 on the various different sources of insurance coverage. These findings give us confidence in a casual interpretation of our results.

Table 3.4 reports the post-treatment implied effects (for the years 2014, 2015, 2016, 2017, and 2018) from the same event study model that generates the results reported in table 3.3. Along with the indicators for statistical significance of the change in insurance coverage relative to the base year of 2013, this table includes indicators to show if the estimates are statistically significantly different in 2017 and 2018 (the first two years of the Trump administration) relative to 2016. Three sets of implied effects are reported in table 3.4: the impact of the ACA without the Medicaid expansion in panel I, the impact of Medicaid expansion alone in panel II, and the impact of the full ACA with Medicaid expansion (which is the sum of the first two effects) in panel III.

[49] The two pre-reform significant effects show up on the coefficient for Uninsurance Rate*2011 and Uninsurance Rate*2012 when the dependent variable is individually purchased insurance. This is the coefficient used to calculate the effect of the ACA without Medicaid expansion.

[50] We run two F tests for each of the six different outcomes (sources of insurance). The first F test for each outcome tests if the coefficients for Uninsurance rate*Medicaid Expansion*2011 and Uninsurance rate*Medicaid Expansion*2012 are jointly significant or not. The second F test for each outcome tests if Uninsurance rate*2011 and Uninsurance rate*2012 are jointly significant or not.

Our primary interest here is whether or not we see changes in the impact of the ACA in 2017 and 2018 as compared to previous years. Panel I reports the impact of the national components of the ACA, which led to statistically significantly smaller overall coverage increases in 2017 and 2018 (3.8 percentage points in each year) as compared to 2016, where we estimate a 5 percentage point increase. These smaller increases in overall coverage due to the national components of the ACA are being driven primarily by smaller gains from employer sponsored insurance (column 3) which in turn contributes to smaller gains in any private insurance.

Panel II reports the impact of the Medicaid expansion alone, where we see statistically significantly larger increases in coverage in 2017 (6.9 percentage points) and 2018 (6.7 percentage points) as compared to 2016 (with a 5.7 percentage point increase). Thus, while panel I suggests the impact of the national components of the ACA fell during the first two years of the Trump administration, panel II suggests that the impact of the Medicaid expansion grew. Not surprisingly, column 5 of panel II shows that the growth in overall coverage due to the Medicaid expansion can be attributed to increases in the likelihood of reporting Medicaid coverage.

The results from panel I and panel II allow us to examine the differential impact of the ACA on insurance coverage in expansion and non-expansion states. Since non-expansion states were only exposed to the national components of the ACA, the results from panel I represents the expected impact of the ACA on a typical non-expansion state. Therefore, in a typical non-expansion state coverage growth due the ACA fell in 2017 and 2018 as compared to 2016. Expansion states, on the other hand, were exposed to both the national components of the ACA (panel I) and the Medicaid expansion (panel II). The combination of these effects are reported in panel III of table 3.4. Panel III suggests that in a typical expansion state increases in coverage growth due to the Medicaid coverage in 2017 and 2018 compared to 2016 is offset by smaller rates of coverage growth due to the national components of the ACA. Thus, the fully implemented ACA led to a 10.8 percentage point increase in coverage both 2016 and 2017 and a 10.5 percentage point increase in 2018 in a typical expansion state (i.e. a plateauing of coverage growth).

Table 3.5 presents the results of a series of additional specification checks to further assess the validity of our combined post-period results from table 3.2. In panel I we exclude 19 to 25-year olds because they should have been mostly "treated' with the ACA dependent care coverage mandate that came into effect at the end of 2010. We observe a somewhat smaller point estimate for the effect of Medicaid expansion on any coverage, which also results in a smaller effect of the fully implemented ACA on any coverage. This is perhaps not surprising given that the young adults population was targeted by the ACA due to their relatively high rates of pre-reform uninsurance.

In the next five panels we address concerns regarding the timing of state Medicaid expansion decisions. Panels II and III drop early expansion states using two different classifications of such states and re-estimates our baseline models. In panel II, the states that expanded between April 2010 and March 2012 (California, Connecticut, Washington DC, Minnesota, New Jersey, and Washington) are dropped.[51] In panel III, states that expanded before 2010 (Delaware, Washington DC, Massachusetts, New Your, and Vermont) according to Kaestner et al. (2017) are dropped. Panel IV restricts the sample to the 13 treatment states and 16 control states that did not have some form of Medicaid expansion prior to January 2014 in order to better isolate the full Medicaid expansion effect. Panel V drops states that expanded after January 2014.

[51] Source: Kaiser Family Foundation (https://www.kff.org/health-reform/issue-brief/states-getting-a-jump-start-on-health/).

Panel VI drops all early expanders before 2014 and late expanders after 2014. In all five models the results are generally similar to our findings in table 3.2.[52]

In panel VII, we test an alternate specification of our DDD model where we add a $UNINSURED_{as} * MEDICAID_s$ component to equation (1). We don't control for this component in the primary specification as this should be absorbed by the area fixed effect α_{as}. We find the same effects as our primary specification in this alternate specification.

Our last specification check presented in panel VIII examines the robustness of our results to a different measure of the local area pretreatment uninsured rate. To do this we aggregate the 2013 uninsurance rate to the state level and add additional controls for labor market and economic conditions at the state level. These state level controls include the percent of healthcare jobs out of all jobs, the percent of government jobs out of all jobs, and state per capita GDP. This alternate specification results in no change in the impact of the fully implemented ACA on coverage, though we do see a larger effect of national components of the ACA and a smaller effect of Medicaid expansion alone.

[52] In panel II, our first version of dropping early expanders results in a statistically insignificant effect of full ACA on individually purchased insurance. This is driven by a slightly larger crowding out effect of the Medicaid expansion on individually purchased insurance compared to table 2 (a 2 percentage points decrease in individually purchased insurance due to the Medicaid expansion in panel II vs. a 1.8 percentage points decrease in table 2) and a larger standard error. In panel IV, where the sample is restricted to the 13 treatment states and 16 control states that did not have some form of Medicaid expansion prior to January 2014, we find a statistically insignificant effect of the fully implemented ACA on employer sponsored and individually purchased insurance. Here, again, the statistically insignificant effect of the fully implemented ACA on individually purchased insurance is driven by a larger crowding out effect of the Medicaid expansion on private purchase of insurance compared to the baseline specification (a 2.1 percentage points decrease in individually purchased insurance due to the Medicaid expansion in panel IV vs. a 1.8 percentage points decrease in table 2) and a larger standard error. The statistically insignificant effects of the fully implemented ACA on employer sponsored insurance is driven by a smaller positive effect of the Medicaid expansion on employer sponsored insurance compared to baseline specification (a 0.7 percentage point increase in employer sponsored insurance due to the Medicaid expansion in panel IV vs a 1.8 percentage point increase in table 2) and a larger standard error. When the sample is restricted to only the states that expanded Medicaid in 2014 and non-expansion states (Pebl VI), we don't find any statuistically significant effect of fully implemented ACA on private coverage. Also, we find that the Medicaid component of ACA leads to some crowding out from the "Other" insurance coverage to Medicaid coverage. None of these changes, however, meaningfully change the estimated coverage impacts of the national components of the ACA, the Medicaid expansion, or the fully implemented ACA.

3.5. Discussion

In this paper we examine the impact of the ACA on insurance coverage during the first two years of the Trump administration. During this time period several changes in the management of the ACA were debated and implemented, such as reductions in outreach funding, the duration of open enrollment, the discontinuation of CSR payments, and the near repeal in 2017. Each of these changes may have potentially influenced the coverage impacts of the ACA.

In non-expansion states, which should be impacted by the national components of the ACA alone, we find statistically significantly smaller coverage increases in 2017 and 2018 (3.8 percentage points in each year) as compared to 2016, where we estimate a 5 percentage point increase. In expansion states, which should be impacted both by the national components of the ACA and the Medicaid expansion, we estimate similar gains in coverage in 2016 through 2018 (about 11 percentage points in each year). This difference between expansion and non-expansion states is due smaller year over year increases in coverage due to the national components of the ACA and larger year over year increases due to the Medicaid expansion. Thus, for expansion states we find no evidence that the administrative changes and political debate surrounding the ACA during 2017 and 2018 led to differential coverage increases as compared to 2016. However, for the non-expansion states we do see smaller coverage increases in 2017 and 2018 compared to 2016.

The relatively lower growth in coverage due to the national components of the ACA as compared to the Medicaid expansion makes sense since many of the changes brought in by the new administration are more likely to negatively influence the Marketplace rather than state Medicaid programs. We find that in 2014 and 2015, 52 percent and 50 percent of the total gains in coverage come from Medicaid. The contribution of Medicaid in coverage gain in 2017 and 2018 increased to 64 percent. This is roughly consistent with Frean et al. (2017) where they find that 60 percent of the ACA coverage gains in 2014 and 2015 come from Medicaid.

Even before the advent of the current coronavirus pandemic, debate surrounding health policy figured to play a major role in both upcoming national and state elections. Whether or not to expand Medicaid is an important policy decision facing many states. The future of the ACA more generally will be an important topic in the presidential election this fall. These ongoing debates all suggest the need to continue monitoring the evolving impact of the ACA on insurance coverage.

3.6. Figures and Tables

Figure 3.1: Changes in Insurance Coverage Over Time

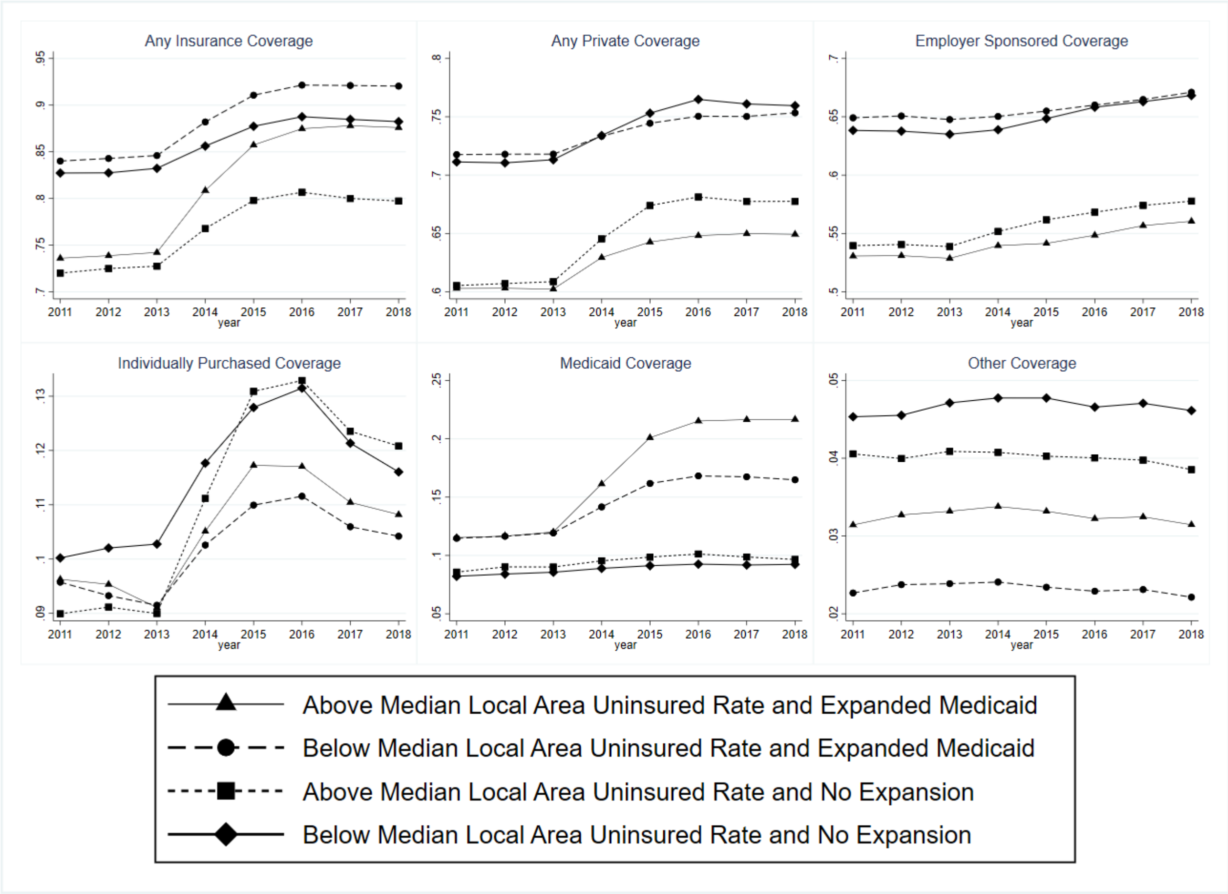


Table 3.1: Descriptive Statistics for Insurance Coverage

	Full sample	Medicaid expansion; at or above median baseline uninsured	Medicaid expansion; below median baseline uninsured	Non-expansion; at or above median baseline uninsured	Non-expansion; below median baseline uninsured
Any insurance coverage	0.792 (0.406)	0.749 (0.433)	0.848 (0.360)	0.727 (0.446)	0.832 (0.374)
Any private	0.668 (0.471)	0.618 (0.486)	0.721 (0.449)	0.608 (0.488)	0.717 (0.451)
Employer-sponsored	0.598 (0.490)	0.546 (0.498)	0.652 (0.476)	0.541 (0.498)	0.641 (0.480)
Individually purchased	0.094 (0.292)	0.093 (0.291)	0.094 (0.292)	0.090 (0.287)	0.102 (0.303)
Medicaid	0.106 (0.307)	0.114 (0.318)	0.118 (0.323)	0.089 (0.285)	0.083 (0.276)
Other	0.032 (0.176)	0.030 (0.172)	0.024 (0.152)	0.041 (0.198)	0.045 (0.207)

Notes: Standard deviations are in parentheses.

Table 3.2: Implied Effects of the ACA at Mean Pre-Treatment Uninsured Rate for Full Sample

	Any insurance	Any private	Employer-sponsored	Individually purchased	Medicaid	Other
Non-elderly adults aged 19-64 (pre-treatment uninsured rate=0.203, sample size=14,091,358)						
ACA without Medicaid Expansion	0.037** (0.012)	0.033*** (0.009)	0.006 (0.006)	0.027 (0.014)	0.006 (0.005)	-0.001 (0.001)
Medicaid Expansion	0.053*** (0.013)	-0.007 (0.011)	0.011 (0.007)	-0.018 (0.015)	0.062*** (0.008)	0.001 (0.001)
Full ACA (w/ Medicaid)	0.091*** (0.006)	0.026*** (0.005)	0.018*** (0.005)	0.009** (0.003)	0.068*** (0.007)	0.001 (0.001)

Notes: Results are effects of the ACA on the proportion of residents with the specified type of insurance, evaluated at the mean pre-treatment uninsured rate. Standard errors, heteroscedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at 0.1% level; ** 1% level; * 5% level. Sampling weights are used. All regressions include area and time fixed effects and the full set of controls.

Table 3.3: Event Study Results for Full Sample – Pre-Reform Coefficients

	Any insurance	Any private	Employer-sponsored	Individually purchased	Medicaid	Other
Non-elderly adults aged 19-64 (pre-treatment uninsured rate=0.203, sample size=14,091,358)						
Unin. Rate* Med. Exp.*2011	0.028 (0.030)	-0.014 (0.036)	-0.013 (0.037)	-0.025 (0.021)	0.033 (0.030)	0.002 (0.010)
Unin. Rate* Med. Exp.*2012	-0.012 (0.038)	-0.003 (0.034)	-0.009 (0.028)	-0.010 (0.025)	-0.003 (0.014)	-0.008 (0.005)
Uninsured Rate*2011	0.006 (0.021)	0.029 (0.032)	0.008 (0.032)	0.040* (0.015)	-0.012 (0.017)	-0.003 (0.008)
Uninsured Rate*2012	0.034 (0.031)	0.035 (0.030)	0.007 (0.023)	0.045** (0.013)	0.002 (0.006)	0.002 (0.003)

Notes: Coefficient estimates are shown. Standard errors, heteroscedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at 0.1% level; ** 1% level; * 5% level. Sampling weights are used. All regressions include area and time fixed effects, the full set of controls, Medicaid Expansion*2011, and Medicaid Expansion*2012.

Table 3.4: Event Study Results for Full Sample – Post-Reform Implied Effects

	Any insurance	Any private	Employer sponsored	Individually purchased	Medicaid	Other
Non-elderly adults aged 19-64 (pre-treatment uninsured rate=0.203)						
PANEL I: ACA without Medicaid Expansion						
ACA w/o Medicaid	0.029**	0.027***	0.010***	0.018	0.002	-0.000
Expansion 2014 (A)	(0.009)	(0.008)	(0.002)	(0.010)	(0.003)	(0.001)
ACA w/o Medicaid	0.047**	0.042**	0.010*	0.034*	0.006	0.000
Expansion 2015 (A)	(0.016)	(0.013)	(0.004)	(0.017)	(0.005)	(0.001)
ACA w/o Medicaid	0.050**	0.045**	0.011***	0.035*	0.008	-0.0003
Expansion 2016 (A)	(0.018)	(0.014)	(0.003)	(0.016)	(0.006)	(0.001)
ACA w/o Medicaid	0.038*	0.033*	0.003	0.034*	0.008	-0.001
Expansion 2017 (A)	(0.017)†††	(0.014)†††	(0.004)††	(0.016)	(0.004)	(0.002)
ACA w/o Medicaid	0.038*	0.038	0.001	0.040*	0.003	-0.002
Expansion 2018 (A)	(0.017)†††	(0.015)†††	(0.003)†††	(0.017)	(0.005)	(0.001)
PANEL II: Medicaid Expansion						
Medicaid Expansion	0.032**	0.001	0.008	-0.008	0.032***	0.000
2014 (B)	(0.010)	(0.009)	(0.004)	(0.010)	(0.007)	(0.001)
Medicaid Expansion	0.047**	-0.010	0.005	-0.018	0.060***	-0.0001
2015 (B)	(0.016)	(0.015)	(0.006)	(0.017)	(0.010)	(0.002)
Medicaid Expansion	0.057**	-0.013	0.009	-0.023	0.072***	0.001
2016 (B)	(0.018)	(0.016)	(0.006)	(0.016)	(0.010)	(0.002)
Medicaid Expansion	0.069***	-0.008	0.011	-0.023	0.078***	0.001
2017 (B)	(0.018)†††	(0.016)	(0.009)	(0.017)	(0.008)	(0.002)
Medicaid Expansion	0.067***	-0.015	0.014	-0.032	0.084***	0.002
2018 (B)	(0.018)	(0.016)	(0.009)	(0.018)	(0.009)	(0.002)
PANEL III: Full ACA						
Full ACA 2014	0.061***	0.028***	0.019***	0.009***	0.035***	-0.0001
(A+B)	(0.004)	(0.005)	(0.004)	(0.003)	(0.006)	(0.001)
Full ACA 2015	0.094***	0.032***	0.015***	0.016***	0.066***	-0.0001
(A+B)	(0.006)	(0.006)	(0.004)	(0.003)	(0.008)	(0.001)
Full ACA 2016	0.108***	0.032***	0.020***	0.012***	0.080***	0.000
(A+B)	(0.005)	(0.007)	(0.006)	(0.003)	(0.008)	(0.002)
Full ACA 2017	0.108***	0.026***	0.015	0.011**	0.086***	0.004
(A+B)	(0.005)	(0.007)	(0.008)	(0.004)	(0.007)	(0.001)
Full ACA 2018	0.105***	0.023***	0.015	0.008	0.087***	0.002
(A+B)	(0.006)	(0.006)††	(0.008)	(0.006)	(0.008)	(0.002)

Notes: Coefficient estimates are shown. Standard errors, heteroscedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at 0.1% level; ** 1% level; * 5% level. Statistically significantly different effect in 2017 and 2018 relative to 2016 is denoted by ††† at 0.1% level, †† at 1% level and † at 5% level. All regressions include area and time fixed effects, the full set of controls.

Table 3.5: Implied Effects of ACA on Health Care Access – Robustness Checks

	Any insurance	Any private	Employer-sponsored	Individually purchased	Medicaid	Other
Panel I: Exclude 19-25 Year Olds						
ACA without	0.037**	0.034***	0.008	0.027	0.004	-0.001
Medicaid Expansion	(0.012)	(0.009)	(0.006)	(0.014)	(0.005)	(0.001)
Medicaid Expansion	0.047***	-0.009	0.009	-0.017	0.057***	0.001
	(0.012)	(0.011)	(0.007)	(0.015)	(0.007)	(0.001)
ACA with Medicaid	0.084***	0.026***	0.016**	0.010**	0.061***	0.000
Expansion	(0.006)	(0.005)	(0.005)	(0.003)	(0.006)	(0.001)
Panel II: Drop ACA Early Expanders Version 1						
ACA without	0.038**	0.033***	0.006	0.027	0.006	-0.001
Medicaid Expansion	(0.012)	(0.009)	(0.006)	(0.014)	(0.006)	(0.001)
Medicaid Expansion	0.045**	-0.004	0.014	-0.020	0.054***	-0.001
	(0.015)	(0.011)	(0.008)	(0.015)	(0.013)	(0.002)
ACA with Medicaid	0.083***	0.029***	0.020***	0.007	0.060***	-0.001
Expansion	(0.009)	(0.007)	(0.006)	(0.005)	(0.012)	(0.001)
Panel III: Drop ACA Early Expanders Version 2						
ACA without	0.037**	0.033***	0.006	0.027	0.006	-0.001
Medicaid Expansion	(0.012)	(0.009)	(0.006)	(0.014)	(0.005)	(0.001)
Medicaid Expansion	0.048***	-0.014	0.003	-0.017	0.062***	0.002
	(0.013)	(0.010)	(0.007)	(0.015)	(0.009)	(0.002)
ACA with Medicaid	0.085***	0.019***	0.009*	0.010*	0.067***	0.001
Expansion	(0.007)	(0.004)	(0.004)	(0.004)	(0.008)	(0.001)
Panel IV: 13 Treatment States and 16 Control States without a Medicaid Expansion Before 2014						
ACA without	0.038**	0.035**	0.007	0.028	0.005	-0.001
Medicaid Expansion	(0.014)	(0.010)	(0.005)	(0.015)	(0.005)	(0.001)
Medicaid Expansion	0.052**	-0.008	0.007	-0.021	0.068***	-0.003
	(0.017)	(0.014)	(0.009)	(0.016)	(0.019)	(0.002)
ACA with Medicaid	0.090***	0.027**	0.015	0.008	0.073***	-0.004*
Expansion	(0.010)	(0.010)	(0.007)	(0.008)	(0.018)	(0.001)
Panel V: Drop ACA Late Expanders						
ACA without	0.037**	0.033***	0.006	0.027	0.006	-0.001
Medicaid Expansion	(0.011)	(0.009)	(0.005)	(0.014)	(0.005)	(0.001)
Medicaid Expansion	0.056***	-0.008	0.010	-0.017	0.066***	0.001
	(0.013)	(0.011)	(0.008)	(0.015)	(0.008)	(0.001)
ACA with Medicaid	0.093***	0.025***	0.016**	0.009**	0.071***	0.000
Expansion	(0.006)	(0.006)	(0.006)	(0.003)	(0.008)	(0.001)
Panel VI: Drop All Early and Late Expanders						
ACA without	0.038*	0.035**	0.007	0.028	0.005	-0.001
Medicaid Expansion	(0.014)	(0.010)	(0.005)	(0.014)	(0.006)	(0.001)
Medicaid Expansion	0.057**	-0.017	-0.008	-0.013	0.086***	-0.006**
	(0.018)	(0.014)	(0.007)	(0.017)	(0.018)	(0.002)
ACA with Medicaid	0.095***	0.018	-0.001	0.015	0.091***	-0.007***
Expansion	(0.010)	(0.010)	(0.004)	(0.008)	(0.017)	(0.001)
Panel VII: Include Medicaid-Uninsurance Rate in DDD Analysis						

ACA without Medicaid Expansion	0.037**	0.033**	0.006	0.027	0.006	-0.001
	(0.012)	(0.009)	(0.005)	(0.014)	(0.005)	(0.001)
Medicaid Expansion	0.053***	-0.007	0.011	-0.018	0.062	0.001
	(0.013)	(0.011)	(0.008)	(0.015)	(0.008)	(0.001)
ACA with Medicaid Expansion	0.091***	0.026***	0.018	0.009	0.068	0.001
	(0.006)	(0.005)	(0.005)	(0.003)	(0.007)	(0.001)
Panel VIII: State Level Uninsurance Rate with additional State Controls						
ACA without Medicaid Expansion	0.054***	0.033***	0.017**	0.018*	0.022*	0.000
	(0.011)	(0.005)	(0.006)	(0.008)	(0.010)	(0.001)
Medicaid Expansion	0.038***	0.003	0.004	-0.003	0.037***	-0.000
	(0.010)	(0.003)	(0.003)	(0.004)	(0.010)	(0.001)
ACA with Medicaid Expansion	0.092***	0.035***	0.021***	0.015**	0.059***	0.000
	(0.010)	(0.006)	(0.006)	(0.006)	(0.010)	(0.001)

Notes: Standard errors, heteroscedasticity-robust and clustered by state, are in parentheses. *** indicates statistically significant at 0.1% level; ** 1% level; * 5% level. Sampling weights are used. All regressions include state*location type and year*location type fixed effects as well as the controls

References

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

Alam, A., Baez, J. E., & Carpio, X. V. (2011). Does cash for school influence young women's behavior in the longer term? evidence from Pakistan. World Bank Policy Research Working Paper Series.

Allen, D., & Dadgar, M. (2012). Does dual enrollment increase students’ success in college? Evidence from a quasi-experimental analysis of dual enrollment in New York City. *New Directions for Higher Education*, *2012*(158), 11-19.

Almond, D. (2006). Is the 1918 influenza pandemic over? Long-term effects of in utero influenza exposure in the post-1940 US population. *Journal of political Economy*, 114(4), 672-712.

Almond, D., & Mazumder, B. (2005). The 1918 influenza pandemic and subsequent health outcomes: an analysis of SIPP data. *American Economic Review*, 95(2), 258-262.

An, B. P. (2013). The impact of dual enrollment on college degree attainment: Do low-SES students benefit? *Educational Evaluation and Policy Analysis*, *35*(1), 57-75.

Angeles, G., Guilkey, D. K., & Mroz, T. A. (2005). The Effects of Education and Family Planning Programs on Fertility in Indonesia. *Economic Development and Cultural Change*, 165-201.

Angrist, J. D., & Keueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings?. *The Quarterly Journal of Economics*, 106(4), 979-1014.

Antecol, H., Eren, O., & Ozbeklik, S. (2015). The effect of teacher gender on student achievement in primary school. *Journal of Labor Economics*, *33*(1), 63-89.

Ashenfelter, O., & Krueger, A. (1994). Estimating the returns schooling using a nex sample of twins. *American Economic Review*, *84*(5), 1157-73.

Attanasio, O., Battistin, E., Fitzsimons, E., Mesnard, A., & Vera-Hernandez, M. (2005). How effective are conditional cash transfers? Evidence from Colombia. London: Institute for Fiscal Studies.

Avery, C., Gurantz, O., Hurwitz, M., & Smith, J. (2018). Shifting college majors in response to advanced placement exam scores. *Journal of Human Resources*, 53(4), 918-956.

Banks, J., & Smyth, E. (2015). ‘Your whole life depends on it’: academic stress and high-stakes testing in Ireland. *Journal of Youth Studies*, *18*(5), 598-616.

Barker, D. J., Bull, A. R., Osmond, C., & Simmonds, S. J. (1990). Fetal and placental size and risk of hypertension in adult life. *Bmj*, 301(6746), 259-262.

Baum, S., Ma, J., Bell, D. W., & Elliott, D. C. (2014). Trends in college pricing, 2014. Trends in Higher Education Series. *College Board*.

Becker, G. S., & Lewis, H. G. (1973). On the Interaction between the Quantity and Quality of Children. *Journal of Political Economy*, 279-288.

Behrman, J. R., & Hoddinott, J. (2005). Program evaluation with unobserved heterogeneity and selective implementation: The Mexican Progresa impact on child nutrition. *Oxford Bulletin of Economics and Statistics*, 547–569.

Berchick, Edward R., Jessica C. Barnett, and Rachel D. Upton. 2019. Current Population Reports, P60-267(RV), *Health Insurance Coverage in the United States: 2018,* U.S. Government Printing Office, Washington, DC.

Bettinger, E. P., & Long, B. T. (2005). Do faculty serve as role models? The impact of instructor gender on female students. *American Economic Review*, *95*(2), 152-157.

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *The Quarterly Journal of Economics*, *127*(3), 1205-1242.

Black, S. E., Devereux, P. J., & Salvanes, K. G. (2007). From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, *122*(1), 409-439.

Blau, F. D., & Kahn, L. M. (2000). Gender differences in pay. *Journal of Economic perspectives*, *14*(4), 75-99.

Bound, J., Lovenheim, M. F., & Turner, S. (2010). Why have college completion rates declined? An analysis of changing student preparation and collegiate resources. *American Economic Journal: Applied Economics*, *2*(3), 129-57.

Braakmann, N. (2011). The causal relationship between education, health and health related behaviour: Evidence from a natural experiment in England. *Journal of Health Economics*, 753-763.

Breierova, L., & Duflo, E. (2004). The impact of education on fertility and child mortality: Do fathers really matter less than mothers? (No. w10513). National bureau of economic research.

Buchmueller, Thomas C., Zachary M. Levinson, Helen G. Levy, and Barbara L. Wolfe. 2016. Effect of the Affordable Care Act on racial and ethnic disparities in health insurance coverage. *American Journal of Public Health* 106: 1416-1421.

Buser, T., Niederle, M., & Oosterbeek, H. (2014). Gender, competitiveness, and career choices. *The Quarterly Journal of Economics*, *129*(3), 1409-1447.

Canes, B. J., & Rosen, H. S. (1995). Following in her footsteps? Faculty gender composition and women's choices of college majors. *ILR Review*, *48*(3), 486-504.

Carrell, S. E., Page, M. E., & West, J. E. (2010). Sex and science: How professor gender perpetuates the gender gap. *The Quarterly Journal of Economics*, *125*(3), 1101-1144.

Chaudhury, N., & Parajuli, D. (2006). Conditional cash transfers and female schooling: the impact of the female school stipend program on public school enrollments in Punjab, Pakistan. Washington DC: World Bank.

Chou, S.-Y., Liu, J.-T., Grossman, M., & Joyce, T. (2010). Parental Education and Child Health: Evidence from a Natural Experiment in Taiwan. *American Economic Journal: Applied Economics*, 33–61.

Clark, D., & Royer, H. (2013). The Effect of Education on Adult Mortality and Health: Evidence from Britain. *The American Economic Review*, 2087-2120.

Collins Sara R, Munira Z. Gunja, Michelle M. Doty, Herman K. Bhupal. 2018. First Look at Health Insurance Coverage in 2018 Finds ACA Gains Beginning to Reverse. New York, NY: Commonwealth Fund. https://www.commonwealthfund.org/blog/2018/first-look-health-insurance-coverage-2018-finds-aca-gains-beginning-reverse. Accessed April 19. 2020.

Commonwealth Fund. 2018. The Affordable Care Act Under the Trump Administration. https://www.commonwealthfund.org/blog/2018/affordable-care-act-under-trump-administration . Accessed April 13, 2020.

Conger, D., Kennedy, A. L., Long, M. C., & McGhee, R. (2019). The effect of advanced placement science on students' skills, confidence and stress. *Journal of Human Resources*, 0118-9298R3.

Courtemanche, C. J., & Zapata, D. (2014). Does universal coverage improve health? The Massachusetts experience. *Journal of Policy Analysis and Management*, *33*(1), 36-69.

Courtemanche, Charles J., Andrew Friedson, Andrew Koller, and Daniel I. Rees, 2019c. The Affordable Care Act and ambulance response times. *Journal of Health Economics* 67: article 102213.

Courtemanche, Charles, Ishtiaque Fazlul, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. (2020a). The Impact of the Affordable Care Act on Disparities in Insurance Coverage After Four Years, in Medicaid: Enrollment, Eligibility, and Key Issues, D. Lanford (editor), Hauppauge, NY: Nova Science Publishers.

Courtemanche, Charles, James Marton, and Aaron Yelowitz. 2016. Who gained insurance coverage in 2014, the first year of full ACA implementation? *Health Economics* 25: 778–84.

Courtemanche, Charles, James Marton, and Aaron Yelowitz. 2020a. Medicaid Coverage Across the Income Distribution under the Affordable Care Act, in Medicaid: Enrollment, Eligibility, and Key Issues, D. Lanford (editor), Hauppauge, NY: Nova Science Publishers.

Courtemanche, Charles, James Marton, and Aaron Yelowitz. 2020b. The Full Impact of the ACA on Political Participation. Russell Sage Foundation Journal of the Social Sciences, forthcoming.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. 2017. 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): 178–210.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. 2018a. Early effects of the Affordable Care Act on health care access, risky health behaviors, and self-assessed health. *Southern Economic Journal* 84(3): 660–691.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. 2018b. Effects of the Affordable Care Act on health care access and self-assessed health after 3 years. *Inquiry* 55: 1–10.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. 2019a. Effects of the Affordable Care Act on health behaviors after 3 years. *Eastern Economic Journal* 45: 7-33.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, and Daniela Zapata. (2020b). The Impact of the ACA on Health Care Access and Self-Assessed Health in the Trump Era (2017-2018). *Health Services Research,* forthcoming.

Courtemanche, Charles, James Marton, Benjamin Ukert, Aaron Yelowitz, Daniela Zapata, and Ishtiaque Fazlul. 2019b. The three-year impact of the Affordable Care Act on disparities in insurance coverage. *Health Services Research* 54: 307-316.

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

Currie, J., & Moretti, E. (2007). Biology as destiny? Short-and long-run determinants of intergenerational transmission of birth weight. *Journal of Labor economics*, 25(2), 231-264.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. *American Economic Review*, *95*(2), 158-165.

Dee, T. S. (2007). Teachers and the gender gaps in student achievement. *Journal of Human Resources*, *42*(3), 528-554.

Denning, J., Eide, E., Mumford, K., Patterson, R., and Warnick, M. (2019). Why have college completion rates increased? *IZA Discussion Paper No 12411*.

Dillon, E. W., & Smith, J. A. (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics*, *35*(1), 45-66.

Dillon, E. W., & Smith, J. A. (2017). Determinants of the match between student ability and college quality. *Journal of Labor Economics*, *35*(1), 45-66.

Dougherty, C., Mellor, L., & Jian, S. (2006). The relationship between advanced placement and college graduation. 2005 AP Study Series, Report 1. *National Center for Educational Accountability*.

East, C. N., Miller, S., Page, M., & Wherry, L. R. (2017). *Multi-generational impacts of childhood access to the safety net: Early life exposure to Medicaid and the next generation's health* (No. w23810). National Bureau of Economic Research.

Ehrenberg, R. G., Goldhaber, D. D., & Brewer, D. J. (1995). Do teachers' race, gender, and ethnicity matter? Evidence from the National Educational Longitudinal Study of 1988. *ILR Review*, *48*(3), 547-561.

Frean, Molly, Jonathan Gruber, and Benjamin D. Sommers. 2017. Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act. Journal of Health Economics 53: 72-86.

Fuwa, N. (2001). The Net Impact of the Female Secondary School Stipend Program in Bangladesh. Munich Personal RePEc Archive.

Gertler, P. (2004). Do Conditional Cash Transfers Improve Child Health? Evidence from PROGRESA's Control Randomized Experiment. *The American Economic Review*, 336-341.

Gong, J., Lu, Y., & Song, H. (2018). The effect of teacher gender on students' academic and noncognitive outcomes. *Journal of Labor Economics*, *36*(3), 743-778.

Goodman, S. (2016). Learning from the test: Raising selective college enrollment by providing information. *Review of Economics and Statistics*, *98*(4), 671-684.

Grépin, K. A., & Bharadwaj, P. (2015). Maternal education and child mortality in Zimbabwe. *Journal of Health Economics.*, 97-117.

Griffith, Kevin N. and Jacob H. Bor. 2020. Changes in health care access, behaviors, and self-reported health among low-income US Adults through the fourth year of the Affordable Care Act. Medical Care 58(6): 574-578.

Griffith, Kevin N., David K. Jones, Jacob H. Bor, and Benjamin D. Sommers. 2020. Changes in health insurance coverage, access to care, and income-based disparities among US adults 2011-2017. Health Affairs 39(2): 319–26.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 223-255.

Gruber, Jonathan and Benjamin D. Sommers, 2019. The Affordable Care Act's effects on patients, providers, and the economy: What we've learned so far. *Journal of Policy Analysis and Management* 38(4): 1028-1052.

Grytten, J., Skau, I., & Sørensen, R. J. (2014). Educated mothers, healthy infants. The impact of a school reform on the birth weight of Norwegian infants 1967–2005. *Social science & medicine*, 105, 84-92.

Güneş, P. M. (2015). The role of maternal education in child health: Evidence from a compulsory schooling law. *Economics of Education Review*, *47*, 1-16.

Gurantz, O. (2019). How college credit in high school impacts postsecondary course-taking: the role of AP exams. *Education Finance and Policy*, 1-43.

Hahn, Y., Islam, A., Nuzhat, K., Smyth, R., & Yang, H.-S. (2018). Education, Marriage and Fertility: Long-Term Evidence from a Female Stipend Program in Bangladesh. *Economic Development and Cultural Change*, 383-415.

Heim, Bradley, Ithai Z. Lurie, Daniel W. Sacks. 2018. Does the individual mandate affect insurance coverage? Evidence from the population of tax returns.

Hinde, Jesse M. 2017. Incentive(less)? The effectiveness of tax credits and cost-sharing subsidies in the Affordable Care Act. *American Journal of Health Economics* 3(3): 346–369.

Hoffmann, F., & Oreopoulos, P. (2009). A professor like me the influence of instructor gender on college achievement. *Journal of Human Resources*, *44*(2), 479-494.

Holt, A., & McGarrity, J. (2018). Overcoming inertia with a nudge: How the AAIMS program increased Advanced Placement participation in Arkansas. *Studies in Business and Economics*, *13*(1), 67-75.

Hong, S. Y., & Sarr, L. R. (2012). Long-term impacts of the free tuition and female stipend programs on education attainment, age of marriage, and married women's labor market participation of in Bangladesh. Retrieved from ImageBank, the World Bank: http://imagebank. worldbank. org/servlet/WDSContentServer/IW3P/IB/2013/09/16/000442464_20130916134309/Rend e red/PDF/810610WP0P10620Box0379826B00PUBLIC0. pdf.

Hoxby, C., & Turner, S. (2013). Expanding college opportunities for high-achieving, low income students. *Stanford Institute for Economic Policy Research Discussion Paper*, *12*, 014.

Hurwitz, M., Smith, J., Niu, S., & Howell, J. (2015). The Maine question: How is 4-year college enrollment affected by mandatory college entrance exams?. *Educational Evaluation and Policy Analysis*, *37*(1), 138-159.

Hyman, J. (2017). ACT for all: The effect of mandatory college entrance exams on postsecondary attainment and choice. *Education Finance and Policy*, *12*(3), 281-311.

Jackson, C. K. (2010). A little now for a lot later a look at a texas advanced placement incentive program. *Journal of Human Resources*, *45*(3), 591-639.

Kaestner, Robert, Bowen Garrett, J. Chen, Anui Gangopadhyaya, and Caitlyn Fleming. 2017. Effects of the ACA Medicaid expansion on health insurance coverage and labor supply. *Journal of Policy Analysis and Management* 36(3): 608–42.

Kaiser Family Foundation [KFF] State Decisions on Health Insurance Marketplaces and the Medicaid Expansion. Available at http://kff.org/health-reform/state-indicator/state-decisions-for-creating-health-insurance-exchanges-and-expanding-medicaid/ . Accessed April 19, 2020.

Khandker, S. R., Pitt, M. M., & Fuwa, N. (2003). Subsidy to Promote Girls' Secondary Education: The Female Stipend Program in Bangladesh. Munich Personal RePEc Archive.

Klasik, D. (2013). The ACT of enrollment: The college enrollment effects of state-required college entrance exam testing. *Educational researcher*, *42*(3), 151-160.

Klopfenstein, K. (2004). Advanced Placement: Do minorities have equal opportunity?. *Economics of Education Review*, *23*(2), 115-131.

Kowalski, Amanda. 2014. The early impact of the Affordable Care Act state-by-state. NBER Working Paper 20597.

Lindeboom, M., Llena-Nozal, A., & van Der Klaauw, B. (2009). Parental education and child health: Evidence from a schooling reform. *Journal of health Economics*, *28*(1), 109-131.

Lleras-Muney, A. (2005). The Relationship Between Education and Adult Mortality in the United States. *The Review of Economic Studies*, 189-221.

Long, B. T., & Osili, U. O. (2008). Does female schooling reduce fertility? Evidence from Nigeria. *Journal of Development Economics*, 57-75.

McCrary, J., & Royer, H. (2011). The effect of female education on fertility and infant health: Evidence from school entry policies using exact date of birth. *American economic review*, *101*(1), 158-95.

McMorrow, Stacey, Genevieve M. Kenney, Sharon K. Long, and Jason A. Gates. 2016. Marketplaces helped drive coverage gains in 2015; Affordability problems remained. *Health Affairs* 35(10): 1810–15.

Miller, P., Mulvey, C., & Martin, N. (1995). What do twins studies reveal about the economic returns to education? A comparison of Australian and US findings. *The American Economic Review*, 85(3), 586-599.

Miller, Sarah 2012. The effect of insurance on emergency room visits: An analysis of the 2006 Massachusetts health reform. *Journal of Public Economics* 96: 893–908.

Miller, Sarah and Laura R. Wherry. 2017. Health and access to care during the first 2 years of the ACA Medicaid expansions. *New England Journal of Medicine* 376: 947-956.

Morgan, R., & Klaric, J. (2007). AP students in college: An analysis of five-year academic Careers. Research Report No. 2007-4. *College Board*.

Morris, S., Olinto, P., Flores, R., Nilson, E. A., & Figueiró, A. C. (2004). Conditional cash transfers are associated with a small reduction in the rate of weight gain of preschool children in northeast Brazil. *Journal of Nutrition*, 2336–41.

Neumark, D., & Gardecki, R. (1996). Women helping women? Role-model and mentoring effects on female Ph. D. student in economics. *National Bureau of Economic Research*, No. w5733.

Nixon, L. A., & Robinson, M. D. (1999). The educational attainment of young women: Role model effects of female high school faculty. *Demography*, *36*(2), 185-194.

Obama, Barack. 2016. United States health reform: Progress to date and next steps. *Journal of the American Medical Association* 316(5): 525–32.

Oreopoulos, P. (2006). Estimating Average and Local Average Treatment Effects of Education when Compulsory Schooling Laws Really Matter. *American Economics Review*, 152-175.

Page, L. C., & Scott-Clayton, J. (2016). Improving college access in the United States: Barriers and policy responses. *Economics of Education Review*, *51*, 4-22.

Patterson, B. F., & Ewing, M. (2013). Validating the use of AP exam scores for college course placement. Research Report 2013-2. *College Board*.

Pope, D. G., & Fillmore, I. (2015). The impact of time between cognitive tasks on performance: Evidence from advanced placement exams. *Economics of Education Review*, 48, 30-40.

Richard Marshall, & Rahman, S. (2013). Internal Migration in Bangladesh: Character, Drivers and Policy Issues. Bangladesh: United Nations Development Program.

Rivera, J., Sotres-Alvarez, D., Habicht, J.-P., & Villalpando, S. (2004). Impact of the Mexican program for education, health, and nutrition (Progresa) on rates of growth and anemia in infants and young children: a randomized effectiveness study. *Journal of the American Medical Association*, 2563–70.

Rothstein, D. S. (1995). Do female faculty influence female students' educational and labor market attainments?. *ILR Review*, *48*(3), 515-530.

Royer, H., & McCrary, J. (2006). *The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth*. National Bureau of Economic Research.

Royer, Heather. 2009. "Separated at Girth: US Twin Estimates of the Effects of Birth Weight." *American Economic Journal: Applied Economics*, 1 (1): 49-85.

Sayeed, Y. (2016). *Effect of girls' secondary school stipend on completed schooling, age at marriage, and age at first birth: Evidence from Bangladesh* (No. 2016/110). WIDER Working Paper.

Schurmann, A. T. (2009). Review of the Bangladesh Female Secondary School Stipend Project Using a Social Exclusion Framework. Journal of Health, *Population and Nutrition*, 505–517.

Sen, A. (1982). *Poverty and famines: an essay on entitlement and deprivation*. Oxford university press.

Shamsuddin, M. (2015). Labour market effects of a female stipend programme in Bangladesh. *Oxford Development Studies, 43*(4), 425-447.

Smith, J., Hurwitz, M., & Avery, C. (2017). Giving college credit where it is due: Advanced Placement exam scores and college outcomes. *Journal of Labor Economics*, *35*(1), 67-147.

Smith, J., Hurwitz, M., & Howell, J. (2015). Screening mechanisms and student responses in the college market. *Economics of Education Review*, *44*, 17-28.

Solorzano, D. G., & Ornelas, A. (2002). A critical race analysis of advanced placement classes: A case of educational inequality. *Journal of Latinos and Education*, *1*(4), 215-229.

Sommers, Benjamin D., Munnira Z. Gunja, Kenneth Finegold, and Thomas Musco. 2015. Changes in self-reported insurance coverage, access to care, and health under the Affordable Care Act. *Journal of the American Medical Association* 314(4): 366–74.

Speer, J. D. (2017). The gender gap in college major: Revisiting the role of pre-college factors. *Labour Economics*, *44*, 69-88.

United States Department of Health and Human Services [HHS]. 2017. *Trump Administration Takes Action to Abide by the Law and Constitution, Discontinue CSR Payments.* https://www.hhs.gov/about/news/2017/10/12/trump-administration-takes-action-abide-law-constitution-discontinue-csr-payments.html . Accessed April 13, 2020.

United States Government Accounting Office [GAO]. 2018. *Health Insurance Exchanges: HHS Should Enhance Its Management of Open Enrollment Performance.* https://www.gao.gov/assets/700/693362.pdf . Accessed April 13, 2020.

Wherry, Laura R., and Sarah Miller. 2016. Early coverage, access, utilization, and health effects associated with the Affordable Care Act Medicaid expansions: A quasi-experimental study. *Annals of Internal Medicine* 164(12): 795–803.

White House. 2017. *Federal Register Volume 82, Number 14, Executive Order 13765.* https://www.govinfo.gov/content/pkg/FR-2017-01-24/pdf/2017-01799.pdf . Accessed April 13, 2020.

WHO, World Bank, & Ministry of Health and Family Welfare Bangladesh. (2015). Success Factors for Women's and Children's Health: Bangladesh. World Bank.

Winters, M. A., Haight, R. C., Swaim, T. T., & Pickering, K. A. (2013). The effect of same-gender teacher assignment on student achievement in the elementary and secondary grades: Evidence from panel data. *Economics of Education Review*, *34*, 69-75.

World Health Organization. (2017). World health statistics 2017: monitoring health for the SDGs, Sustainable Development Goals. Geneva: World Health Organization.

Vita

Ishtiaque Fazlul was born and raised in a Mymensingh, Bangladesh. He received his Bachelor's degree in Economics in 2011 and his Master's degree in Economics in 2012 from the University of Dhaka. Ishtiaque worked as a research associate for Innovations for Poverty Action (IPA) for three years till the summer of 2015. At IPA, he worked on a number of randomized controlled trial studies evaluating various interventions for developing ready-made garments industry workforce in Bangladesh.

Ishtiaque started his doctoral studies in economics at Georgia State University in the fall of 2015. His research interests include Education Economics and Health Economics. He is the recipient of the Andrew Young School Dean's Fellowship and the Center for Economic Analysis of Risk (CEAR) Fellowship at Georgia State University.

Ishtiaque will join the University of Missouri as a postdoctoral research associate from the fall of 2020.