THREE ESSAYS IN APPLIED ECONOMICS

By

Quinton James Baker

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics – Doctor of Philosophy

2020

ABSTRACT

THREE ESSAYS IN APPLIED ECONOMICS

By

Quinton James Baker

My dissertation covers areas in applied economics: labor economics, public economics, and the economics of education. I use a variety of econometric tools and other economic analysis to study welfare program rules and regulations as well as assess the efficacy of a high school science curriculum. My first chapter uses data from the Survey of Income and Program Participation (SIPP), spanning 2003-2013, to estimate whether the ABAWD-specific 20 hour per week minimum work requirement influences their labor supply outcomes and SNAP participation. I employ binary response models to estimate average partial effects (APE) and find the work requirement has statistically significant effects: ABAWDs are 1 percentage point (pp) less likely to participate in SNAP and are 2.6 pp more likely to meet the 20-hour work requirement. This negative effect of the work requirement on SNAP participation is larger among non-white (1.37 pp), specifically blacks (2.09 pp), suggesting that the impacts of a work requirement must be considered in areas with higher percentages of minorities. This paper contributes to the study of ABAWDs, a relatively understudied population in the context of SNAP.

Chapter 2 uses the 2014 Survey of Income and Program Participation (SIPP), spanning from January 2013 to December 2016, to study the effect of SNAP and Medicaid expansion on labor market outcomes (income, hours worked and employment status) and SNAP participation. Using a suite of empirical methods, I find no evidence that the interaction of the SNAP and Medicaid expansions has an effect on labor outcomes of the head and second adult in a household. However, I do find that the Medicaid expansion increases SNAP participation in states with the least generous state-level SNAP policy options. These findings demonstrate the importance of analyzing the effect of both expansions jointly, as both SNAP

and Medicaid serve low income households that may simultaneously choose their labor supply and program participation.

Chapter 3 uses data from the NSF funded project Crafting Engaging Science Environments (CESE), a cluster-randomized controlled trial to study the effect of project-based learning on the scientific achievement of high school chemistry and physics students in Michigan and California during the 2018-2019 school year. I extend the analysis conducted in Schneider et al. (2021) and use pooled OLS with school level fixed effects to estimate the treatment effect. I find sound evidence to support the findings in Schneider et al., 2021 that the CESE intervention had a positive and significant effect on students' scientific learning, even in the presence of multiple levels of attrition. The point estimates range from 0.24 to 0.34 standard deviations. Additionally, I compute the Lee Bounds for the estimates and find the bounds do not contain zero, suggesting that differential attrition alone likely does not drive the entire treatment effect.

For my wife, Samantha Lucía Padilla.
I could not have done it without your loving support all these years.
For my daughter, Eleonora Lila.
Submitting this dissertation was the second best thing to happen this year.
The first was becoming your papá.
I love you both dearly.

ACKNOWLEDGEMENTS

I am incredibly grateful and indebted to many people and groups that helped shape me into the person I am today.

Sincerest thanks to many at the Michigan State University Department of Economics for the opportunity to study. In particular: Dr. Leslie Papke, she was the first person I spoke to about admission to MSU econ, and the last to sign-off on my dissertation as committee chair. Dr. Stacy Dickert-Conlin, who taught a wonderful course on public expenditures and helped me greatly with my first chapter of this thesis. Dr. Todd Elder, for telling it like it is and being willing to help. Dr. Steven Haider, for the numerous meetings in which he encouraged me to continue.

Deepest gratitude to Dr. Barbara Schneider, for hiring me to work on two fantastic projects managing and analyzing data on the economics of education. This opportunity was the air beneath my wings.

I also thank the Department of Agricultural, Food, and Resource Economics. First, for using the same preliminary exams as the Economics department as it is how I met my wife. Second, thank you professors Dr. James Hilker and Dr. Dave Weatherspoon for hiring me to assist in their courses; both ended up giving me great program advice and introducing me to fascinating subjects along the way. My placement at the U.S. Department of Agriculture, Economic Research Service, is in part because of them. Lastly two friends whom are also the godparents of my firstborn child: Dr. Sophia Tanner, for providing feedback on my job market paper, for listening to my presentation, and for her friendship over the years. Dr. Stephen Morgan for providing assistance directly and via proxy through my wife. I look forward to being your coworker!

Great thanks to many fine people I have met at the University of Wisconsin-Whitewater. In particular: Dr. David Welsch, for the mentorship and guidance to graduate school. Dr. Yamin Ahmad for making macroeconomics a great pleasure to study. Dr. Jaeger Nelson, for pushing my love for economics, helping me understand and apply for graduate school, and for being there for me as a friend as I struggled through the PhD program.

Special thanks to the village of Amherst, Wisconsin and the fine teachers at the Tomorrow River School District. There are too many great people to name from this small school in a small town that helped me become who I am today.

I am eternally grateful to my parents, Patrick and Katherine Baker. They made many sacrifices and worked very hard to provide me the opportunity to choose a life for my own. Apparently that path ended up with me earning a PhD in economics from a highly respected university.

Lastly, though most importantly, I thank my wife Dr. Samantha Lucia Padilla. For the majority of my PhD program, Samantha has had an immense amount of patience, support, and love for me. She also got her loving family to care for me and pray for me whenever I needed it to get me through the next milestone. We even had a beautiful daughter, Eleonora Lila, whose smiles and coos gave me a renewed sense of importance of the degree. I cannot say it enough times, I could not have done it without you, Samantha. Luckily for me, we have vowed to spend the rest of our lives together. May this time frame allow me enough opportunities to express my gratitude.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

ABLE-BODIED ADULTS WITHOUT DEPENDENTS: PARTICIPATION, LABOR SUPPLY, AND THE SNAP TIME LIMIT

1.1 Introduction

During the Great Recession, the caseloads in the Supplemental Nutrition Assistance Program (SNAP, formerly the Food Stamp Program) increased from 26.5 million households per month in 2007 to 44.7 million households per month in 2011 (Wolkwitz, 2007; Strayer, Eslami, and Leftin, 2012; Ganong and Liebman, 2018). The sub-population of able-bodied adults without dependents (ABAWDs) grew substantially during this period, accounting for 4 percent of the caseload before the recession to more than 10% from 2011 through 2014, as shown in Figure 1.1 (Strayer, Eslami, and Leftin, 2012; K. F. Gray, 2014; K. F. Gray and Eslami, 2014). An ABAWD is defined as an individual (1) aged between 18 to 49 years old, (2) with neither children nor other dependents, and (3) with no medically-certified physical or mental disabilities (USDA, 2015).

ABAWDs are the only SNAP participants subject to a time limit on the receipt of benefits. Through the Welfare Reform Act of 1996, an ABAWD can only receive three months of SNAP benefits within a 36-month period, unless they engage in work activities for at least 20 hours per week (Falk, McCarty, and Aussenberg, 2014). This policy is known as the ABAWD time limit and is the focus of this paper.

The relative growth in ABAWD caseload during the Great Recession may be due to changes in program eligibility requirements made during the financial crisis. Specifically, the American Recovery and Reinvestment Act of 2009 (ARRA), for the first time since the welfare reform of 1996, waived the ABAWD time limit nationwide and increased maximum benefit amounts.[1]

[1] Under ARRA, ABAWDs must still meet the general SNAP requirement. This entails actively looking for employment.

In early 2016, several years into post-recession growth, 31 states and territories were receiving state-wide time limit waivers, with 12 more states receiving waivers for specific regions inside the state. Although the waivers of many states have expired as of 2017, the ABAWDs' share of SNAP caseload (7.3%) and benefit amount (10.6%) remained double the pre-recession levels (Cronquist, 2019). The drastic increase and gradual decline of ABAWDs participation in SNAP motivate the questions of interest in this work: Does the time limit effectively nudge ABAWDs into the labor market? And, does the time limit discourage program participation?

In this paper, I use the variations in time limit waivers by state from pre- and post-ARRA to study the effect of the the 20 hour per week work requirement on labor market outcomes using data from the 2001, 2004, and 2008 Survey of Income and Program Participation (SIPP) panels, which span from October 2003 to December 2013. I use a logit model[2] to estimate whether the time limit influences two main outcomes of interest: (1) SNAP program participation and (2) whether participants met the 20-hour work requirement. I compute the average partial effects of the policy, that is, the average effectiveness of the ABAWD time limit to alter behavior. In addition, I examine the intensive margin of the labor supply by analyzing the effect of the time limit on the reported number of hours worked using OLS estimation with state and time dummies.

The additional work requirement may have important considerations for this population, as ABAWDs are generally not eligible for other welfare programs. That is, SNAP may be their only safety net in the US welfare system. By virtue of being able-bodied, ABAWDs do not qualify for programs such as Disability Insurance or Supplemental Security Income (SSI). By being adults, defined in this context as being aged 18-49 years, they do not qualify for programs aimed at serving the elderly, and by not having dependent[3] children, they do not qualify for any programs designed to promote the welfare of children. Hence, SNAP is an essential element of the safety net for this population and changes to participation

[2]Probit results are also available in Appendix A.

[3]Dependent in this context is unrelated to a dependent for tax purposes.

requirements affect ABAWDs food security and health outcomes (Rosenbaum, 2013). In addition, emphasis on this population provides for a study of the influence of welfare programs on labor supply, whereas this is generally difficult to estimate because many of participants who are eligible for one particular program are often eligible for multiple other programs (Czajka et al., 2001; Fraker and R. Moffitt, 1989).

I find that the ABAWDs facing the time limit are 1 pp less likely to participate in SNAP and are 2.6% more likely to meet the 20-hour work requirement than ABAWDs who are living in regions where the time limit is waived. These effects are larger for non-white heads of household (1.37% and 2.77%, respectively) and even larger when examining only black heads of household (2.1% and 3.2%). These results are similar in sign to recent findings by Harris (2018) and C. Gray et al., 2019. Harris (2018) applies difference-in-difference estimation on SNAP Quality Control (QC) data (2010-2016) and finds the time limit decreases the ABAWD's SNAP participation by 10%. Similarly, using regression discontinuity design and administrative data from the State of Virginia, C. Gray et al. (2019) find the work requirement significantly reduces SNAP participation. Both of these studies use the aging out of ABAWD status for identification and consistently demonstrate the negative effect of an additional work requirement on program participation.

This paper contributes to the labor and SNAP-ABAWD literature by using survey based data (SIPP) and a model to examine the labor supply effects of the ABAWD time limit. The benefit of using the SIPP is that it asks about income and program participation on a monthly basis, which is the same time interval as SNAP administration. The use of survey data in general, as opposed to administrative data such as the SNAP Quality Control Sample, is important because those data sets do not contain information about people who could potentially participate in SNAP. This is because SNAP administrators do not have cases on file from non-applicants. Administrative data can also fail to identify which participants could potentially become subject to the ABAWD time limits (Czajka et al., 2001). A second contribution of this paper is it focuses on the entire United States, as opposed to case studies centered on a single state or county and does not rely on estimating local average treatment effects from ABAWDs aging out of ABAWD status (C. Gray et al., 2019; Hahn et al., 2019; Shaw and Hooker, 2016).

This paper is organized as follows. Section 2 discusses SNAP and the specific provisions for ABAWDs. Section 3 contains a review of existing studies and positions my work within the literature. Sections 4 and 5 describe the SIPP data used and models employed. Sections 6 and 7 present the main results from the empirical estimations and the discussion. Section 8 concludes.

1.2 Supplemental Nutrition Assistance Program Details

1.2.1 Brief History of the ABAWD Time Limit

In 1996, President Bill Clinton signed the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA), a major welfare reform into law. The act established additional restrictions to participate in welfare programs and the new restrictions were less strict on families with children and the elderly. For Food Stamp Program participants, PRWORA imposed a general work requirement that to receive benefits, individuals must be actively searching for a job, not voluntarily leaving work and/or reducing work hours. In addition to this general requirement, ABAWDs had an extra restriction imposed on them, requiring that they work at least 80 hours per month, or roughly 20 hours per week in order to receive more than three months of benefits in a 36-month period.

The Balanced Budget Act of 1997 grants time limit waivers to states equal to 15 percent of their total ABAWD caseload. Through these waivers, SNAP administrators can give SNAP benefits to ABAWDs who have exhausted eligibility, i.e. are not meeting the work requirements and have received benefits for three months. If a state receives a time limit waiver, such as the waiver received through ARRA, the state does not receive any 15 percent waivers. The states get to determine how the waivers are used. When states overuse these 15 percent waivers, they are liable to refund the federal government for the benefits administered

(Foley, 2007). In practice, however, the states with a current negative balance of waivers will simply not be allowed to grant more 15 percent waivers until they their balance is positive.

The American Recovery and Reinvestment Act of 2009, in addition to increasing the maximum food stamp benefit amounts, granted nation-wide time limit waivers from April 2009 to October 2010. This provides a unique time to examine the participation rates of ABAWDs nationwide, as there was no uncertainty with respect to the individual's time limit. The waivers granted through ARRA and statewide waivers issued before and after the Great Recession are the focus of this paper.

1.2.2 ABAWD Time Limit Waivers

As the policies currently stand, every ABAWD is entitled to receive benefits for three full months out of a 36-month period, conditional on meeting the general SNAP participation requirements. If the ABAWD does not meet the ABAWD work requirement, then the ABAWD can only participate in SNAP in one of the following ways:

1. By living in a region receiving a time limit waiver- a region deemed to have "insufficient jobs."

2. By participating in a state-sponsored training program as opposed to, working for 20 hours per week.

3. By receiving a "15 percent" waiver from their state's allocation.

The definition of "insufficient jobs" generally means that there is a recent unemployment rate in the area either 10 percentage points higher or more than 20% higher than the national unemployment rate in the previous 3- or 12- month period. In subsequent legislation and U.S. Department of Agriculture Food and Nutrition Service (FNS) policies, various other measures of insufficient jobs have been added, such as declining employment-to-population ratio and Department of Labor designated "Labor Surplus Area." The additional work requirement for ABAWDs has been made more flexible since the late 1990s by allowing for broader definitions of what constitutes insufficient jobs, as well as states increasing funding training programs for ABAWDs to participate in lieu of work to satisfy the 20 hours per week requirement. States also have the discretion to allow certain kinds of volunteer work to count toward the work requirement.

Time limit waivers are generally applied for a 12-month period, with new two-year waivers available for areas with particularly poor employment conditions. For a 12- to 24- month period after receiving a waiver, the ABAWD work requirement does not apply in that waived region. Waivers are typically effective in the same month as their approval. Most of the other waivers effective in a month month different than their approval month were extensions of previous waivers or additions to current waivers. The timing of when these waivers is important as SNAP administrators need to know whether the waivers will be available during any SNAP re-certification meetings. In this paper, I examine these waivers as they, unlike the 15 percent waivers, are completely determined by geography and are certain to any knowledgeable participants.

One possible shortcoming of the insufficient job criteria is that it depends only on previous employment conditions and is not forward-looking. This has consequences for regions experiencing a large employment shock such as a major employer closing in a county. The state must first wait to qualify for a waiver, then they must apply for one. Upon receipt, the waiver would expire 15 months after the initial unemployment shock, making the waivers not available at the time where they could have been most useful. Exceptions to the lags include shocks such as Hurricane Katrina in 2005, where waivers were granted to Louisiana before the unemployment data became available (Bolen and Dean, 2018).

The eligibility consequences for the ABAWDs after the waiver expires are also ambiguous. These waivers prevent states from collecting more 15 percent waivers for the ABAWDs living in these regions, so there are fewer 15 percent waivers available than there may have been otherwise. However, participation in SNAP during a time limit waiver does not count against the three months out of three years policy. Thus, for a one year waiver, ABAWDs can regain universal eligibility with previous usages falling off of their history.

1.2.3 Budget Constraints and SNAP Eligibility

There are three tests that households must satisfy to receive food stamp benefits: gross income, net income, and assets. Both income tests refer to the federal poverty levels (FPL) from the previous fiscal year, with the thresholds being 130% and 100% respectively for the gross and net income tests.

Generally, the gross income standard includes all forms of income aside from small, infrequent receipts such as tax returns or bonuses. States use broad based categorical eligibility (often referred to as BBCE), a relatively recent state policy option, to increase the gross income limit from 130% to as high as 200% of the FPL.

The net income test is more complicated, as deductions for earned income, child support, dependent care, excess housing[4] and medical deductions[5] (HC, MC), in addition to a standard deduction (SD) are involved. To determine program eligibility, net income (Net) is determined by Adjusted Gross Income (AGI), net of child support (CS), medical deductions (MC), and excess housing (HC):

$$Net = AGI - CS - MC - HC \tag{1.1}$$

Adjusted Gross Income (AGI) is calculated as gross income[6] (GI) less 20% of earned income (EI), the standard deduction (SD), and dependent care (DC):

$$\begin{aligned} AGI &= GI - 0.2EI - SD - DC \\ &= NonEI + 0.8EI - SD - DC \end{aligned} \tag{1.2}$$

The SNAP benefit amount (B^*) is the maximum amount for a given household size, less 30 percent of net income. That is, from the disposable income left after deductions, participants

[4] Housing cost deduction is for the difference between actual housing costs and half of the Adjusted Gross Income, subject to a maximum amount. I omit this in the budget constraint discussion, but the housing cost deduction will increase the phase-out rate of food stamp benefits with respect to income.

[5] Medical deductions are for elderly only and applicable to spending beyond $35 per month.

[6] Gross income is the sum of earned and non-earned income, referred to as $NonEI$ in Equation 1.2 and 1.3

are expected to contribute 30 percent of their income to food expenditures while participating in the program. This amount is then subject to established maximum and minimum thresholds:

$$\begin{aligned} FSB^* &= MaxFSB - 0.3Net \\ &= MaxFSB - 0.3[NonEI - Deduc] - 0.24EI \end{aligned} \tag{1.3}$$

$$FSB = \max[MinFSB, \min[FSB^*, MaxFSB]] \tag{1.4}$$

The maximum food stamp benefit amount $MaxB$ is determined by the cost of the Thrifty Food Plan[7], the most affordable of four nutritionally adequate diets created by the U.S. Department of Agriculture (Carlson et al., 2007). This maximum benefit also depends on the size of the household, accounting for economies of scale in grocery shopping. In 2008, the maximum benefit amounts were \$162, \$298, \$426, and \$542 for household sizes 1-4, respectively. These amounts increased during and after the Great Recession (Table 1.4).

As ABAWDs have smaller household sizes with larger maximum benefit amounts per person, they account for a larger share of SNAP expenditures than SNAP caseload, as shown in Figure 1.1. ABAWDs should have zero dependent care costs and are also not eligible for excess medical cost deductions. The budget constraint with the food stamps then depends on the wage rate they face, whether ABAWDs have excess housing costs, and, lastly, if they can receive a waiver for the ABAWD work requirement.

For instance, an ABAWD living alone with the standard deduction in 2008 would only receive benefits beyond the \$10 per month minimum when their hourly wage was less than \$10.50. At the 2008-2009 federal minimum wage of \$6.55, an ABAWD with the same deductions as one making \$10.50 an hour would receive \$76 in benefits, making the total income available to both ABAWDs roughly equivalent after accounting for the food stamp benefits. Hence, the wage levels faced by the participant are a key driver in SNAP participation. Thus, I do not expect many people with high wages to participate in the program beyond the periods for which they are unemployed.

[7] Typically, it is 100% of the plan cost. The American Recovery and Reinvestment Act temporarily increased the maximum benefit amounts.

Using the program rules and the SNAP Quality Control Sample, I generated various static budget constraints faced by households with 25th - 75th percentile of deductions in Figure 1.2. The "cliff" (R. A. Moffitt, 2002) faced for ABAWDs with low wages represents the drop off in income available to them depending on whether they are eligible to receive benefits for that month.

1.3 Related Literature

The literature on the SNAP is rich and well-developed, given the size of the program (R. A. Moffitt, 2002; Fraker and R. Moffitt, 1989; Hoynes and Schanzenbach, 2012; Hagstrom, 1996). However, only a relatively small portion of papers focus on labor supply effects of ABAWD. The following section introduces seminal and recent studies related to SNAP and the ABAWD time limit.

In a review article on welfare programs and labor supply, R. A. Moffitt (2002) stated that there is little work disincentive from the food stamp program, and that food stamps affect consumption decisions largely the same way as equivalent increases in income, a claim later bolstered by Hoynes and Schanzenbach (2009).

Hoynes and Schanzenbach (2012) used the gradual roll-out of the food stamp program across counties to examine the labor supply effects of the program between the 1960s and 1970s- the same identification strategy used in their 2009 paper. Using the Panel Study of Income Dynamics (PSID), they found no evidence of labor force distortions for the general population, but did find small decreases in labor supply for specific groups who were more likely to participate in the welfare program. These decreases in labor supply were found for low education households as well as female-headed households and were mostly on the intensive margin, with a 25% reduction in annual hours.

The ABAWD-specific literature is less developed. Much of the research from the 90s and early 2000s focused on the new rules that came with PRWORA, such as (Stavrianos and Nixon, 1998; Czajka et al., 2001) and studies focused on single states (Richardson et al., 2003; Ribar, Edelhoch, and Liu, 2010; C. Gray et al., 2019). The findings were that many ABAWDs who left the Food Stamp Program after the welfare reform did not return to the program, and that 68% of ABAWDs were meeting the work requirement. Surveys with SNAP administrators suggested that many ABAWDs not meeting the work requirement were "unmotivated and did not want to participate in qualifying work [activities]" (Czajka et al., 2001) which would have allowed the ABAWD to receive food stamps while not meeting the 20 hours/week work requirement. (Stavrianos and Nixon, 1998) stated that the job prospects for ABAWDs participating in the food stamp may be hindered by structural unemployment, spatial mismatch, and poor professional networks, citing that many survey respondents have relied on only family contacts for employment applications. In other words, the ABAWD population participating in the food stamp program may have extra difficulties acquiring and keeping employment.

A few papers studied the ABAWD population specifically using administrative data from South Carolina (Richardson et al., 2003; Ribar, Edelhoch, and Liu, 2010). Both papers focused on spells of food stamp participation, with Richardson et al. (2003) finding that the reasons for exiting a food stamp spell did not vary by county waiver status. Ribar, Edelhoch, and Liu (2010) estimated that the ABAWD caseload in South Carolina would be 20% higher if the time limit waivers were suspended. Ribar, Edelhoch, and Liu (2010) also found that a disproportionate amount of exits from SNAP were at re-certification periods and encouraged more studies about ABAWDs and SNAP participation, noting their study focused on South Carolina's state options and may not be generalizable to the rest of the country.

Additionally, Moulton, Graddy-Reed, and Lanahan (2016) examined mothers phasing out of the Earned Income Tax Credit (EITC) eligibility by comparing single mothers with and without age eligible children. This is because the maximum EITC credit for a family with no children is relatively small and available at relatively low levels of earned income compared to one with one, two or more children. They found evidence of mothers leaving the labor force once they were no longer eligible to receive the EITC.

There are two analogous approaches for the ABAWD population. First, the "without dependents" requirement of ABAWD generally means that no children younger than 18 live in the household. Estimating the differences in labor force outcomes between new ABAWDs in waived versus non-waived regions could give an estimate of the influence of the time limit waiver; however, other household expenses may be drastically changing if the child is preparing to move out of the household.

Second, unlike the EITC, ABAWDs are able to age out of the additional work requirement when they turn 50 years old. Examining people around the upper age cutoff could also give an estimate of how effective the work requirement is influencing work. These local average treatment effects from aging out have been estimated with state administrative data (C. Gray et al., 2019; Harris, 2018), with evidence that the time limit does affect both participation and labor market outcomes.

1.4 Data

1.4.1 Survey of Income and Program Participation

This paper uses the Survey of Income and Program Participation (SIPP) panels from 2001, 2004, and 2008, which jointly cover monthly surveys from October 2002 through December 2013. The survey, conducted by the U.S. Census Bureau, interviews individuals every four months and the interview questions pertain to income, employment, work search, and participation in welfare programs for each of the previous months since the last interview. Once a household begins participating in a SIPP panel, it is included in future waves of interviews for that panel. Because the purpose of the SIPP is to inquire about programs primarily aimed at low-income households, families from low-income and low-socioeconomic status are over-sampled relative to a simple random sample of the population, making the SIPP ideal to study program participation. To account for the non-random sampling and attrition of SIPP respondents, I use the provided SIPP weights in all analysis. In addition, the SIPP has the additional benefit over the Current Population Survey (CPS), for example, of inquiring about income and participation monthly.

One drawback of the SIPP is evidence for "seam bias" as noted in Ham and Shore-Sheppard (2005). The bias stems from individuals under-reporting changes of status, such as employment, program participation, or source of insurance, within the four months of recall in a particular interview, and consequentially over-reporting changes between interviews. While seam bias has attenuated in more recent SIPP panels, Moore (2008) claims that seam bias is still prevalent and should be addressed when using the data. For instance, Table 1.2 shows a drastically higher transition to employment after four consecutive months of unemployment (13.19%) compared to a roughly 0.5% transition rate to employment after three or five reported months of unemployment. This is consistent with seam bias from the interview interval of four months.

Due to potential seam bias the under-reporting of changes within months covered by a single interview, and over-reporting of changes between interviews, I follow Ham and Shore-Sheppard, 2005 and estimate models using only the data from the most recent recall month in each interview. This entails dropping three fourths of the observations, as the SIPP asks about four months of information in each interview.

A second limitation of the SIPP is it does not directly ask whether a person in the household is an ABAWD. I address this limitation by inferring ABAWD status from reported ages, disability statuses, and number of children in a household. The definition of a household for SNAP does not necessarily match these definitions. For example, the presence of an elderly member in the same physical household may or may not count as a member of that household. That is, if the economic conditions are poor enough, the elderly member may qualify as their own household.

There are various measures of disability status in the SIPP, including whether the person receives Supplemental Security Income (SSI) or whether the respondent reported having a "work-limiting physical or mental condition." Pregnancy status, another exempting characteristic for ABAWD status, is offered as a reason for why the person did not have employment in a reference month. The question is only asked of respondents who did not work, so pregnancy status is otherwise unknown in the SIPP. Pregnant women are likely to be exempt from ABAWD status for other reasons, such as already having a dependent child. In March 2000, for instance, pregnant women accounted for 7.1% of the SNAP adult caseload in Wisconsin, but only 7.3% of that group were excluded from ABAWD status solely due to their current pregnancy (Czajka et al., 2001).

When ARRA was in effect and all ABAWDs were receiving time limit waivers in 2008 and 2009, there were 593,611 heads of household in the SIPP. Among them, 252,048 are households headed by a person aged 50 or more, above the ABAWD age threshold. Among the remainder, 197,783 households had at least one child living in the household and 15,552 heads of household had a disability. This leaves 128,228 ABAWD heads of household in the two years worth of data, about 38% of the adult-headed households.

I analyze heads of household because food stamp eligibility and benefits are allocated at the household level, not for individuals or families. Only head of household adults aged 18-49 years old and without disabilities are included in the estimation. Thus, whether the household has children is the sole observable factor differentiating ABAWDs from non-ABAWDs in the estimations.

The designation of ABAWD can vary across an individual's life. During nation-wide waivers granted by ARRA, ABAWDs accounted for 48% of households headed by a person aged 18-32, 26% with a head aged 33-39, and 36% for heads aged 40-49. These trends reflect the fact that households leave ABAWD status by having a child in the household, and that they become ABAWDs again once their children move away.

Table 1.3 shows the educational attainment distribution by race, participation, and ABAWD status in the SIPP. Among ABAWDs participating in SNAP between 2008 and 2009, 13% have no high school degree, 52% completed high school, and 34% have a college degree. Overall, the majority of ABAWDs (59%) in the SIPP, including both participants and non-participants of SNAP, have education beyond high school. These means do not vary much between white and nonwhite ABAWD, as can be observed from Panels B and C in Table 1.3.

Lastly, Tables 1.5 and 1.6 contain descriptive statistics for key variables and socioeconomic characteristics. Table 1.5 gives the statistics for all household heads, while Table 1.6 displays the descriptive statistics of only households who report receiving SNAP benefits. Unlike Table 1.3, Tables 1.5 and 1.6 are constructed using all three SIPP panels, not just during the time period in which the American Recovery and Reinvestment Act (ARRA) of 2009 is in effect. Consequently, the unemployment averages do not simply reflect the Great Recession. Compared to all household heads, we see that SNAP participants are less likely to be white, are more likely to not have a high school diploma and to live in a non-metro area.

For ABAWDs, we see that the current state unemployment rate is half of a percentage point higher for SNAP participants despite being no different from the whole population, suggesting that the waivers may influence their take-up rates. Unsurprisingly, participants have much less reported earned and total income than non-participants.

1.4.2 Time Limit Waivers

I use time limit waiver data obtained from a Freedom of Information Act (FOIA) request to the Food and Nutrition Service of the US Department of Agriculture. In the data, I have lists of counties, cities, and other regions that received ABAWD time limit waivers from April 2009 through June 2016. The waivers from the FOIA request show that either the entire state was covered by a waiver, or specific counties and/or Indian reservations were covered. Generally, states either received state-wide waivers or for a cluster of counties in a certain area. The most common non-county area is the Indian reservation, though particular cities and towns received waivers in Vermont and New Hampshire and South Dakota received waivers for “Economic Areas” corresponding to their three largest cities.

For waivers before and after that time period, I collected records publicly available from the FNS and the state SNAP administration documents as applicable. However, I utilize the public use version of the SIPP that only contains identifiers for state of residence[8] and whether the household is in a metro area. Thus, to incorporate the waiver information, I use the percentage of persons in the state living in waived regions as a proxy of exposure to the time limit for households in that state. A future expansion of this work will incorporate SIPP restricted data. Using both the uncensored version of SIPP and full waiver details would allow for clearer identification of which ABAWDs lived in waived areas in different time periods.

While the measure of whether the ABAWD is currently subject to the time limit waiver is imperfect, the presence of the waiver somewhere in the state may have implications for ABAWDs in non-waived regions. This is because the stock of 15 percent waivers available in the state does not immediately decrease when the waivers go into effect. Thus, there are more individual waivers available for a smaller population.

Table 1.1 reports the sample weighted SNAP participation rates, where one can see how drastically participation rates vary by education and race. This finding is largely consistent with unemployment rates by the same factors, and at least with respect to education, consistent with the fact that high-skill workers are unable to participate in SNAP if they receive high wages.

These participation rates are *ex post*, which is not in line with how SNAP eligibility works: if a household experiences a change in their employment that makes them eligible to participate, they may apply for benefits immediately, forgoing any waiting to prove their new status. From Table 1.3, we see that 4% of ABAWD households and 20% of non-ABAWD households reported participation in SNAP during ARRA.

1.5 Model

I examine the effect of the additional work requirement imposed on ABAWDs on three main outcomes of interest. The first two outcomes are measures of labor supply: (1) whether

[8] In the 2001 Panel, some states with relatively small populations were merged together, such as North and South Dakota.

the head of household works 20 hours or more per week and (2) the number of hours worked, conditional on being employed. The third is an indicator variable for SNAP participation. For the two binary dependent variables, I estimate equation (1.5) with the logit estimator. For a continuous measure of labor supply, I rely on the OLS estimator with survey weights. I compare the differences in the outcomes between ABAWDs who do or do not face the additional work requirement in the time period to non-ABAWDs, whom are not subject to the time limit waivers. With the public SIPP data, I estimate model (1.5) below:

$$\begin{aligned} y_{ist} =& \alpha + \gamma \text{ABAWD}_{it} + \sigma \text{WPctPopWaived}_{st} + \delta(\text{ABAWD}_{it} \text{ x WPctPopWaived}_{ist}) \\ &+ X_{ist}\beta + \Sigma_s \eta_{ist} S_s + \Sigma_t \lambda_{ist} T_t + \epsilon_{ist} \end{aligned} \tag{1.5}$$

Where i, s, and t index head of household, state, and time, respectively and y_{it} refers to one of the three outcomes of interest. For the time effects, t can either index the year and month of the observation, or the SIPP wave if there are too few observations in a cell for non-linear models to converge. $ABAWD_{it}$ is an indicator variable for able-body adult status, $WPctPopWaived$ refers to the percent of state s's population living in a waived county, and X_{ist} includes demographic characteristics such as race, sex, age, veteran status, and educational attainment. Indicator variables are used to control for state- and time- specific effects, as both employment and SNAP details vary across those dimensions.[9]

While households in SIPP remain in the panel for every wave after their initial wave, estimating individual fixed effects models are difficult. First, the paper uses four different SIPP panels appended together. An individual household may be in the data set for the duration of a single panel, but no longer. They also may leave or join the panel after the initial wave, so that the length of observation is rather short. The ABAWD status does vary within an individual across SIPP, but for less than 10 percent of the population. Lastly, state invariant characteristics are represented by S_s.

[9] I have also estimated this model with state-specific time trends, allowing them to change while under the waiver. The average partial effects of the coefficient of interest remain within 0.1 percentage points. The analysis with county level data will allow me to conduct this analysis without disregarding partially-waived states.

Equation (1.5) allows for the difference in outcomes of ABAWDs in waiver states vs non-waived states, relative to non-ABAWDs competing in the same labor markets. The coefficient of interest, δ, can be interpreted as the change in the dependent variable for an ABAWD when the time limit is waived. With y_{ist} being a positive measure of labor supply, a $\hat{\delta} < 0$ would be consistent with the time limit affecting the labor supply of ABAWDs. To be able to claim that the waivers influence labor supply through the welfare channel, and not through some other means, the estimates of δ should be non-zero.

When y_{ist} is a dummy variable for food stamp receipt, then a positive sign would imply that the time limit waiver acts as a barrier for participation in SNAP. If $\hat{\sigma} \neq 0$, then the percent of counties waived variable is picking up on something beyond SNAP eligibility, as the waiver status for non-ABAWDs has no direct channel to influence their behavior in either the labor market or SNAP participation. As the labor supply for household heads, as well as other covariates are generally different between ABAWDs and non-ABAWDs, I estimate versions of equation (1.5) on the ABAWD population only. This allows for a more comparable control group at the cost of losing state-specific information on economic and SNAP conditions. The equation estimated in this case is then:

$$y_{ist} = \alpha + c_i + \delta \text{WPctPopWaived}_{st} + X_{ist}\beta + \Sigma_s \eta_{ist} S_s + \Sigma_t \lambda_{ist} T_t + \epsilon_{ist} \tag{1.6}$$

Lastly, high-skilled workers' labor force outcomes should not be influenced by SNAP, as wages even within a few dollars of the federal minimum wage yield little to no benefits while working. Hence, I only include heads of household with a high school education or less in the regression population, except when I specifically examine those with higher educational attainment. This serves as a falsification test, where non-zero estimations of $\hat{\delta}$ suggest that a factor correlated with ABAWD time limit waivers is affecting SNAP participation. Further, the decision to only regress upon a lower education population is consistent with Hoynes and Schanzenbach (2009) and Hoynes and Schanzenbach (2012).

1.6 Results

1.6.1 Estimations on Total Sample of Household Heads

Estimation results of the logit model can be found in Tables 1.7 and 1.8, with the average partial effects (APEs) from the logit estimations are displayed in Table 1.9. Table 1.10 contains the OLS results for number of hours worked. Overall, I find evidence that the ABAWD time limit has a statistically significant effect on SNAP participation and whether ABAWDs meet the work requirement, though the magnitude of the effect is small. In cases where the logit point estimates are not statistically significant, the average partial effects from the estimation are non-zero. In contrast, I find no evidence of the time limit has an effect on the number of hours worked (the continuous measure of labor supply).[10]

The average partial effects (Table 1.9, Column 1) of the time limit waiver on SNAP participation suggest the waivers increased SNAP participation for ABAWDs by 0.99 percentage points, when including state- and time-invariant factors and the full specification of (1.5). The effects of waivers for non-ABAWDs are statistically zero, providing evidence the ABAWD population is responding to the waivers differently than non-ABAWDs, and that the presence of the time limit waivers is not confounded with some other economic factor (i.e. poor employment conditions, high food prices) that could potentially explain participation.

The results for the work requirement are in Table 1.8. I find a negative and statistically significant effect of the waiver on whether the ABAWD meets the 20 hour per week work requirement. The average partial effects in Table 1.9, Column 2 for these estimates suggest that an ABAWD is 2.6 pp less likely to meet the 20-hour work requirement when living in a waived region. Thus, the work requirement has a small influence on labor supply for the targeted population. This effect size is slightly larger (2.0 pp) than the local average treatment effect found in Harris (2018), but generally consistent with the literature. The

[10] Though not implied by the finding of the policy not having an influence on hours worked, further (unreported) analysis yields no evidence that the distribution of hours worked around 20 hours per week was affected by these policies. That is, I find no evidence to suggest that ABAWDs under the time limit are working "just above" the limit as opposed to meeting the requirement generally.

time limit requirement makes it more difficult for ABAWDs to participate in SNAP and ABAWDs are less likely to meet the work requirement when the time limit waiver is in effect.

1.6.2 Estimations on ABAWD Sample

The "B" panels of Tables 1.7, 1.8, 1.10 show the results when restricting the estimation sample to ABAWDs only. In this case, the indicator for ABAWD and interaction term are omitted. I find a mix of no significant effects and coefficients similar in magnitude than in Panel A of the waiver status on SNAP participation and labor market outcomes. The point estimates and marginal effects are of similar magnitude and have the same sign (APE of +0.7 pp for SNAP participation and -2.8 pp for meeting the 20-hour work requirement) but are statistically insignificant. Estimating on this group alone drastically reduces the number of observations to estimate the state- and time-fixed effects that describe the accessibility of SNAP and labor market conditions. The inclusion of SNAP policy variables, or an index as proposed by Stacy, Tiehen, and Marquardt (2018), as well as detailed county-level economic conditions may allow for more precision in these estimates. However, I currently do not have access to county of residence so I am unable to utilize the variations in economic conditions at the county level.

1.7 Discussion

The ABAWD time limit is a policy that provides additional requirements to receive food assistance for a population that receives virtually no other forms of welfare assistance. Hence, policy makers should carefully consider whether the only means-tested welfare program they can use needs additional means-testing for this population. Monitoring and record keeping for this time limit comes with additional administrative work as well.

An additional complication to the ABAWD time limit is the salience of rules for qualification. Prior to the Great Recession, many states did not receive blanket waivers for their ABAWDs, and since the Great Recession, states have improved their ability to communicate how the time limits work. Still, it can be difficult for ABAWDs to know for precisely when they may qualify or how much they may receive and may be deterred from applying to receive food assistance in the first place.

The purpose of the work requirement is also unclear and other policies might be more effective in fulfilling set objectives. For example, if the goal of the work requirement is to encourage ABAWDs to work while receiving public assistance, perhaps an increase in the relatively small Earned Income Tax Credit (EITC) maximum benefit amount for households with zero children would be effective. For 2019, that maximum amount is \$529 a year, or \$44 per month, which is only available to individuals making earning less than \$9,000 in a year, with the entire benefit amount going to zero when earning less than \$14,350. Such a benefit would help alleviate the disincentive for earning more income on the margin while receiving SNAP benefits, as they phase out at \$0.24 per dollar of earned income.

For a goal of increasing the available food budget to ABAWDs (improving food security) while maintaining an incentive to work, the government may change the budget constraints faced by ABAWDs participating in SNAP. Increasing the maximum benefit amount or increasing the deductions for either excessive shelter costs or earned income would all provide higher SNAP benefit amounts for ABAWDs. Increasing the benefit amount for smaller households to the Low-Cost USDA Food plan would increase the maximum benefit about by about 25% or \$50 per month in early 2020. For an ABAWD earning \$10 an hour, they could potentially work an additional 20.8 hours per month while maintaining the same SNAP benefit amount before the policy change.

A more generous deduction for excessive shelter costs may benefit ABAWDs living alone, as they may not have an additional person to share rent expenses with. SNAP administrators, should they choose, could then target advertisements to areas with higher rent to income areas saying that SNAP may be able to help with their tight budgets. Increasing ABAWD participation in SNAP may make the work requirement more binding should they be aware that their financial situation is eligible for SNAP, and that their excess housing makes their benefits slightly larger.

Lastly, any potential reform to the ABAWD time limit is the heterogeneous effects found the APE tables of this policy across racial minorities. One potential reason could be due to higher need for SNAP in areas where such groups are located. Another could be related to the generally worse labor market outcomes obtained by racial minorities and people with lower educational attainment, even when controlling for sensible observable factors. More detailed information on county of residence to better control for local economic conditions to begin to attempt the answer, though a case study and other research into the disparities is warranted to be able to suggest anything conclusively.

1.8 Conclusion

The ABAWD time limit is an additional work requirement for able-bodied adults without dependents to participate in the Supplemental Nutrition Assistance Program. In order to participate beyond three months in a 36 month period, they must work at least 20 hours per week, and it doing so they earn wages which decreases their SNAP benefit amounts. This policy has been the law since the major welfare reform of 1996, with its intention to have people who ought to be able to work do so while participating in welfare programs.

I study the effectiveness of the ABAWD time limit for SNAP receipt using data from three panels of the SIPP (2001, 2004, and 2008). I find that the ABAWD time limit, which is designed to encourage labor force participation for those seeking welfare benefits, does increase labor force participation on the extensive margin (2.6 pp) but not on the intensive margin. Further, the time limit waivers decrease the likelihood an ABAWD receives SNAP benefits by 1.1 pp.

Further study is needed to declare whether the time limit achieves the goals of increasing labor force participation with sufficiently low administrative cost. The demographic characteristics that constitute an ABAWD make them categorically ineligible for many other forms of federal welfare programs, so this policy should be examined carefully for this population. Lastly, formal welfare analysis is needed with special attention to racial minority groups and persons with low-levels of education to determine whether deterring SNAP participation is worth the trade off of higher labor force participation.

APPENDICES

APPENDIX 1A: Tables

Table 1.1: SNAP Participation Rate by Race, Education Status, and ABAWD Status

		Education Level		
	All	Less than HS	HS Only	More than HS
Panel A: All Races				
All family types	.137	.365	.172	.071
ABAWDs	.037	.104	.052	.021
Non-ABAWDs	.199	.436	.243	.108
Panel B: White				
All family types	.109	.313	.136	.055
ABAWDs	.027	.083	.037	.017
Non-ABAWDs	.161	.377	.196	.084
Panel C: Non-white				
All family types	.235	.552	.300	.131
ABAWDs	.072	.178	.108	.039
Non-ABAWDs	.334	.659	.400	.200

Note: Household participation rates for taken from 2009-March to 2010-September, when ARRA waived time limits for all ABAWDs. Education and race represent the household head's values. Source: Survey of Income and Program Participation, author's calculations.

Table 1.2: Reported Transition Rates to Unemployment

Length of Unemployment Spell (months)	Percent reporting: Employment	Unemployment
0	95.48	4.52
1	0.68	99.32
2	0.62	99.38
3	0.52	99.48
4	13.19	86.81
5	0.55	99.45
6	0.41	99.59
7	2.19	97.81
8	1.19	98.32
9	2.94	97.06
10	0.00	100.00

Source: 2008 Survey of Income and Program Participation, author's calculations. The estimations reported above are for the reference month.

Table 1.3: Education Level by Race, SNAP Participation, and ABAWD Status

	All	ABAWD			Non-ABAWD		
		Overall	SNAP	No SNAP	Overall	SNAP	No SNAP
Panel A: All Races							
All education levels	1.00	.380	.038	.962	.620	.201	.799
No HS Degree	.091	.046	.133	.043	.119	.256	.084
HS Degree Only	.377	.360	.523	.353	.387	.472	.366
Advanced Degree	.532	.594	.343	.601	.594	.272	.550
Panel B: White							
All education levels	.785	.383	.030	.970	.617	.162	.838
No HS Degree	.091	.044	.112	.042	.120	.275	.090
HS Degree Only	.373	.360	.520	.355	.381	.464	.365
Advanced Degree	.536	.596	.360	.603	.499	.261	.545
Panel C: Non-white							
All education levels	.215	.367	.069	.931	.633	.339	.661
No HS Degree	.092	.055	.156	.047	.114	.223	.058
HS Degree Only	.391	.357	.527	.345	.410	.486	.370
Advanced Degree	.518	.588	.316	.608	.477	.291	.572

Note: Participation rates taken from 2009-March to 2010-September, when ARRA waived time limits for all ABAWDs. Source: Survey of Income and Program Participation, author's calculations.

Table 1.4: Maximum Allotment Benefits for SNAP Participants by Household Size

	Household Size			
Year	1	2	3	4
2004	\$141	\$259	\$371	\$471
2005	\$149	\$274	\$393	\$499
2006	\$152	\$278	\$399	\$506
2007	\$155	\$284	\$408	\$518
2008	\$162	\$298	\$426	\$542
2009	\$176	\$323	\$463	\$588
2010	\$200	\$367	\$526	\$668
2011	\$200	\$367	\$526	\$668
2012	\$200	\$367	\$526	\$668
2013	\$200	\$367	\$526	\$668
2014	\$189	\$347	\$497	\$632
2015	\$194	\$357	\$511	\$649
2016	\$194	\$357	\$511	\$649
2017	\$194	\$357	\$511	\$649

Source: U.S. Department of Agriculture, Food and Nutrition Service (FNS)

Table 1.5: Descriptive Statistics of All Household Heads

	All		Abawds		Non-Abawds	
	Mean	SD	Mean	SD	Mean	SD
White	.796		.816		.786	
Age	36.8	8.02	36.2	8.93	37.2	7.44
No High School Diploma	.099		.053		.124	
High School Diploma Only	.423		.399		.437	
Household Size	3.1	1.6	1.8	.9	3.8	1.4
# Children in HH	1.2	1.3	0	0	1.8	1.2
Live in Metro Area	.773		.800		.759	
N	563,770		200,273		363,497	
Pct Counties Waived, no ARRA	.253		.257		.251	
N	263,542		94,581		168,961	
Pct Counties Waived, w/ ARRA	.432		.435		.430	
N	310,885		111,565		199,320	
State Unemployment Rate	6.6	2.1	6.6	2.1	6.6	2.1
Employed Last Month	.796		.816		.786	
Hours Worked/Week, if positive	42	13	43	13	41	13
Worked 20 or more hours/week	.796		.816		.786	
Household Earned Income	5,077	5,360	5,177	5,011	5,056	5,545
Household Earned Income/person	1,995	2,384	3,106	3,133	1,383	1,535
Household Total Income	5,389	5,393	5,315	5,076	5,429	5,559
Household Total Income/person	2,112	5,360	3,212	3,152	1,383	1,535
SNAP Participant	.119		.028		.168	
N	563,770		200,273		363,497	
Food Stamp Benefit (if positive)	320	220	213	165	330	221
Food Stamp Benefit per person (if positive)	95	63	121	100	93	57
N	66,866		5,703		61,163	

Note: Heads of Household responding to the most recent reference month only. Dollar amounts are monthly and in 2009 dollars using the personal consumption expenditure price index (PCEPI).

Table 1.6: Descriptive Statistics of SNAP Participants

	All		Abawds		Non-Abawds	
	Mean	SD	Mean	SD	Mean	SD
White	.627		.624		.627	
Age	35.1	8.3	36.3	9.5	34.9	8.1
No High School Diploma	.261		.157		.270	
High School Diploma Only	.51		.541		.508	
Household Size	3.7	1.8	2.08	1.1	3.8	1.8
# Children in HH	1.9	1.5	0	0	2.1	1.5
Live in Metro Area	.715		.733		.713	
N	66,866		5,703		61,163	
Pct Counties Waived, no ARRA	.270	.363	.312	.393	.266	.361
N	31,091		2,371		28,720	
Pct Counties Waived, with ARRA	.482	.450	.554	.455	.476	.449
N	38,283		3,116		35,167	
State Unemployment Rate	7.1	2.2	7.5	2.2	7.0	2.2
Employed Last Month	.546		.735		.528	
Hours Worked/Week (if positive)	37	14	38	15	41	37
Worked 20 or more hours/week	.455		.640		.438	
Household Earned Income	1,339	2,268	1,890	2,645	1,288	2,222
Household Earned Income/person	361	608	829	1,028	317	1,535
Household Total Income	1,883	2,394	2,287	2,744	1,846	2,356
Household Total Income/person	555	661	1,015	1,036	512	596
N	66,866		5,703		61,163	

Note: Heads of Household responding to the most recent reference month only. Dollar amounts are monthly and in 2009 dollars using the personal consumption expenditure price index (PCEPI).

Table 1.7: Logistic Regression Results for SNAP Participation

Panel A: Total Sample of Household Heads	(1)	(2)	(3)
ABAWD	-1.88***	-1.84***	-1.81***
	(0.72)	(0.08)	(0.08)
Waiver	0.53***	0.16***	0.05
	(0.04)	(0.05)	(0.05)
ABAWD * Waiver	0.27***	0.20*	0.16
	(0.11)	(0.11)	(0.11)
N	118,571	118,571	118,571
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			X

Panel B: Sample of ABAWD Household Heads	(1)	(2)	(3)
Waiver	0.80***	0.23*	0.19*
	(0.10)	(0.13)	(0.11)
N	32,394	32,394	32,236
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level, respectively. Survey weights from SIPP used in estimation and calculation of standard errors.

Table 1.8: Logistic Regression Results for Meeting the Work Requirement

Panel A: Total Sample of Household Heads	(1)	(2)	(3)
ABAWD	1.28***	1.38***	1.38***
	(0.05)	(0.06)	(0.06)
Waiver	-0.32***	-0.04	-0.01
	(0.03)	(0.04)	(0.05)
ABAWD * Waiver	-0.23***	-0.23***	-0.23***
	(0.07)	(0.07)	(0.07)
N	103,286	103,286	103,286
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			X
Panel B: Sample of ABAWD Household Heads	(1)	(2)	(3)
Waiver	-0.55***	-0.12	-0.20**
	(0.07)	(0.10)	(0.09)
N	27,628	27,628	27,590
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level, respectively. Survey weights from SIPP used in estimation and calculation of standard errors.

Table 1.9: Average Partial Effects, Logit Estimation

	(1)	(2)
	Snap Participation	Work Requirement
All-ABAWDs	0.99%**	-2.61%***
	(0.50)	(0.80)
Race		
Non-White	1.37**	-2.77***
	(0.69)	(0.85)
Black	2.09**	-3.22***
	(1.05)	(0.98)
Hispanic (non-white)	0.94*	-2.41***
	(0.48)	(0.73)
Education		
No High School Degree	0.77**	-2.23***
	(0.40)	(0.70)

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level, respectively. Survey weights from SIPP used in estimation and calculation of standard errors. Estimates and standard errors reported in percents. APEs taken from fully specified models only, including controls demographics, regional employment, state- and time- effects.

Table 1.10: Regression Results for Log Number of Hours Worked

Panel A: Total Sample of Household Heads	(1)	(2)	(3)
ABAWD	0.05***	0.04***	0.04***
	(0.01)	(0.01)	(0.01)
Waiver	-0.02***	0.01	0.01
	(0.01)	(0.01)	(0.01)
ABAWD * Waiver	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
N	76,344	76,344	76,344
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			X

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level, respectively. Survey weights from SIPP used in estimation and calculation of standard errors.

APPENDIX 1B: Figures

Figure 1.1: ABAWD Share of SNAP over Time

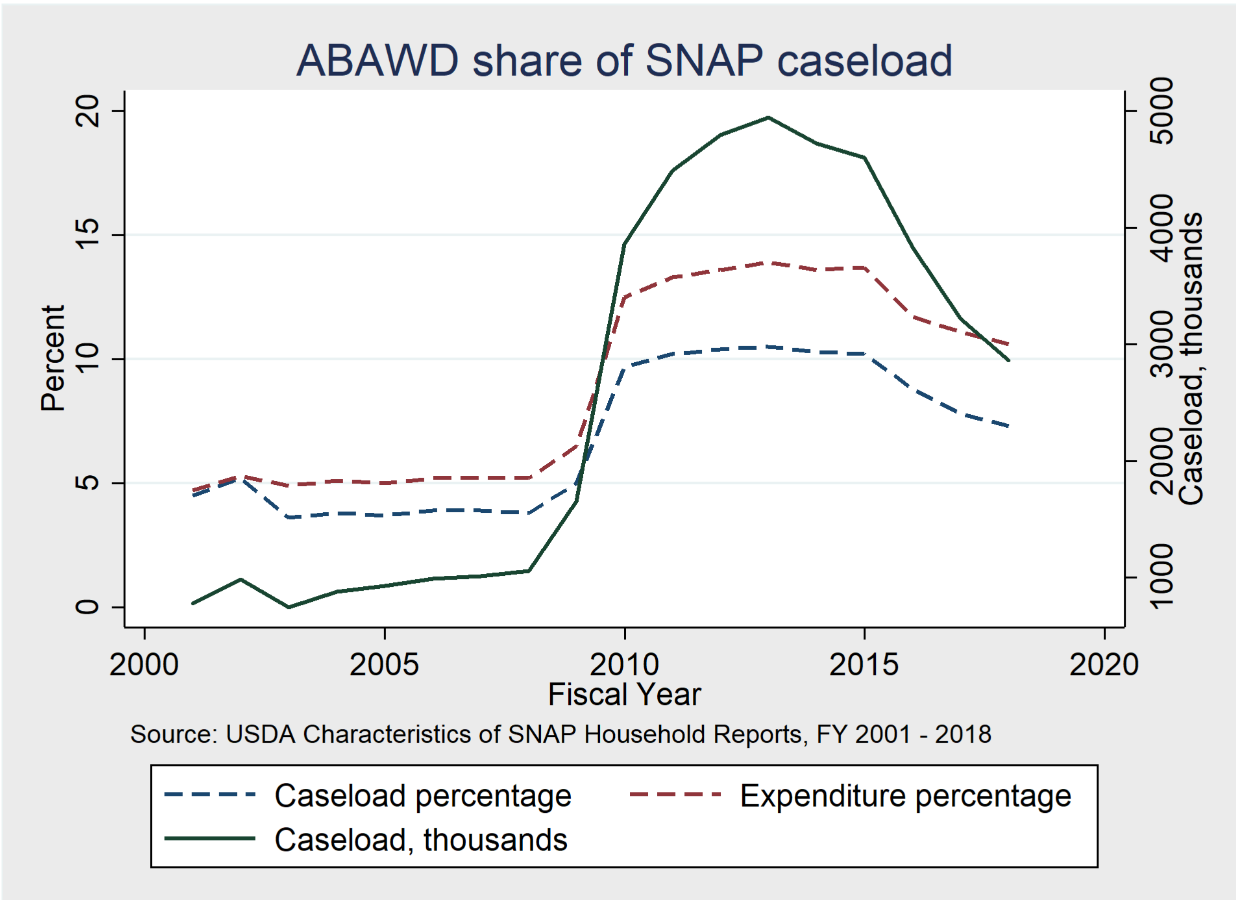


Figure 1.2: Static Budget Constraints for Single ABAWDs

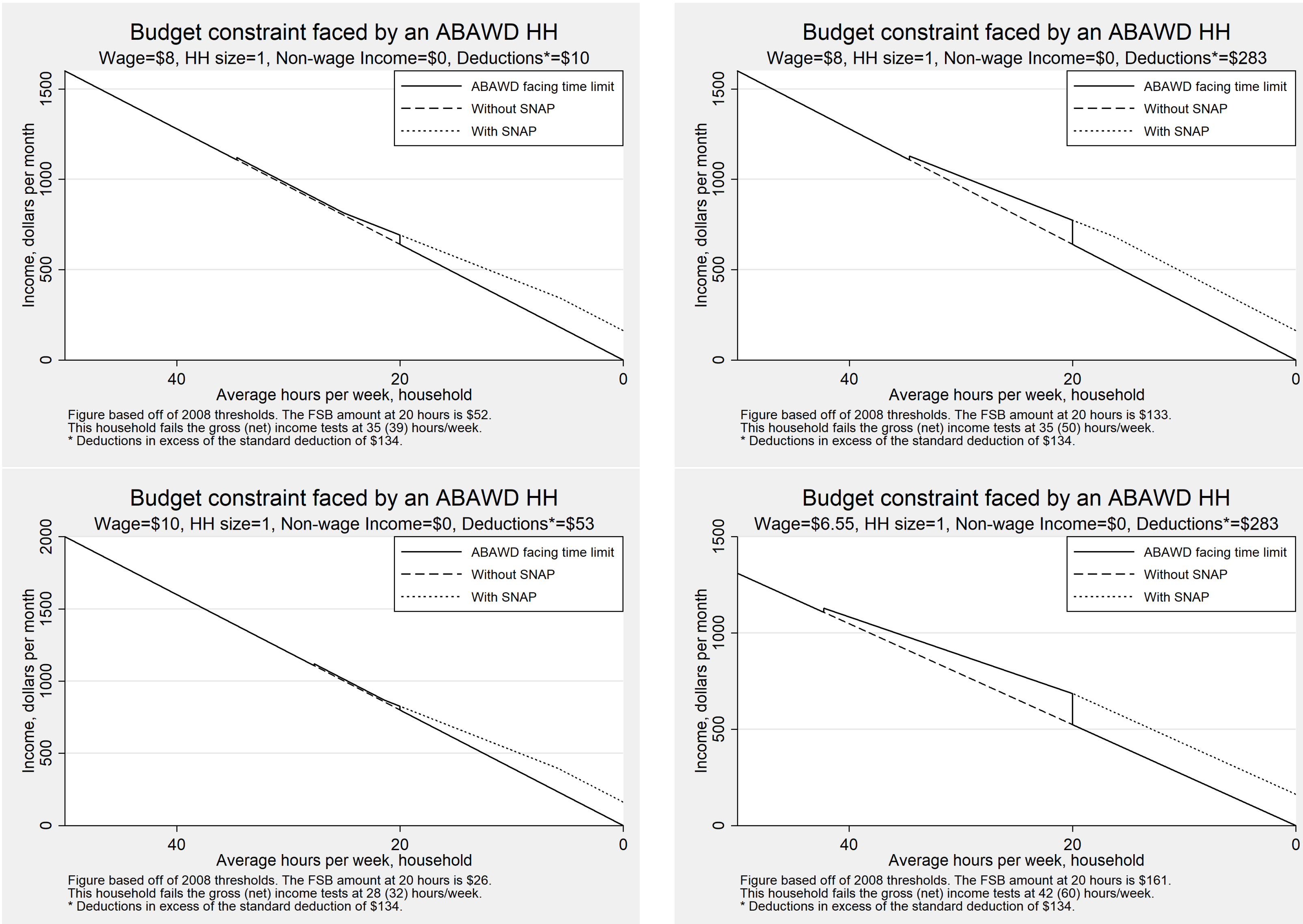
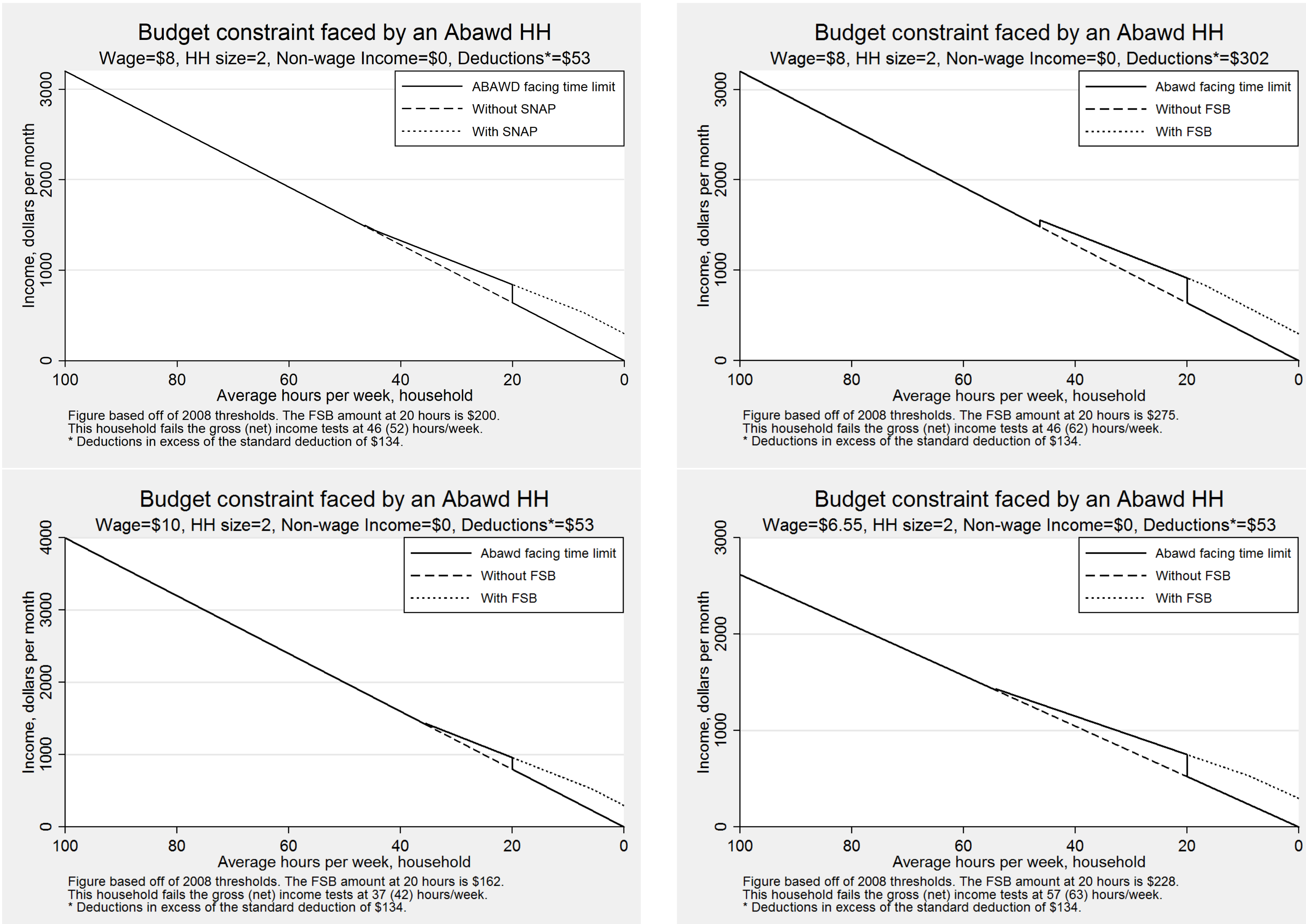


Figure 1.3: Static Budget Constraints for ABAWD Households



APPENDIX 1C: Probit Model Results

In addition to estimating a logit model to analyze the impact of the ABAWD time limit and its waiver, I estimated probit model. For SNAP participation, I find insignificant effects of the waivers for both the probit point estimates and the probit average partial effects. However, the magnitude of the coefficients (Table 1.11) and APEs (Table 1.13) are similar to those found in logit.

A key difference between logit and probit may explain the differences in the findings for SNAP participation, as the probit probability function has flatter tails. The implication is that even modest changes in the indexing function (from adding $\hat{\delta}$ when the time limit is waived) only slightly changes the predicted probabilities for people who are already very likely or very unlikely to be participating in SNAP. Nevertheless, these probit models must be more closely analyzed, especially since the full specification (3) has the opposite sign as found in the logit estimation.

For the work requirement, where ABAWDs are found to be more responsive to changes in waiver status, the point estimates (Table 1.12) are similar to those found in Table 1.8. The APEs were also similar, and slightly larger in magnitude than the logit APEs (Table 1.9).

Table 1.11: Probit Regression Results for SNAP Participation

Panel A: Total Sample of Household Heads	(1)	(2)	(3)
ABAWD	-0.94***	-0.85***	-0.86***
	(0.03)	(0.03)	(0.03)
Waiver	0.31***	-0.15***	0.03
	(0.02)	(0.03)	(0.03)
ABAWD * Waiver	-0.06	0.03	-0.02
	(0.05)	(0.05)	(0.05)
N	118,051	118,051	118,051
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			X

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level. Survey weights from SIPP used in estimation and calculation of standard errors.

Table 1.12: Probit Regression Results for Meeting the Work Requirement

Panel A: Total Sample of Household Heads	(1)	(2)	(3)
ABAWD	0.72***	0.61***	0.61***
	(0.03)	(0.03)	(0.03)
Waiver	-0.20***	-0.05*	-0.00
	(0.02)	(0.03)	(0.03)
ABAWD * Waiver	-0.09**	-0.12***	-0.12***
	(0.04)	(0.04)	(0.04)
N	102,846	102,846	102,846
Demographic Characteristics		X	X
Economic Conditions		X	X
State FE			X
Time FE			X

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level. Survey weights from SIPP used in estimation and calculation of standard errors.

Table 1.13: Average Partial Effects, Probit Estimation

	(1)	(2)
	APE (Snap Participation)	APE (Work Requirement)
All-ABAWDs	0.84%	-2.67%***
	(0.72)	(0.90)
Race		
Non-White	0.79	-2.82***
	(0.74)	(0.95)
Black	1.12	-3.20***
	(1.09)	(1.07)
Hispanic (non-white)	0.56	-2.51***
	(0.54)	(0.85)
Education		
No High School Degree	0.89	-3.34***
	(0.84)	(1.12)

Note: ***, **, and * correspond to significance at the .01, .05, and .1 level. Survey weights from SIPP used in estimation and calculation of standard errors. Estimates and standard errors reported in percents. APEs taken from fully specified models only, including controls demographics, regional employment, state- and time- effects.

APPENDIX 1D: Paper Extensions

A possible extension and improvement of this paper would be to use restricted SIPP data. For this version of the paper, I do not have access to the county of residence and use only state-level data. Using county-level data would allow us to observe the specific counties receiving time limit waivers and the counties in which the survey respondents reside, yielding a richer model. I would then estimate equation (1.7):

$$\begin{aligned} y_{ict} =& \alpha + \gamma \text{ABAWD}_i t + \sigma \text{WaivedCounty}_{ct} + \delta \text{ABAWD x WaivedCounty}_{ict} \\ &+ X_{ict}\beta + \Sigma_s \eta_{ict} S_s + \Sigma_t \lambda_{ict} T_t + \epsilon_{ict} \end{aligned} \tag{1.7}$$

Additional considerations for controls include the USDA Economic Research Service's SNAP Policy data set (Stacy, Tiehen, and Marquardt, 2018) to examine the extent that general SNAP policies influence ABAWD participation. Additionally, leveraging SIPP supplemental questionnaire information as child support payments paid can help inform the budget constraints faced by some of the ABAWDs as well as their potential benefit amounts. Previous to 2002, states were not allowed to adjust net income for child support payments paid, but the impacts of this policy option remains unstudied for ABAWDs.

REFERENCES

REFERENCES

Bolen, Ed and Stacy Dean (2018). *Waivers Add Key State Flexibility to SNAP's Three- Month Time Limit*. en. Tech. rep. Washington DC: Center on Budget and Policy Priorities, p. 10.

Carlson, Andrea et al. (2007). *Thrifty Food Plan 2006*. en. Tech. rep. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion.

Cronquist, Kathryn (2019). *Characteristics of Supplemental Nutrition Assistance Program Households: Fiscal Year 2018*. en. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

Czajka, John L et al. (2001). *Imposing a Time Limit on Food Stamp Receipt: Implementation of the Provisions and Effects on Food Stamp Program Participation*. en. Tech. rep. U.S. Department of Agriculture, Food and Nutrition Service.

Falk, Gene, Maggie McCarty, and Randy Aussenberg (2014). "Work Requirements, Time Limits, and Work Incentives in TANF, SNAP, and Housing Assistance". en. In: *Congressional Research Service*, p. 35.

Foley, Arthur (2007). *Overuse of the 15 Percent Able Bodied Adults Without Dependents (ABAWD) Ecemptions by State Agencies.*

Fraker, Thomas and Robert Moffitt (1989). *The Effect of Food Stamps on the Labor Supply of Unmarried Adults without Dependent Children.*

Ganong, Peter and Jeffrey B. Liebman (Nov. 2018). "The Decline, Rebound, and Further Rise in SNAP Enrollment: Disentangling Business Cycle Fluctuations and Policy Changes". en. In: *American Economic Journal: Economic Policy* 10.4, pp. 153–176. ISSN: 1945-7731, 1945-774X.

Gray, Colin et al. (2019). "Employed in a SNAP? The Impact of Work Requirements on Program Participation and Labor Supply". en. In: p. 68.

Gray, Kelsey Farson (2014). *Characteristics of Food Stamp Households: Fiscal Year 2013*. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

Gray, Kelsey Farson and Esa Eslami (2014). *Characteristics of Food Stamp Households: Fiscal Year 2012*. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

Hagstrom, Paul A. (1996). "The Food Stamp Participation and Labor Supply of Married Couples: An Empirical Analysis of Joint Decisions". en. In: *The Journal of Human Resources* 31.2, p. 383. ISSN: 0022166X.

Hahn, Heather et al. (2019). *SNAP Work Requirements in Arkansas for Adults without Dependents or Disabilities*. en. Tech. rep. The Urban Institute.

Ham, John C. and Lara Shore-Sheppard (Jan. 2005). "The effect of Medicaid expansions for low-income children on Medicaid participation and private insurance coverage: evidence from the SIPP". en. In: *Journal of Public Economics* 89.1, pp. 57–83. ISSN: 00472727.

Harris, Timothy F. (Dec. 2018). *Do SNAP Work Requirements Work?* en. Tech. rep. W.E. Upjohn Institute.

Hoynes, Hilary Williamson and Diane Whitmore Schanzenbach (2009). "Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program". In: *American Economic Journal: Applied Economics* 1.4, pp. 109–139.

— (Feb. 2012). "Work incentives and the Food Stamp Program". en. In: *Journal of Public Economics* 96.1-2, pp. 151–162. ISSN: 00472727.

Moffitt, Robert A (2002). "Welfare Programs and Labor Supply". en. In: *Handbook of Public Economics* 4, pp. 2393–2430.

Moore, Jeffrey (2008). "Seam Bias in the 2004 SIPP Panel: Much Improved, Much Bias Still Remains". In: *Survey Methodolgy Research Report*. Vol. 3. Washington DC: U.S. Census Bureau, Statistical Research Division.

Moulton, Jeremy G., Alexandra Graddy-Reed, and Lauren Lanahan (June 2016). "Beyond the EITC: The Effect of Reducing the Earned Income Tax Credit on Labor Force Participation". en. In: *National Tax Journal* 69.2, pp. 261–284. ISSN: 00280283.

Ribar, David C, Marilyn Edelhoch, and Qiduan Liu (2010). "Food Stamp Participation among Adult-Only Households". en. In: *Southern Economic Journal* 77.(2), p. 28.

Richardson, Philip et al. (2003). *Food Stamp Leavers Research Study—Study of ABAWDs Leaving the Food Stamp Program in South Carolina: Final Report*. en. Tech. rep. Economic Research Service, p. 152.

Rosenbaum, Dorothy (2013). *The Relationship Between SNAP and Work Among Low-Income Households*. en. Tech. rep. Center on Budget and Policy Priorities, p. 36.

Shaw, Justin and Neal H Hooker (2016). "A SNAP in Time: ABAWD Work Requirements". en. In: p. 31.

Stacy, Brian, Laura Tiehen, and David Marquardt (2018). *Using a Policy Index To Capture Trends and Differences in State Administration of USDA's Supplemental Nutrition Assistance Program*. en. Tech. rep. ERR-244. U.S. Department of Agriculture, Economic Research Service, p. 38.

Stavrianos, Michael and Lucia Nixon (1998). *The Effect of Welfare Reform on Able-Bodied Food Stamp Recipients*. en. Tech. rep. Alexandria, VA: Food and Nutrition Service, U.S. Department of Agriculture.

Strayer, Mark, Esa Eslami, and Joshua Leftin (2012). *Characteristics of Food Stamp Households: Fiscal Year 2011*. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

USDA (2015). *Guide to Serving ABAWDs Subject to Time-limited Participation*. en. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

Wolkwitz, Kari (2007). *Characteristics of Food Stamp Households: Fiscal Year 2006*. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

CHAPTER 2

THE EFFECT OF MEDICAID AND SNAP EXPANSIONS ON LABOR MARKET OUTCOMES AND SNAP PARTICIPATION

2.1 Introduction

Many households in the United States have struggled to overcome medical costs and secure access to affordable, nutritious foods. For example, in 2018, medical expenses were the primary reason behind the increase in the number of individuals in poverty (Fox, 2019) and over the same period, 12% of households were considered food insecure (Gundersen, 2019). The key policy mechanisms used to address these growing concerns are the Supplemental Nutrition Assistance Program (SNAP) and Medicaid. SNAP is the largest food assistance program in the U.S., over 40 million Americans currently enrolled (Canning and Stacy, 2019), while Medicaid provides insurance to 67 million individuals (CMS, 2020). Households that are 100% to 138% from the federal poverty line (FPL) are generally eligible for both SNAP and Medicaid. Approximately, between 6-7 million households are in this income range from 2013 to 2016 (Table 2.1).[1]

In the last decade, income requirements for these programs changed, allowing more individuals to enroll. Due to the Great Recession of 2008, the Affordable Care Act of 2010 (ACA) and the American Recovery and Reinvestment Act of 2009 (ARRA) increased these income thresholds to 138% and 200% for Medicaid and SNAP, respectively (Nord and Prell, 2009; Frean, Gruber, and Sommers, 2017). Prior to 2010, only households earning 100% (130%) of the federal poverty level (FPL) could participate in Medicaid (SNAP).[2] However, in 2012, the Supreme Court ruled in *Sebelius* that states do not have to participate in the Medicaid expansion provided by the ACA, as the states would have to pay a portion of the

[1] This table was computed using the Survey of Income and Program Participation (2013-2016) with survey weights.

[2] Households are generally able to participate in SNAP with 130% of the FPL, but the benefit amount may be small as they phase out with income. The Medicaid expansion allows households to earn above the FPL and receive insurance and SNAP benefits.

cost of the expansion (McMorrow et al., 2017; Rosenbaum, 2013).

As a consequence of the *Sebelius* ruling, only 24 states (plus the District of Columbia) implemented the Medicaid expansion when it first became available in January 2014. By the end of the 2014 SIPP panel (December 2016), 31 states had opted into the expansion. I use this variation in Medicaid expansion in combination with the variation in state SNAP policies[3] to study how these policy expansions affect labor outcomes and SNAP participation.

Both SNAP and Medicaid are need-based programs with income thresholds for participation, but the formulas for "clawing back," or decrease in benefits, are different across the programs. For SNAP benefits, participating households lose \$0.24[4] for each dollar earned in income provided that their net income is sufficiently high. For Medicaid, the benefit is whether households receive insurance. Earning above the income limit would make the household ineligible for Medicaid and the household would then qualify for subsidized health insurance through other insurance markets.

The clawing back of these benefits effectively increases marginal income tax rates for participating households, distorting the normal labor-leisure trade-off of a marginal hour's worth of labor and its corresponding wage. Furthermore, since SNAP and Medicaid both have economic value, it may behoove households to strategically choose incomes to maximize a combination of income, leisure, SNAP benefits and health insurance. However, to manipulate their position on any "kinks" the household would both have to understand the rules that govern the programs as well as be able to manipulate their incomes to choose their optimal bundles, leading to interesting labor market questions.

This paper examines the effect of the SNAP and Medicaid expansion policies and the interaction of the two on labor market outcomes. Using panel data (January 2013-December 2016) from the 2014 Survey of Income and Program Participation (SIPP), I study 1) if household labor market outcomes (income, hours worked and employment status) change

[3] States may choose simplified or more generous ways of calculating net income and/or re-certifying SNAP eligibility. See section 2.2.2 for more details.

[4] Households are expected to spend 30% of their net income on food, up to the estimated cost for food of a household their size. Only \$0.80 of each dollar from wage income is counted toward net income.

as Medicaid and SNAP expansions made certain income levels more desirable,[5] 2) if SNAP participation changed in response to states opting out of Medicaid expansion and 3) whether state SNAP policy options, aside from expansion of the gross income limit, affected SNAP participation differently with the Medicaid expansion.

It is important to study these welfare expansions jointly, as these two programs serve overlapping households. In addition, analyzing these expansions in isolation ignores some of the interactions between the policies, especially as some states have made households receiving insurance through Medicaid automatically eligible for SNAP. A rational household, with full information of program rules and requirements, will simultaneously decide on labor, insurance and program participation. The focus on state policy options is important for understanding the linkages in the social safety net and in predicting changes in participation in these programs as the policies change.

The results from the analysis suggest that the interaction between the Medicaid and SNAP expansions has no statistically significant effect on household labor outcomes. I find no evidence for the influences of these policies on measures of head of household's employment and hours worked, non-head employment, non-head hours and household income. I do find that the Medicaid expansion's impact on SNAP participation is positive, but only in areas that do not have generous SNAP policies. That is, the negative and statistically significant interaction coefficient nullifies the effect of the Medicaid expansion. These findings overall reveal the importance of jointly analyzing policy expansions and changes.

This paper contributes to the literature in three ways. First, SNAP is a federally funded program where the states are granted flexibility with how they administer the program. Earlier studies looking at the Medicaid expansion and SNAP ignore many of these details, nor have they investigated the demographics of which households are most likely affected by these two programs changing over the last ten years. In this paper, I include state specific

[5] Desirable income in this context refers to income levels that would grant program eligibility. With the expansions in place, a household could increase their hours work and therefore their income without losing program eligibility. Similarly, if a household is near the income cutoff, a slight decrease in income could result in program eligibility.

policies such as changes to the federal asset tests, broad-based categorical eligibility and simplified applications and re-certifications of eligibility. Further, I study the interaction of these SNAP policies and Medicaid expansion, as a household would have to live in a state with both Medicaid and SNAP income expanded in order to be eligible for both programs.

Second, this is the first study exploring the interaction of SNAP and Medicaid using the SIPP. Many papers rely on the Current Population Survey (CPS) -Economic supplement and/or look at other annualized measures of income and participation (Han, 2020; Burney, Boehm, and Lopez, 2018; Kaestner et al., 2017). SNAP is a monthly program and analysis of annual income may under-report months of SNAP eligibility. Further, the panel nature of SIPP provides informative glimpses into household assets, child support payments and other important information for SNAP eligibility and participation. It is important to note that SIPP respondents under-report SNAP participation, but Meyer, Mittag, and George (2020) finds that this happens at a lower rate in the SIPP compared to the American Community Survey (ACS) and the CPS.[6]

Lastly, existing literature does not investigate whether households are changing labor decisions in response to the Medicaid expansion and SNAP income limits, nor do they test across different members of the household. In this paper, I explore these questions and also, I empirically test whether participation in these programs is a plausible explanation for any changes in self-reported income.

All of these considerations are important to understand as the social safety net in the United States continues to change. For instance, Figure 2.1 shows the states that have expanded Medicaid under the ACA. Since the 2014 SIPP Panel, four additional states have expanded Medicaid and Missouri voted in 2020 to opt-in to the expansion for 2021. Despite the upward trend in state's take-up in Medicaid, provisions of the ACA face many legal and political battles that threaten this very policy.

[6]Specifically, 23% of true SNAP recipients do not report participation in the SIPP, compared to 35% and 50% in the ACS and CPS, respectively. False positives in all of these surveys make estimates of SNAP participation less attenuated.

The remainder of the paper is organized as follows: Section 2.2 provides more policy details and policy context for SNAP and Medicaid, section 2.3 discusses related literature and section 2.4 provides details on the 2014 SIPP panel and other data sources. The model (Section 2.5) is followed by results, discussion and paper conclusion.

2.2 Policy Background

2.2.1 Medicaid

Medicaid is jointly funded by states and the federal government, but is administered by the states. It primarily provides health insurance to low income households, with zero insurance premiums. The state Medicaid programs may or may not have other medical expenses related to consuming health care, such as co-payments, office fees and deductibles. Individuals and families[7] must apply through their state's application system to receive insurance through Medicaid. Upon meeting the eligibility criteria, the applicants are then given a list of health insurance plans to choose from to best suit their needs, though the details may vary by state. Households generally must apply to receive Medicaid and some states have imposed work requirements and additional income limits for receiving insurance funded through Medicaid.

Under the Affordable Care Act (ACA), households earning between 100% and 400% FPL[8] are eligible for health insurance subsidies provided they are neither eligible for Medicaid nor able to acquire insurance through their employer. This subsidy removes the "cliff" from earning just above the Medicaid income limit and having to pay for the full cost of health insurance. The subsidy amount is based on the "Second Lowest Cost Silver Plan" (SLCSP) available to the particular household's health insurance market.The ACA created four metallic categories for health insurance: bronze, silver, gold and platinum. Each metal tier has to

[7]A spouse and/or any dependent children.

[8]This range is for states that did not expand Medicaid. For states that expanded, individuals are eligible for a subsidy if household earnings are between 138% and 400% FPL. Wisconsin opted to increase Medicaid coverage up to the federal poverty line, removing previous restrictions, while some states kept Medicaid limits below 100% FPL, such as Texas.

satisfy a set of minimum requirements pertaining to the expected percentage of health care costs being covered by the plan as opposed to out of pocket. The subsidy is either received as a refundable tax credit or the estimated credit is paid directly to the insurer, reducing the health insurance bill for the household.

Households are expected to contribute a slightly larger share of their income on health insurance the more they earn. For instance, within 2020, households in 100-133% FPL would have to pay no more than 2.06% of their income on the SLCSP in their area. The difference between SLCSP and that percentage of their income is the subsidy amount. Households earning above 300% FPL are capped at paying no more than 9.78% of their income for SLCSP.

2.2.2 SNAP

SNAP is mostly funded by the federal government through the United States Department of Agriculture's budget and is administered at the state level.[9] SNAP benefits are transfers from the government to an electronic benefits transfer (EBT) card, similar to a debit card. Households use the EBT card to purchase most kinds of food (not prepared for immediate consumption) from participating markets such as grocery stores, convenience stores and farmer's markets.

Household benefits are based on income as well as family size. The maximum food stamp benefit amount is determined by the cost of the Thrifty Food Plan, the most affordable of four nutritionally adequate diets created by the U.S. Department of Agriculture (Carlson et al., 2007). SNAP allows for all households to be able to afford the Thrifty Meal Plan for their household size, with the expectation that households will spend at least 30% of their net income on food. The maximum benefit also depends on the size of the household, accounting for economies of scale in grocery shopping. Table 1.4 displays the maximum benefit amounts from 2004-2017 for household sizes 1-4.

[9] Though as of 2020, several states administer the program at a county level and many other use regional offices to facilitate program administration.

There are three means tests that households must satisfy to receive food stamp benefits: gross income, net income and assets (Ganong and Liebman, 2018). Both income tests refer to the federal poverty levels (FPL) from the previous fiscal year, with the thresholds being 130% and 100% respectively for the gross and net income tests.

Generally, the gross income standard includes all forms of income aside from small, infrequent receipts such as tax returns or bonuses. The net income test is more complex, as deductions for earned income, child support, dependent care, excess housing[10] and medical deductions[11], in addition to a standard deduction are involved. To determine program eligibility, net income is determined by adjusted gross income, net of child support, medical deductions and excess housing.

While states do not have the option to change the benefit calculations with federal funding, they are able to change the eligibility criteria for SNAP participation. The set of policy options have grown over time, with states expanding the types of households eligible to participate in SNAP.

The first drastic change to SNAP eligibility by a state was the option to modify or completely remove the household asset tests. Generally, households with more than a few thousand dollars in bank accounts, or a high dollar value in automobile assets were ineligible for SNAP. Most states have exempted the value of at least one vehicle per adult in the household and many have stopped using the asset test altogether.

Perhaps the largest change to eligibility in SNAP was the introduction of Broad-Based Categorical Eligibility (BBCE). BBCE allowed states to use Temporary Assistance to Needy Family (TANF) block grants to fund SNAP participants at higher levels of net income. States were able to raise the gross income limit from 130% FPL to as high as 200%. Further, BBCE allowed for households eligible for other welfare programs to be automatically qualify for SNAP.

[10] Housing cost deduction is for the difference between actual housing costs and half of the Adjusted Gross Income, subject to a maximum amount.

[11] Medical deductions are for elderly only and applicable to spending beyond $35 per month.

Other important state policy options include the transition to a debit card, referred to as EBT, increasing the amount of months between re-certification of eligibility, expansion of job training programs and other opportunities to satisfy the work requirement while not being employed in the formal labor market and the general simplification of calculating and reporting net income for a household. Generally, states have become more generous with respect to SNAP policy options to make more households eligible to participate.

2.2.3 The interaction of SNAP and Medicaid

Medicaid and SNAP are linked in two substantial ways. First, they are both designed to assist households near or below the poverty line and for both programs, the upper income limits that determine eligibility are between 100% and 200% FPL. Second, some states have simplified applications for welfare programs, including SNAP and Medicaid. The applicants are either automatically enrolled for one program if they qualify for the other, or some of the information is used to begin the application of the other program.

The Medicaid expansion allows participating households to earn additional income while still maintaining insurance through Medicaid, just beyond the gross income limit for SNAP participation. Expansions to SNAP that increase the gross income test or relax asset tests also make SNAP accessible to more households.

The interactions of the welfare programs may have the largest considerations for whether the second adult in a household supplies labor and if they do, how many hours they work. If a household can have adequate health insurance, food and shelter provisions from a single source of income, the additional wages of a second person make the household earn too much income to qualify for SNAP and/or Medicaid. For households with small children, the second job may then require enrolling their children in child care, which may result in less disposable income.

To illustrate how these programs can jointly affect labor decisions, I will consider two hypothetical households, both with four family members in 2013, the last year before the ACA's Medicaid expansion provision took effect. The federal poverty guideline was $23,550 for a family of this size.[12] The income cutoffs were $32,499 in 2014 Medicaid expansion states and $30,615 for SNAP under the federal cutoff of 130% FPL.

A family at 125% FPL with household income of $29,440 living in a Medicaid expansion state and paying for health insurance out of their own pocket is then able to qualify for Medicaid. This can decrease medical expenses in multiple ways, including decreasing premiums, co-pays and coinsurance payments. A typical household could receive $200 in SNAP benefits per month, with $0.24 being clawed back for every additional dollar earned.[13] This household faces a higher marginal tax rate from clawed back benefits up until their benefits become zero or their income is sufficiently high for them to fail a SNAP income test.

A second household may earn at 150% of the FPL ($35,325 annually, or on average $2,944 per month). This household is neither eligible for SNAP nor Medicaid, unless their state used BBCE to expand the gross income limit. However, the household may stand to benefit from strategically reducing their income to qualify for both programs. Earning $3,000 less would place the household at 130% FPL. This decrease in monthly income ($250) can be made up in SNAP benefits in circumstances similar to the household previously discussed. Additionally, the household would be eligible for Medicaid.

The importance of the interaction of these policies are two-fold. First, SNAP claw backs add on a relatively high marginal tax rate in addition to normal income and payroll taxes. Reducing income to take advantage of Medicaid can be compensated by the SNAP benefits. Second, application to SNAP may increase awareness of Medicaid and vice-versa.

[12] This number is for the 48 contiguous states and the District of Columbia. Alaska and Hawaii use different guidelines and FPL are not defined for U.S. territories.

[13] The maximum benefit amount was $640 after the American Recovery and Reinvestment Act's increase of the benefits expired. The average SNAP net income deductions for households with countable income was $515(Gray and Kochhar, 2015). The benefit is then $640 - 0.3 * (0.8 * 2450 - 515) = 206.5$ Where $2,450 is monthly gross income and 0.8 is the policy that forgives 80% of earned income and 0.3 is the proportion of net income a household is expected to contribute for their own grocery purchases.

2.3 Related Literature

Several papers look at the relationship between the SNAP, Medicaid expansion and welfare program participation. Overall, the findings have been consistent across studies, but some studies have not analyzed both SNAP and Medicaid expansions jointly.

Both Schmidt, Shore-Sheppard, and Watson (2019) and Lanese, Fischbein, and Furda (2018) found evidence suggesting Medicaid expansion lead to an increase in SNAP participation. Schmidt, Shore-Sheppard, and Watson (2019) look at EITC and SNAP take-up from the Medicaid expansion using a county-pair analysis to compare counties from states with and without the expansion with respect to county caseload. They use administrative tax data for EITC and USDA Food and Nutrition Service (FNS) National Data Bank data for biannual SNAP caseload information for the counties. Not all states had county level data available. They find Medicaid expansion leads to a 4% increase in SNAP participation, or 0.6 percentage points in their data set which had 15% of the sample participating in SNAP. Like other studies, a limitation of this study is that the data set used contains annual measures of income and SNAP participation. The eligibility criteria are based on resources available for a particular month, not yearly.

Lanese, Fischbein, and Furda (2018) analyzes the American Community Survey (ACS), the Annual Social and Economic Supplement (ASEC) to the Current Population Survey (CPS) from 2011 to 2016, small area income and poverty estimates and total state caseloads of SNAP. The ASEC of the CPS provides poverty and income annual estimates based on a survey of roughly 75,000 households. Using a DID estimation, the authors find that Medicaid expansion states increases SNAP caseload by 10%. The paper only considered expansion of SNAP in terms of outreach, not changes to program income and asset limits.

Conversely, Han (2020) looks at the effect of SNAP expansions on Medicaid enrollment. The author explores the impact of the SNAP expansion on SNAP participation, WIC participation, private insurance and Medicaid takeup. The results show the SNAP expansion to correlate with a loss of private insurance, but no increase in Medicaid participation. Han

(2020) used variation in SNAP policies and multiple panels of SIPP to estimate how take-up in SNAP affected the other programs. A potential shortcoming is that it selects on income for limiting the regression population as opposed to selecting on education of the household head.

2.4 Data

2.4.1 2014 Survey of Income and Program Participation (SIPP)

This paper uses the 2014 Survey of Income and Program Participation (SIPP), a panel data set spanning from January 2013 to December 2016. The SIPP is a long-running survey of U.S. households that over-samples households in lower income areas and asks many detailed questions about assets, income and participation in government sponsored welfare programs. The 2014 SIPP followed households across four waves, where in each of those four years they were interviewed about income, health, assets and program participation for the previous calendar year. The United States Census Bureau conducts and manages the SIPP, with the interviews for the 2014 panel occurring between February and April each year. The SIPP tracks household members as they move and form new households, although some exit the SIPP for various reasons. Thus, the 2014 SIPP is an unbalanced panel.

A unique advantage of using the SIPP over other large surveys is the SIPP asks about monthly changes to income and program participation. When limiting questions to annual income, a researcher may conclude that a household was ineligible to participate in SNAP. However, when analyzing monthly income, the household may have been eligible for several months during the year due to a reduction of earnings, such as a spell of joblessness.

The 2014 SIPP was significantly revised from the format of the 1996, 2001 and 2004 panels. In the previous panels, the wave was four months long and households were interviewed about the last four months (Table 2.2). Further, the interviews within each wave were staggered so some households were interviewed every month. These changes were made in part to ease respondent burden and cost of administration (US Census Bureau - Economic and

Statistics Administration, 2019). The revision to the 2014 SIPP also gives more detailed information about household relationships, allowing for more flexibility in defining households to approximate different criteria.

The previous SIPP panels experienced a "seam bias" (Moore, 2008) where respondents were more likely to report changes in employment or program participation between interviews and not within the four months in which they were asked about during the survey. The new format has a 12-month recall, which potentially changes the errors in recall. Notably, the earliest month they are asked about was over 12 months prior to the interview date.

The new format of recall over the previous year uses an "event-history calendar." This system asks the respondent to list notable dates that occurred over the calendar and uses those dates to aid them in recall of what their circumstances were for the reference months. The results from field testing in 2011 and 2012 suggest that the 2014 SIPP data better matches administrative records (US Census Bureau, 2013).

As SIPP questions do not directly determine SNAP household eligibility, I estimate models using the SNAP definition for household unit. In this paper, the household is defined as all persons living in the home. I exclude any households that have more than two adults aged 18-64 years old, households with more than 8 members,[14] and households for which the head of household is aged 65 or older, as they categorically eligible for Medicare and not subject to similar rules for participation.[15] Additionally, I limit the sample to households whose head of household has a high school diploma or less education, as they are most likely to be affected by changes in welfare policy and comprise 56% of the sample earning at or less than 200% FPL despite accounting for 39% of households. Lastly, headship in this paper is self reported, based on the respondent's answer to the head of household question. I define the secondary wage earner as the adult who is not the head of household in a household with only two adults.

[14] I remove populations where the household size is above 8 to remove other group housing scenarios that are generally different than the unit of interest.

[15] In addition, the requirements for SNAP eligibility for senior citizens are different than the requirements for adults aged 18-64.

In addition to those household constructions, I also restrict my estimation to the sample of households that did not move out of state during the 2014 SIPP. This accomplishes two things: first, it alleviates concerns over households moving in response to state policy options as well as allow for clustering of standard errors at the state level in unobserved heterogeneity models.

For computational efficiency, survey weights are not used and thus, the estimation is based on a sample that is not reflective of the U.S. population as a whole. I summarize descriptive variables for this subsample in Table 2.4.

The proportion of households participating in SNAP and Medicaid can be observed in Figure 2.2, where I calculated the rates across all states that either expanded Medicaid as of January, 2014 (i.e. the first month they were allowed to do so under the ACA) or did not expand Medicaid through the 2014 SIPP coverage period. For the "always expanded" Medicaid states, the proportion of households in SNAP and Medicaid were essentially the same at 12%, then Medicaid coverage increased year after year while SNAP receipt remained flat. Both of these facts are notable when considering that the national unemployment rate dropped from 8.0% to 4.7% during the 2014 SIPP coverage period.

The states that did not expand Medicaid had a higher proportion of households receiving SNAP benefits for the 2014 SIPP. The proportion of the population receiving Medicaid was below 10% in 2013 and gradually rose to meet the proportion receiving SNAP benefits. One reason why Medicaid usage may have increased in the non-expansion states is because they may have used the ACA as an opportunity to eliminate previous work and income requirements for households earning less than the federal poverty line.

These participation rates are simply the ratio of the number receiving benefits to total persons in that group. That is, it does not take into consideration the proportion of the population being eligible for these programs, therefore it does not say anything about take-up across those states.

2.4.2 Other Data

I obtained publicly available health insurance plan data through the Center for Medicare and Medicaid Services. The data published included information on premiums, county coverage, metal level of the plan[16] and effective dates for those prices. The premiums offered generally depend on age of the insured, number of persons on the plan and whether the insured uses tobacco products. To calculate the SLCSP premium, I determine the second-lowest cost silver plan for each county and then take a population-weighted average to arrive at a singular number for the state.

I use and extend the U.S. Department of Agriculture (USDA) Economic Research Service's (ERS) State SNAP Policy Database and Index as a measure of how inclusive the state's policies are for SNAP eligibility. The SNAP Policy Index includes weighted and unweighted rankings of the importance of different policies to program participation. These policies, which include eligibility criteria (income and asset tests), program outreach to nonparticipating households and transaction costs associated with enrolling and maintaining benefits, differ across states (Stacy, Tiehen, and Marquardt, 2018). The ERS policy index was constructed via estimation on how each policy option influences participation in SNAP. As noted by the authors, by 2011, many states have already adopted most of the policies included in the index. Thus the identification of the specific policies may be difficult to disentangle from state fixed effects.

The ERS does not have complete data on the underlying variables appearing in the SNAP policy index. This may be because the reports from which the ERS collects the data no longer track each of the original policies in a level of detail sufficient for the index. Where possible I found dates for when states may have changed their policies and alson inferred no change. A summary of the variables for this paper can be found in Table 2.5.

I extend the ERS' policy index by separating the broad-based categorical eligibility

[16] There are four metal tiers of health insurance plans, Bronze, Silver, Gold and Platinum. The higher the tier, the better the coverage in terms of minimizing out-of-pocket medical expenses. Platinum plans are expected to pay 90+% of health care costs and the bronze at least 60%.

(BBCE) policy based on whether it is used to expand the SNAP eligibility income limit, remove or relax an asset test, or make changes to both. The separate record keeping is especially important as the default income limit for SNAP eligibility is 130% FPL, which is above the non-expanded Medicaid income limit, but below the Medicaid expanded limit. During the 2014 SIPP time period, California and Illinois expanded their SNAP income limits through BBCE, while 23 states had not expanded income limits through 2016.

In the construction of my modified SNAP Policy Index, I follow the ERS' unweighted index. I award one point for each of the six state policy options used by the state for that period. This index includes income expansion, vehicle exclusion from assets, simplified reporting, online applications, EBT issuance and not requiring fingerprints to be filed for participation. This index ranges from 0 to 6; however, from 2013 to 2016, the minimum value is three policies. For ease of interpretation, I then transformed the index so that 0 represented the most restrictive policy set in the sample and 1 the most generous.

Lastly, I collect state unemployment statistics and minimum wage data from Bureau of Labor Statistics via the St. Louis Federal Reserve Bank's *FRED*.

2.5 Model

I examine the effect of SNAP expansion, Medicaid expansion and the interaction of the two on various employment related outcomes. Specifically, my main outcomes of interest are (1) income as a percentage of the federal poverty level (FPL), (2) hours worked by the head of household and the second adult in the household, (3) employment status of the head and second adult and (4) SNAP participation. These outcomes reveal different insights about how expansion policies influence labor market decisions. First, there are different ways in which a household may change their income in response to incentives. To qualify for SNAP, the household must satisfy certain income requirements, but the combination of hours worked[17] and wages are not restricted. Thus, households could choose a different set

[17] Some households, such as able-bodied adults without dependents, are subject to an hours "worked" requirement.

of job opportunities with different hours, wages and non-wage benefits in order to change their income.

Second, examining hours worked reveals whether employment changes on the intensive margin with respect to the program incentives. For the first estimation, I use the total reported hours worked by the household head for a given month. For the second estimation, I model the adult non-head of household's hours worked as a function of the household head's hours. This is similar to the "second wage earner" model, where the primary wage earner is likely to supply their labor inelastically as a means to provide for the household. However, I use the SIPP's designation for head of household,[18] which is not necessarily the same as the person with the higher earnings potential or higher propensity to participate in the formal labor market. In this model, the second wage earner is deciding their labor market participation while taking as given welfare policies, earnings of the head, temporal economic conditions and household composition.

Households may change their employment status to adjust their income into the desired range. A secondary wage earner may reduce their hours to suppress income, or conversely, take on more employment to increase it. Households earning above the income limits for SNAP and Medicaid may be inclined to reduce their earnings in exchange for better job attributes or fringe benefits. That is, they may opt for a lower hourly wage for a more enjoyable work experience.

The availability of food and heath insurance benefits to higher levels of income could increase the amount of search time for employment of a second adult in the family. This is because the benefits from the government relax some of the budget constraints a family may have in the short run. For this reason, coupled with the fact that the difference between 100 and 138 FPL is small relative to one person's annual employment at a modest wage rate,[19] I expect that households in these ranges would spend more time in-between jobs for the

[18] In the SIPP, respondents self-identify as the head of household. I designate the second of two adults in the household (non-head) as second earner.

[19] In 2020 for a family of 4, the difference is \$9,956. This is roughly 20 hours a week at \$11/hour.

second earner. Such a finding would be consistent with Garthwaite, Gross, and Notowidigdo (2014) that find that improved access to insurance outside of employment leads to decreased employment.

Lastly, examining responses to changing eligibility standards for SNAP participation allows me to check if participation in welfare programs could plausibly explain why households would change their income. That is, if households manipulate their income in order to be eligible for multiple programs, then participation in those programs should increase as a result. The model for these outcome variables is as follows:

$$\begin{aligned} y_{ist} =& \alpha + ExpandS_{st}\beta_1 + ExpandM_{st}\beta_2 + ExpandS_{st} * ExpandM_{st}\beta_3 \\ &+ X_{ist}\gamma + \Sigma_s\eta_{st}S_s + \Sigma_s\tau_{0s}ExpandM_{st}T_tS_s + \Sigma_s\tau_{1s}(1 - ExpandM_{st})T_tS_s \\ &+ \Sigma_t\lambda_tT_t + \epsilon_{ist} \end{aligned} \tag{2.1}$$

Where i, s and t index head of household, state and time, respectively and y_{it} refers to one of the outcomes of interest. For the time effects, t can either index the year and month of the observation or quarter depending on the specification of the model estimated. State-specific time trends are also used in all models. These allow for each state to have their own specific slope to reflect gradually changing conditions over time. The time trends are interacted with the Medicaid expansion so that the outcomes may experience a discrete jump or change the time trend when the major policy shock occurs.

The βs are the coefficients of interest. $ExpandS$ and $ExpandM$ take the value of 1 in states during the time periods they have expanded the SNAP and Medicaid income limits, respectively. β_3 is the coefficient on the interaction of the policies, indicating the difference in the outcome in states where both expansions are in effect.

Household-level controls (X_{ist}) include age of the head of household, marital status, family size, veteran status, race and ethnicity, whether the household is in an urban area and highest educational attainment inside the household. I provide summary statistics for these variables in Table 2.4.

State-level controls include variables to capture economic conditions and state policies.

The economic variables include the unemployment rate, minimum wage and the second lowest cost of a silver health insurance plan. The policy variables are either summarized in the ERS' State Policy Index or as a vector of policy variables pertinent to SNAP participation.

The coefficients for state-specific intercepts are represented by η_s. I include state-specific linear time trends (coefficients $\tau_0 s$ and $\tau_1 s$), where 0 or 1 reflects the Medicaid expansion policy for time period t. That is, the slope of the time trend was allowed to vary across Medicaid expansion within a state. Time-specific intercepts are represented by T_t with coefficients λ_t.

I also estimate modifications to the model above using correlated random effects (CRE) framework. The models are modified such that the error term is separated into a unique household intercept and a stochastic part. In the estimation, the means of the X_{ist} variables are included to allow for the individual heterogeneity parameter to be correlated with time-varying covariates. This differs from fixed effects estimation by restricting the correlation between the unobserved household heterogeneity to be linear in the means of the individual characteristics, whereas fixed effects is less restricted.

The inclusion of the time averages allows for a test on the random effects assumption that the error terms are uncorrelated with the time-variant individual variables. A rejection of the null of joint insignificance of the coefficients on the time averages is sufficient to say the random effects assumption does not hold. This implies that the fixed effects estimator is more appropriate (Wooldridge, 2019).

The estimation model is as below, but the interpretation is the same as the model above - there is no interpretation of the coefficients on the time averages, ι.

The model for these outcome variables is as follows:

$$\begin{aligned} y_{ist} =& \alpha + ExpandS_{st}\beta_1 + ExpandM_{st}\beta_2 + ExpandS_{st} * ExpandM_{st}\beta_3 \\ &+ X_{ist}\gamma + \Sigma_s \eta_{st} S_s + \Sigma_s \tau_{0s} ExpandM_{st} T_t S_s + \Sigma_s \tau_{1s}(1 - ExpandM_{st}) T_t S_s \\ &+ \Sigma_t \lambda_t T_t + \iota \bar{X}_i + u_{ist} \end{aligned} \tag{2.2}$$

Where $\bar{X}$ is a vector of household time-averaged characteristics, $\bar{X} = \Sigma_t X_i t$. The model only includes the time periods for which the household responded to the SIPP survey. The error term, u_{ist} in (2) is constructed to be uncorrelated with the other regressors, whereas the error term in (1) can be written to include these time means of the household characteristics which are omitted in that model.

For income, I estimate the above model with pooled OLS and CRE with state- and time-fixed effects. I limit the estimation sample to households that did not move out of state in the 2014 SIPP to be able to cluster standard errors at the state-level in all estimations.

For binary response variables, e.g. income within a range and SNAP participation, I estimate a linear probability model and a logistic regression (logit). As with the other linear models, I include state- and time- fixed effects and state clustered standard errors.

When estimating expansion effects on SNAP participation, I add SIPP responses to food security questions as control variables. Some argue that SNAP benefits amounts are infra-marginal, so that benefits can be treated similar to receipt of cash (Hoynes and Schanzenbach, 2009). However, households facing higher levels of food insecurity for several months have a higher propensity to join SNAP and some measures of food security remain low for households receiving SNAP benefits (Nord and Golla, 2009).

2.6 Results

I report the estimation results below. Overall, I find no evidence to suggest that households strategically manipulate their income to secure eligibility for SNAP and Medicaid receipt. However, I do find that the Medicaid expansion increased SNAP participation in states with the least generous state SNAP policy options.

2.6.1 Household Income

I estimate the effect of SNAP and Medicaid expansion on continuous measures of household income using pooled OLS with state fixed effects. I use two different measures of the SNAP expansion. The first measure is a modified version of the unweighted ERS Policy Index. The second measure is a binary variable for whether the gross income limit was raised from the default 130% FPL.

The results for household income and log income using the novel index discussed above are found in Tables 2.6 and 2.7. Column 2 in both tables show the results of the CRE estimation. The results of the F-tests for joint significance of the $\bar{X}_i$ coefficients are reported at the bottom of the tables. The null hypothesis that they are jointly insignificant is rejected at any reasonable level of confidence. This result provides evidence that the random effects assumption, that the household unobserved heterogeneity is uncorrelated with the X_{it} variables, is not reasonable. Therefore, I use fixed effects estimation in this paper over random effects (Column 3).

I find no consistent evidence the interaction of the SNAP policy index and Medicaid expansion is influencing both income and log income. In Table 2.6, the coefficient on the interaction term is positive and significant, suggesting that household incomes are higher in states where both policies are expanded. However, the coefficient on the Medicaid expansion is large and oppositely signed. Taken at face value, the 38% of the FPL increase from the interaction of the policies seems implausibly large, especially when logged income (table 2.7 does not provide evidence to support this finding.

The analogous results with SNAP income expansion as the policy change are in Table 2.8. This measure of SNAP expansion varies less within and across states through time than the SNAP policy index. In contrast to the estimations with the novel index, the interaction of the policies does not have a significant effect on household income. As the examination of income alone does not provide a consistent story, I study the components of income, specifically employment status and hours worked.

2.6.2 Hours Worked

For the remainder of the paper, I use the novel modified SNAP policy index as a measure of SNAP expansion, as opposed to a single indicator for a particular policy option. The use of the index allows for slightly more variation in SNAP administration across time than the use of one particular SNAP policy.

I use POLS, CRE and FE to examine the effect of welfare program expansions on number of hours worked for both the the head of household and the non-head of household.[20]Using OLS estimation with clustered standard errors, I estimate the effect of welfare program expansions on number of hours worked for the head of household and, for households with two adults, the hours worked of the non-head of household.

The results for head of household's hours worked are in Table 2.9. The estimations provide evidence that the number of hours worked do not depend on the separate program expansion polices, or the interaction of the two. The finding of an insignificant effect of Medicaid expansion on labor supply is consistent with the findings in Kaestner et al. (2017).

Further, for log hours worked (Table 2.10) I find a no statistical evidence of a effect of the expansions on hours worked for the head of households. The point estimate of +4% would suggest that if the effect were statistically significant, the economic effect would be small. With a mean of 100 hours of work per month, that would imply an additional 4 hours of work per month.

The results for the non-head of household are not directly comparable to the head of household estimation, as the population is further restricted to households with only two adults aged 18-64.[21] Additionally, I include the number of hours worked as well as employment status of the head of household as controls. That is, the results (Tables 2.11 and 2.12) are conditional on whatever the employment status of the head of household is for that month.

I find that interaction of both welfare expansions is negatively signed, though statistically indistinguishable from zero. The consistency in signs of the coefficient of interest leaves open the possibility that a small effect may exist for some sub-populations, however further investigation would be needed to bolster this claim.

[20]The estimation of the non-head regressions is restricted to households with two adults.

[21]The number of children is controlled not restricted (included in the control variables).

2.6.3 Employment Status

I examine employment status of both the head of household (Table 2.13) and the second non-head adult in the household (Table 2.14). I estimate three variations of the linear probability model (LPM). These variations are pooled OLS, CRE and FE. The rationale for estimating the LPMs is that the coefficients, under certain assumptions, consistently estimate the average partial effects from non-linear models. For the non-linear models, I estimate logit and CRE logit. I do not estimate FE logit as it is computationally expensive to do so.

I find no evidence that the program expansions positively increase the probability of the household head being employed. The CRE and FE show the interaction coefficient does not have a significant effect on employment status. Additionally, neither the pooled OLS, the logit coefficients or average partial effects are statistically significant for either the policies or the interaction. The one exception is the APE for logit estimation, with an effect of +1.1 percentage points, or nearly 2% increase in employment probability. This is because the interaction term is estimated to be small and negative, but the coefficient on the Medicaid expansion is positive and larger in magnitude. The finding is not supported in other estimations of this variable with the same sets of controls.

Lastly, I find no significant effects for the second adult's employment status. The findings of small to no policy effects on employment status are consistent with theoretical models of household labor supply, where labor may be supplied inelastically (Ashenfelter and Heckman, 1974).

2.6.4 SNAP Participation

The results for the binary variable of SNAP participation are reported in Table 2.15. Overall, households were less likely to participate if they were white or Hispanic and more likely to participate in SNAP if they reported higher levels of food insecurity and had more children living in the household.

I find that the SNAP expansion increased the likelihood of a household participating in SNAP when controlling for other factors. The estimate is +0.3 percentage points for the household fixed effects estimation, which corresponds to a 14% increase in SNAP for this population, which has a mean take-up rate of 22%. The Medicaid expansion appears to have increased participation in SNAP. However, the magnitude is similar to the magnitude of the interaction term and those coefficients are oppositely signed. This means that while the Medicaid expansion is associated with an increase in SNAP participation, the net effect of the expansion is zero in states with generous state SNAP policy options. In other terms, states with less generous SNAP policies may see an increase in SNAP participation if and when Medicaid expansion would occur in their state.

The insignificance of the SNAP index on SNAP participation in some estimations could potentially be an artefact of how few states changed their options during the 2014 SIPP panel: Only one state added BBCE expansion, though changes on the intensive margin for some of the policies did change during this time period. Those changes were not captured in this index. Further, most states have added the policy options discussed in the ERS' SNAP Policy index by 2014, with the state average increasing from 6.16 out of 10 in 1996 to 9.23 in 2014 (Stacy, Tiehen, and Marquardt, 2018).

2.7 Discussion

The interactions of welfare policies on household labor market decisions are meaningful, yet with varying rules and benefit amounts across households and across states, the study of these interactions can be complex.

This paper studies the interaction of SNAP and Medicaid expansions and ignores the presence of other programs, such as the EITC that changes financial incentives and WIC, which provides food resources for some households. Although the interaction of these two programs has important considerations, the estimations showed mixed evidence of the statistical significance of the interaction coefficient.

This paper has a couple of limitations. First, income recall is a potential issue of the SIPP. A possible extension for this work to alleviate concerns over income recall would be to access administrative records on income from the Social Security Administration. Further, linking these data with county of residence through the private-use SIPP would allow for better integration of information on local economic conditions as well as market rates for health insurance costs for families obtaining insurance neither from employment or Medicaid.

Another limitation of this study is the salience of the changes to eligibility may not be well-known for households, especially as the complexity of those rules increases. It may take time for the changes in eligibility to be common knowledge. Further, a lag may exists from when a household decides the optimal number of jobs and hours in a household from obtaining the bundle of employment and programs they wish to consume, if they are able to choose their hours or income at all. Additionally, the SIPP will identify ex-post eligibility for SNAP and Medicaid. Studying the spells of unemployment and welfare program receipt may yield different results, as the lags from recognizing eligibility, applying to programs and receiving benefits could result substantially different outcomes than what is reported in the SIPP.

Lastly, further improving the SNAP policy index, following the approach of Dickert-Conlin et al. (2020) may better capture the variations that still exist in SNAP administration across states today. Many of the important factors for reducing stigma for participation (e.g. no fingerprinting, electronic receipt of benefits) and remaining in the program (e.g. more time before re-certifying eligibility) have been implemented by many states. But the ease of navigating and participating in the program may depend on other factors, such as marketing and access to stores that accept EBT for payment.

2.8 Conclusion

The data in this study cover the United States transition into full employment after a long and deep recession that began four years prior to the study. While unemployment rates decreased from 8.7% to below 5%, the number of Americans on Medicaid increased substantially and the percentage of households receiving SNAP benefits in the SIPP remained roughly constant. This paper examines whether the policy options, more specifically the general expansions of program eligibility, affects household labor outcomes and SNAP participation.

This paper uses data from the 2014 Survey of Income and Program Participation (SIPP) to examine whether the interaction of a state SNAP policies and Medicaid expansion affected the income and program participation of households where the head of household does not have a post-secondary degree. For measuring state SNAP policies, I used a novel policy index as well as a binary indicator for whether the state expanded the gross income limit. I find limited evidence that the interaction of these policies has an effect on the outcomes of interest.

For employment, I find no evidence that expansion of the programs influence the head's supply of labor in terms of hours and employment status. Additionally, I found no evidence that the adult non-head of household in a household has a high school diploma or less decreases labor supply on either the extensive or intensive margins when Medicaid and SNAP programs are expanded to include households with higher incomes.

Lastly, I find that Medicaid expansion is associated with an increase (4.7 percentage points, or 21% increase) in reported SNAP participation, but only in states that do not have generous state SNAP policy options already in place. In those states with generous SNAP polices, the net effect of Medicaid expansion is roughly zero. Further study of the interactions of welfare programs is warranted and the use of restricted SIPP data would increase the confidence in these results.

APPENDICES

APPENDIX 2A: Tables

Table 2.1: Number of Households in the 100% to 138% FPL Range

	(1)
	Weighted SIPP
2013	6,975,306
2014	6,503,895
2015	6,729,129
2016	5,906,960

Note: Computed using data from the 2014 SIPP. Survey weights are used.

Table 2.2: Differences between 2014 and 2008 SNAP Panels

	2008 Panel	2014 Panel
Instrument	DOS-based	Blaise/C#
Interview Frequency	3x/year	Annual
Reference Period	Previous 4 months	Previous year
Sample Size	65,522 households	53,070 households
Collection Format	Traditional question/answer	Question/Answer+EHC

Source: SIPP 2014 Panel Users' Guide

Table 2.3: Maximum Allotment Benefits for SNAP Participants by Household Size

	Household Size			
Year	1	2	3	4
2004	$141	$259	$371	$471
2005	$149	$274	$393	$499
2006	$152	$278	$399	$506
2007	$155	$284	$408	$518
2008	$162	$298	$426	$542
2009	$176	$323	$463	$588
2010	$200	$367	$526	$668
2011	$200	$367	$526	$668
2012	$200	$367	$526	$668
2013	$200	$367	$526	$668
2014	$189	$347	$497	$632
2015	$194	$357	$511	$649
2016	$194	$357	$511	$649
2017	$194	$357	$511	$649

Note: Valid for contiguous USA only - benefit amounts higher in Alaska and Hawaii. Source: Food and Nutrition Service (FNS)

Table 2.4: Summary Statistics

	SNAP Participant		Non-Participant	
VARIABLE	Mean	Std. Dev.	Mean	Std. Dev.
White (0/1)	0.43	0.50	0.70	0.46
Black (0/1)	0.30	0.46	0.11	0.32
Asian (0/1)	0.02	0.13	0.04	0.20
Hispanic (0/1)	0.20	0.40	0.12	0.33
Married (0/1)	0.31	0.46	0.58	0.49
Age of household head	46.11	13.49	47.82	14.05
Average household income	1,654.42	3,154.03	5,356.68	6,652.37
Employed (0/1)	0.42	0.49	0.73	0.44
Not in the labor force	0.51	0.50	0.25	0.43
Income % FPL	137.56	182.69	450.93	485.47
Has health insurance	0.81	0.39	0.87	0.34
Has Medicaid	0.55	0.50	0.07	0.25
Number of children in HH	1.13	1.37	0.58	1.01
Number of children aged less than 5	0.36	0.69	0.16	0.47
N	36,192		185,280	

Note: Data from the 2014 Survey of Income and Program Participation and are based on heads of household with a high school education or less. Survey weights are not used - the SIPP over-samples households from low-income areas so these numbers may not be representative of the United States as a whole.

Table 2.5: Status of SNAP State Policy Options

	2013 Status	2016 Status	Notes
SNAP State Policy			
Exempts one vehicle from asset test	ID, MI, PA, TX	ID, MI, TX	All other states: either no BBCE of any kind, or exempted multiple/all vehicles
Exempts all vehicles from asset test			
Broad-based categorical eligibility	No: AK, AR, IN, KS, MO, SD	No: AK, AR, IN, KS, LA, MO, SD	
Eligibility restrictions for noncitizens	Yes: CA, CT, WA	Yes: CA, CT, WA	Same 3 states for entire sample
Proportion of households with short recertification period (1-3 mo)	No states	No states	All states have recertification periods longer than 1-3 months.
Simplified reporting	All states except CA	All states	CA adopted in FY2014
Online application	No: AK, CT, DC, HI, ID, KY, MS, NC, OK, SD, WY	No: AK, DC, GA, HI, ID, MS, WY	

Note: These state policy options are described in Stacy, Tiehen, and Marquardt (2018). The mean EBT policy is numerically 100% for the entire sample.

Table 2.6: Results for Income (with SNAP Policy Index)

Dep. var: HH income as % of the FPL	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	-47.33*	-57.20**	-57.20**
	(18.64)	(14.27)	(14.27)
SNAP policy index	8.618	-7.589	-7.589
	(20.69)	(16.84)	(17.00)
Medicaid expansion x SNAP policy index	27.67	38.04*	38.04*
	(21.43)	(15.94)	(16.01)
Constant	193.7**	160.4**	188.7**
	(38.37)	(29.52)	(30.40)
Observations	325,421	325,421	325,421
R-squared	0.164		0.011
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	262.35	262.35	262.35
Number of HHs		13,001	13,001
F-test $\bar{X}_i$		509.8 (0.00)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

Table 2.7: Results for Log Income (with SNAP Policy Index)

Dep. var: Log HH income as % of the FPL	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	-0.145*	-0.0990+	-0.0990+
	(0.0680)	(0.0582)	(0.0588)
SNAP policy index	-0.0163	-0.0912+	-0.0912+
	(0.0581)	(0.0532)	(0.0542)
Medicaid expansion x SNAP policy index	0.0863	0.0897	0.0897
	(0.0751)	(0.0624)	(0.0631)
Constant	4.916**	4.807**	4.990**
	(0.196)	(0.143)	(0.180)
Observations	309,339	309,339	309,339
R-squared	0.200		0.013
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	5.23	5.23	5.23
Number of HHs		12,604	12,604
F-test $\bar{X}_i$		628.8 (0.00)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

Table 2.8: Results for Income (with SNAP Income Limit)

Dep. var: HH income as % of the FPL	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	-21.32*	-22.33*	-22.33*
	(9.646)	(10.11)	(10.36)
SNAP income limit	-4.374	-3.614	-3.614
	(13.81)	(15.81)	(16.12)
Medicaid expansion x SNAP income limit	-4.692	-4.696	-4.696
	(13.23)	(11.96)	(12.07)
Constant	206.7**	155.6**	185.7**
	(29.64)	(23.09)	(27.44)
Observations	325,421	325,421	325,421
R-squared	0.164		0.011
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	261.70	261.70	261.70
Number of HHs		13,001	13,001
F-test $\bar{X}_i$		510 (0.00)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

Table 2.9: Estimation Results for Hours Worked

Dep. var: number of hrs worked by HH head (month)	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	4.943	3.704	3.704
	(5.003)	(2.963)	(3.029)
SNAP policy index	0.179	-1.648	-1.648
	(4.329)	(2.886)	(2.887)
Medicaid expansion x SNAP policy index	-0.107	-1.433	-1.433
	(5.503)	(3.367)	(3.416)
Constant	121.2**	102.1**	99.70**
	(15.05)	(14.21)	(15.84)
Observations	325,421	325,421	325,421
R-squared	0.227		0.046
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	99.86	99.86	99.86
Number of HHs		13,001	13,001
F-test $\bar{X}_i$		516.1 (0.00)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

Table 2.10: Estimation Results for Log Hours Worked

Dep. var: log number of hrs worked by HH head	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	0.0271	-0.0110	-0.0110
	(0.0671)	(0.0645)	(0.0651)
SNAP policy index	-0.0196	-0.0173	-0.0173
	(0.0332)	(0.0386)	(0.0397)
Medicaid expansion x SNAP policy index	0.0106	0.0395	0.0395
	(0.0711)	(0.0685)	(0.0690)
Constant	5.163**	4.652**	4.568**
	(0.167)	(0.286)	(0.344)
Observations	189,463	189,463	189,463
R-squared	0.070		0.015
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	5.07	5.07	5.07
Number of HHs		8,821	8,821
F-test $\bar{X}_i$		137.2 (0.00)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

Table 2.11: Results for Hours Worked

Dep. var: number of hrs worked by second earner (month)	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	10.36	2.435	2.435
	(13.76)	(15.50)	(16.10)
SNAP policy index	13.37+	5.757	5.757
	(6.668)	(6.010)	(6.014)
Medicaid expansion x SNAP index	-6.421	-4.013	-4.013
	(14.59)	(16.00)	(16.58)
Constant	43.48+	69.70*	116.5**
	(22.51)	(31.70)	(42.44)
Observations	141,991	141,991	141,991
R-squared	0.107		0.063
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	93.66	93.66	93.66
Number of HHs		6,893	6,893
F-test $\bar{X}_i$		154.7 (0.00)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the “X” indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household. These estimations also includes employment status of the head of household.

Table 2.12: Results for Log Hours Worked

Dep. var: log number of hrs worked by second earner	(1)	(2)	(3)
VARIABLES	POLS	CRE	FE
Medicaid expansion	-0.0141	-0.000498	-0.000498
	(0.156)	(0.168)	(0.167)
SNAP policy index	0.0368	0.0200	0.0200
	(0.0841)	(0.0579)	(0.0568)
Medicaid expansion x SNAP index	-0.00667	-0.00234	-0.00234
	(0.168)	(0.179)	(0.178)
Constant	5.129**	5.185**	5.206**
	(0.190)	(0.170)	(0.244)
Observations	77,550	77,550	77,550
R-squared	0.055		0.035
Demographics	X	X	X
Time FE	X	X	X
State specific t-trends	X	X	X
State FE	X	X	
y mean	5.07	5.07	5.07
Number of HHs		4,318	4,318
F-test $\bar{X}_i$		6.0 (0.54)	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household. These estimations also includes employment status of the head of household.

Table 2.13: Estimation Results for Employment Status of Household Head

Dep. var: whether HH head is employed	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	POLS	Logit	Logit APE	CRE	CRE Logit	CRE Logit APE	FE
Medicaid expansion	0.0142	0.0660		0.000702	0.0118		0.000702
	(0.0199)	(0.107)		(0.00984)	(0.200)		(0.0103)
SNAP policy index	-0.000145	0.00789		-0.00976	-0.325		-0.00976
	(0.0223)	(0.121)		(0.0162)	(0.390)		(0.0162)
Medicaid expansion x SNAP index	-0.00278	-0.00681	0.0112+	0.00287	0.00199	0.000378	0.00287
	(0.0224)	(0.121)	(0.00667)	(0.0113)	(0.238)	(0.00301)	(0.0117)
Constant	0.656**	0.764*		0.633**	4.956*		0.658**
	(0.0658)	(0.349)		(0.0876)	(2.400)		(0.0978)
Observations	325,421	325,421	325,421	325,421	325,421	325,421	325,421
R-squared	0.215						0.039
Demographics	X	X		X	X		X
Time FE	X	X		X	X		X
State specific t-trends	X	X		X	X		X
State FE	X	X		X	X		
y mean	0.59	0.59		0.59	0.59		0.59
Number of HHs				13,001	13,001		13,001
F-test $\bar{X}_i$				591.4 (0.00)	20.7 (0.00)		

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the “X” indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

Table 2.14: Estimation Results for Employment Status of Second Earner

Dep. var: whether second earner is employed	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	POLS	Logit	Logit APE	CRE	CRE Logit	CRE Logit APE	FE
Medicaid expansion	0.0159	0.0775		-0.000286	-0.00981		-0.000286
	(0.0527)	(0.233)		(0.0379)	(0.747)		(0.0395)
SNAP policy index	0.0556	0.245		0.0359	0.923+		0.0359
	(0.0362)	(0.159)		(0.0325)	(0.502)		(0.0323)
Medicaid expansion x SNAP index	-0.0128	-0.0647	0.00285	0.00105	0.0199	0.000309	0.00105
	(0.0540)	(0.240)	(0.0174)	(0.0433)	(0.825)	(0.00546)	(0.0449)
Constant	0.279**	-1.017*		0.363+	-0.901		0.642*
	(0.0944)	(0.412)		(0.202)	(1.497)		(0.259)
Observations	141,991	141,991	141,991	141,991	141,991	141,991	141,991
R-squared	0.094						0.050
Demographics	X	X		X	X		X
Time FE	X	X		X	X		X
State specific t-trends	X	X		X	X		X
State FE	X	X		X	X		
y mean	0.55	0.55		0.55	0.55		0.55
Number of HHs				6,893	6,893		6,893
F-test $\bar{X}_i$				163.9 (0.00)	143.9 (0.00)		

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household.

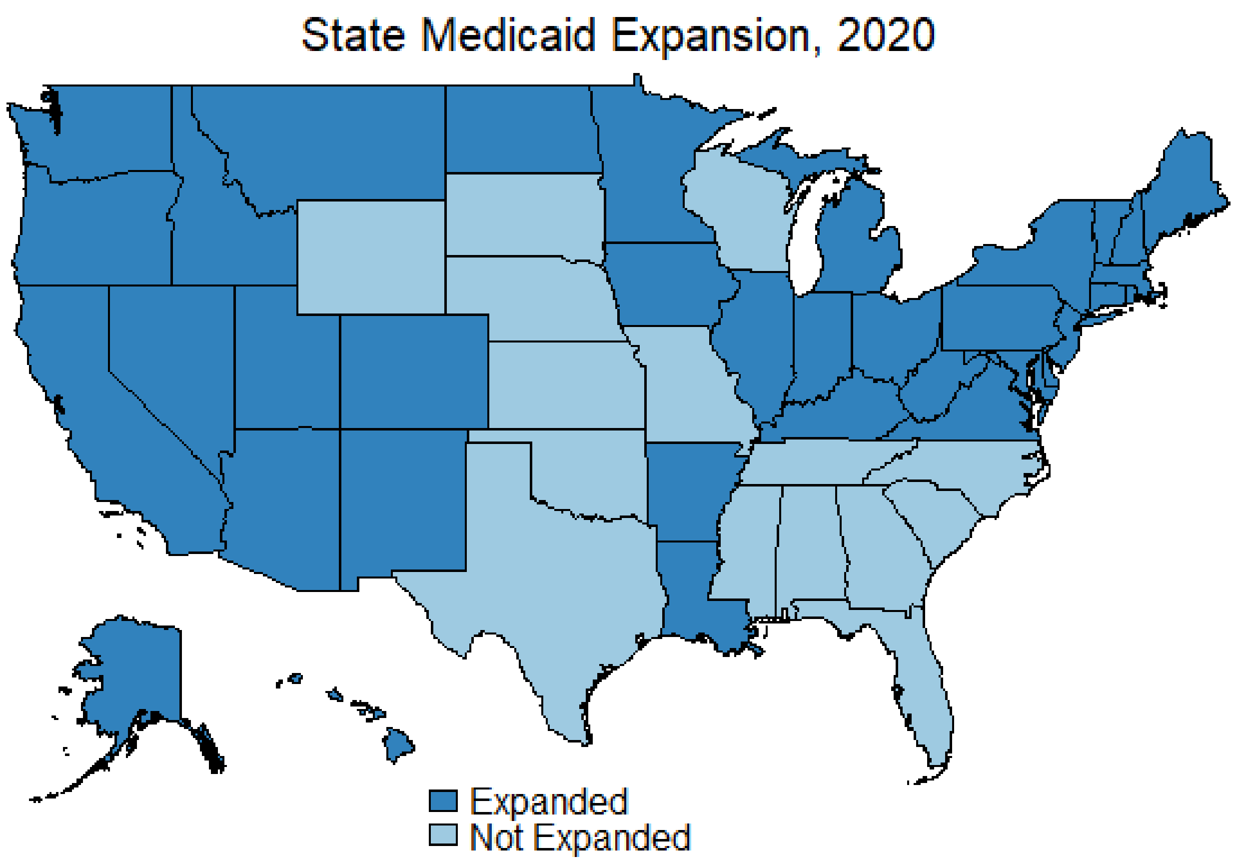
Table 2.15: Estimation Results for SNAP Participation

Dep. var: whether HH is enrolled in SNAP	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	POLS	Logit	Logit APE	CRE	CRE Logit	CRE Logit APE	FE
Medicaid expansion	0.0484**	0.300**		0.0466**	1.277**		0.0466**
	(0.0142)	(0.105)		(0.0131)	(0.323)		(0.0135)
SNAP policy index	0.0122	0.140		0.0296*	0.880+		0.0296*
	(0.0121)	(0.106)		(0.0141)	(0.479)		(0.0142)
Medicaid expansion x SNAP index	-0.0533**	-0.328**	-0.00381	-0.0555**	-1.584**	-0.00704	-0.0555**
	(0.0169)	(0.126)	(0.00792)	(0.0164)	(0.447)	(0.00486)	(0.0167)
Constant	0.0889*	-2.599**		0.0517	-13.97*		0.0829
	(0.0419)	(0.292)		(0.0590)	(6.666)		(0.0659)
Observations	325,421	325,421	325,421	325,421	325,421	325,421	325,421
R-squared	0.201						0.010
Demographics	X	X		X	X		X
Time FE	X	X		X	X		X
State specific t-trends	X	X		X	X		X
State FE	X	X		X	X		
y mean	0.22	0.22		0.22	0.22		0.22
Number of HHs				12,919	12,919		12,919
F-test $\bar{X}_i$				738.4 (0.00)	7.6 (0.27)		

Note: **, *, and + correspond to significance at the .01, .05, and .1 level, respectively. Demographic controls include race, age, marital status, sex, number of children, perceived health and educational attainment. Geographic controls include an indicator for urban, suburban, or rural household and state unemployment rate. In the table above, the "X" indicates the model includes the corresponding controls discussed in the model section. All three estimations include demographics, time FE and state specific t-trends. Population is households where the head of household has no degrees beyond high school and no persons over 65 live in the household. These estimates include two measures of food insecurity as controls for SNAP participation.

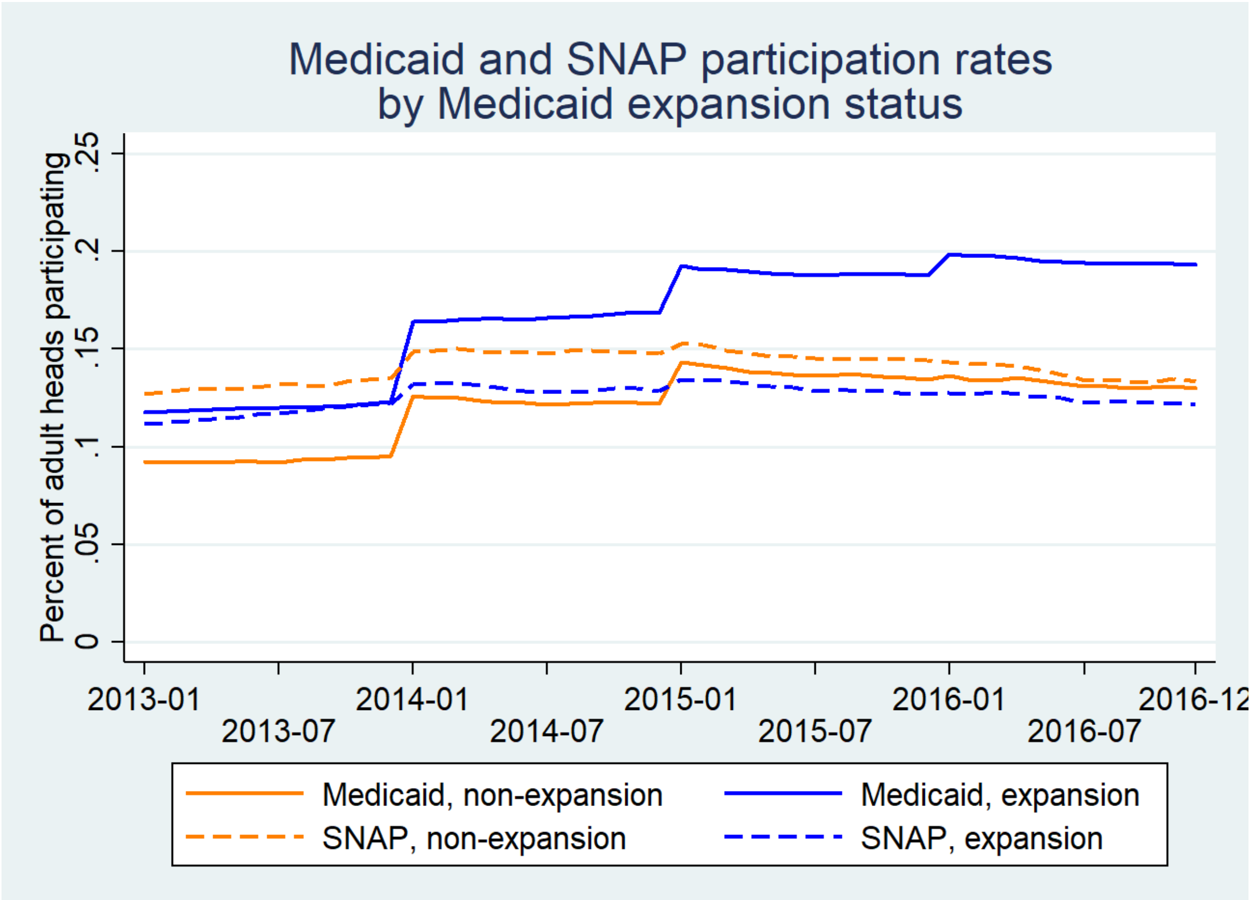
APPENDIX 2B: Figures

Figure 2.1: Status of Medicaid Expansion by State



Source: The Kaiser Family Foundation, 2020.

Figure 2.2: Medicaid and SNAP Participation Rates by Medicaid Expansion Status



Source: 2014 Survey of Income and Program Participation, author's calculations.

APPENDIX 2C: State-Level Statistics

Table 2.16: State Unemployment Rate by Year

	2013	2014	2015	2016
Alabama	7.5	7.2	6	6
Alaska	7	7	6.5	6.7
Arizona	7.9	7.3	6.4	5.6
Arkansas	7.4	6.6	5.6	4.2
California	9.6	8.2	6.8	5.7
Colorado	7.4	6	4.3	3.5
Connecticut	8.2	7.2	6.1	5.5
Delaware	7.2	6.1	5.1	4.6
District of Columbia	8.7	8	7.4	6.4
Florida	7.9	6.6	5.8	5
Georgia	8.7	7.5	6.4	5.5
Hawaii	5.1	4.7	4	3.1
Idaho	6.7	5.3	4.2	4
Illinois	9.2	8.3	6	6.1
Indiana	8.5	6.4	5.4	4.7
Iowa	4.9	4.4	3.9	3.7
Kansas	5.6	4.8	4.3	3.9
Kentucky	8.1	7.5	5.4	5.3
Louisiana	7	5.7	6.9	6.1
Maine	7.1	6.1	5	3.9
Maryland	6.9	6.1	5.4	4.6
Massachusetts	6.8	6.2	5.1	4.3
Michigan	9.1	8.1	6	4.9
Minnesota	5.3	4.6	3.8	3.7
Mississippi	8.9	8	6.7	6.1
Missouri	6.9	6.7	5.5	4.4
Montana	5.6	5	4.3	4.2
Nebraska	3.9	3.5	2.9	3
Nevada	10.4	8.6	7.2	6.2
New Hampshire	5.5	4.7	3.8	2.9
New Jersey	9	7.2	6.5	5
New Mexico	7	6.9	6.4	6.5

Table 2.16: (cont'd)

	2013	2014	2015	2016
New York	8.1	7	5.7	4.9
North Carolina	8.9	6.8	5.7	5.3
North Dakota	3.1	2.7	2.6	3.1
Ohio	7.5	6.6	5.1	5
Oklahoma	5.3	5	4.2	4.6
Oregon	8.5	7.3	6	5.1
Pennsylvania	7.9	6.5	5.4	5.2
Rhode Island	9.5	8.8	6.6	5.4
South Carolina	8.4	6.4	6.5	5.4
South Dakota	4	3.6	3.3	2.9
Tennessee	7.9	6.8	6.2	4.8
Texas	6.5	5.7	4.5	4.4
Utah	5	4.1	3.6	3.6
Vermont	4.6	4.1	3.8	3.3
Virginia	5.8	5.5	4.9	4.1
Washington	7.4	6.5	5.8	5.5
West Virginia	7.3	6.9	6.6	6.4
Wisconsin	7	6.1	4.8	4.3
Wyoming	5.1	4.4	3.9	5.1

Source: United States Bureau of Labor Statistics.

REFERENCES

REFERENCES

Ashenfelter, Orley and James Heckman (1974). “The Estimation of Income and Substitution Effects in a Model of Family Labor Supply”. In: *Econometrica* 42.1, pp. 73–85. ISSN: 0012-9682.

Burney, Shaheer, Rebecca Boehm, and Rigoberto Lopez (Aug. 2018). “The Impact of the 2014 Medicaid Expansion on SNAP Participation and Benefits”. In: Washington, D.C.

Canning, Patrick and Brian Stacy (July 2019). *The Supplemental Nutrition Assistance Program (SNAP) and the Economy: New Estimates of the SNAP Multiplier*. en. Economic Research Report 265. Economic Research Service, p. 60.

Carlson, Andrea et al. (2007). *Thrifty Food Plan 2006*. en. Tech. rep. U.S. Department of Agriculture, Center for Nutrition Policy and Promotion.

CMS (2020). *May 2020 Medicaid & CHIP Enrollment Data Highlights | Medicaid*.

Dickert-Conlin, Stacy et al. (2020). “The Downs and Ups of the SNAP Caseload: What Matters?” en. In: *Applied Economic Perspectives and Policy*. ISSN: 2040-5804.

Fox, Liana (2019). *The Supplemental Poverty Measure: 2018*. en. Tech. rep. U.S. Census Bureau, p. 33.

Frean, Molly, Jonathan Gruber, and Benjamin D. Sommers (May 2017). “Premium subsidies, the mandate, and Medicaid expansion: Coverage effects of the Affordable Care Act”. en. In: *Journal of Health Economics* 53, pp. 72–86. ISSN: 01676296.

Ganong, Peter and Jeffrey B. Liebman (Nov. 2018). “The Decline, Rebound, and Further Rise in SNAP Enrollment: Disentangling Business Cycle Fluctuations and Policy Changes”. en. In: *American Economic Journal: Economic Policy* 10.4, pp. 153–176. ISSN: 1945-7731, 1945-774X.

Garthwaite, Craig, Tal Gross, and Matthew J. Notowidigdo (2014). “Pubic Health Insurance, Labor Supply, and Employment Lock”. In: *The Quarterly Journal of Economics* 129.2, pp. 653–696. ISSN: 0033-5533.

Gray, Kelsey Farson and Shivani Kochhar (2015). *Characteristics of Food Stamp Households: Fiscal Year 2014*. Tech. rep. U.S. Department of Agriculture, Food, Nutrition Service, Office of Analysis, Nutrition, and Evaluation.

Gundersen, Craig (Oct. 2019). “The Right to Food in the United States: The Role of the Supplemental Nutrition Assistance Program (SNAP)”. en. In: *American Journal of Agricultural Economics* 101.5, pp. 1328–1336. ISSN: 0002-9092, 1467-8276.

Han, Jeehoon (2020). “SNAP Expansions and Participation in Government Safety Net Programs”. en. In: *Economic Inquiry* 58.4, p. 58.

Hoynes, Hilary Williamson and Diane Whitmore Schanzenbach (2009). “Consumption Responses to In-Kind Transfers: Evidence from the Introduction of the Food Stamp Program”. In: *American Economic Journal: Applied Economics* 1.4, pp. 109–139.

Kaestner, Robert et al. (2017). “Effects of ACA Medicaid Expansions on Health Insurance Coverage and Labor Supply”. en. In: *Journal of Policy Analysis and Management* 36.3, pp. 608–642. ISSN: 1520-6688.

Lanese, Bethany G., Rebecca Fischbein, and Chelsea Furda (Oct. 2018). “Public Program Enrollment Following US State Medicaid Expansion and Outreach”. In: *American Journal of Public Health* 108.10, pp. 1349–1351. ISSN: 00900036.

McMorrow, Stacey et al. (May 2017). “Medicaid Expansion Increased Coverage, Improved Affordability, And Reduced Psychological Distress For Low-Income Parents”. en. In: *Health Affairs* 36.5, pp. 808–818. ISSN: 0278-2715, 1544-5208.

Meyer, Bruce D., Nikolas Mittag, and Robert M. George (Aug. 2020). “Errors in Survey Reporting and Imputation and their Effects on Estimates of Food Stamp Program Participation”. en. In: *Journal of Human Resources*, p. 0818. ISSN: 0022-166X, 1548-8004.

Moore, Jeffrey (2008). “Seam Bias in the 2004 SIPP Panel: Much Improved, Much Bias Still Remains”. In: *Survey Methodolgy Research Report*. Vol. 3. Washington DC: U.S. Census Bureau, Statistical Research Division.

Nord, Mark and Anne Marie Golla (2009). *Does SNAP Decrease Food Insecurity? Untangling the Self-Selection Effect*. en. Number: 1477-2016-121152.

Nord, Mark and Mark Prell (2009). *Food Security Improved Following the 2009 ARRA Increase in SNAP Benefits*. en. Economic Research Report 116. Economic Research Service, p. 52.

Rosenbaum, Dorothy (2013). *The Relationship Between SNAP and Work Among Low-Income Households*. en. Tech. rep. Center on Budget and Policy Priorities, p. 36.

Schmidt, Lucie, Lara Shore-Sheppard, and Tara Watson (Nov. 2019). *The Impact of Expanding Public Health Insurance on Safety Net Program Participation: Evidence from the ACA Medicaid Expansion*. en. Tech. rep. w26504. Cambridge, MA: National Bureau of Economic Research, w26504.

Stacy, Brian, Laura Tiehen, and David Marquardt (2018). *Using a Policy Index To Capture Trends and Differences in State Administration of USDA's Supplemental Nutrition Assistance Program*. en. Tech. rep. ERR-244. U.S. Department of Agriculture, Economic Research Service, p. 38.

US Census Bureau (2013). *SIPP-EHC: 2011 and 2012 Field Test Evaluation*. Tech. rep.

US Census Bureau - Economic and Statistics Administration (2019). *Survey of Income and Program Participation, 2014 Panel Users' Guide*. Tech. rep., p. 164.

Wooldridge, Jeffrey (2019). *Introductory Econometrics: A Modern Approach*. 7th ed. Boston, MA: Cengage.

CHAPTER 3

THE IMPACT OF PROJECT-BASED LEARNING ON HIGH-SCHOOL STUDENTS' SCIENCE LEARNING OUTCOMES

3.1 Introduction

Positive educational outcomes are the goal of many educational reforms, as they have positive effects for employment opportunities (Altonji, 1995), health outcomes (Cutler and Lleras-Muney, 2006), and civic engagement (Dee, 2004). However, schools face trade-offs for funding and other resources to meet or exceed all of their objectives. It is imperative, then, to study the effectiveness of potential classroom reforms and assess the costs of implementing them. This paper assesses whether a study from the 2018-2019 school year yielded an academic benefit to the students enrolled in its treated classrooms.

One method of instruction that has gained traction over the last decades has been project-based learning (PBL). PBL is student-driven, teacher-led comprehensive framework in which teaching is motivated by engaging students in investigation: students ask questions, debate ideas in peer groups, and create/design experiments to solve problems (Blumenfeld et al., 1991). Under PBL, items of a curriculum are learned through a project that contains a "unit driving question." Students engage with the question through discussion, demonstrations, videos, and experiments (Schneider, Krajcik, et al., 2020). Then, the project concludes with a product or presentation thus promoting creativity and hands on opportunities, the application of acquired knowledge, and the development of necessary work and life skills (Bell, 2010).

Chen and Yang (2019) conduct a meta-analysis of the effect of project based learning on students' academic achievement. After evaluating 30 journal articles published from 1998-2017, they find PBL has a positive and large significant average weighted effect (0.71 standard deviations) on students' academic achievement, compared to traditional learning.

Despite these findings, the economics of education literature has exclusively focused on other interventions that affect outcomes, such as class sizes (Krueger and Whitmore, 2001) and the use of technology in classrooms (Machin, McNally, and Silva, 2007). No economic studies (that I am aware of) focus on the impact of curriculum and pedagogy on student outcomes.[1] This gap in the literature is the primary motivation for this work.

This paper uses the same primary collected data (2018-2019 school year) as Schneider, Krajcik, Joseph, et al. (2021). The project, Crafting Engaging Science Environments (CESE), is an NSF funded cluster-randomized controlled trial to study the effect of project-based learning on the scientific achievement of high school chemistry and physics students in Michigan and California.

The CESE intervention provided teachers in the treatment group training, the curriculum for partial school year, and laboratory supplies. The training occurred over multiple multi-day sessions, covering the Next Generation Science Standards (NGSS), pedagogy in PBL, and specific instructions for how to conduct the lab experiments. The curriculum contained daily lesson plans for three scientific units. The teachers in the control were not restricted in the instruction of their own classes, but did receive training on understanding the NGSS.

I extend the analysis conducted in Schneider, Krajcik, Joseph, et al. (2021) and use pooled OLS with school level fixed effects to estimate the treatment effect, as opposed to the hierarchical linear model (HLM) used in Schneider, Krajcik, Joseph, et al. (2021). The estimation strategy detailed in this paper accounts for the unobserved, time constant heterogeneity across schools. The use of the fixed effects reduces the presence of omitted variable bias by allowing heterogeneity to be correlated with student and school characteristics. I also test for teacher and student program attrition based on observables that could bias any treatment effect, and bound the treatment effect by calculating Lee bounds. Lastly, this paper seeks to contextualize the effectiveness of this particular intervention with other economics of education papers.

The results show there is a positive, statistically significant treatment effect (0.30 standard deviations) and the effect is robust to program attrition, as the Lee bounds of the estimates

[1] There are studies focused on teacher quality, such as Rivkin, Hanushek, and Kain (2005), Rockoff (2004), Aaronson, Barrow, and Sander (2007), and Nye, Konstantopoulos, and Hedges (2004).

do not contain zero. The effect is larger in magnitude than the finding of 0.20 standard deviations in Schneider, Krajcik, Joseph, et al. (2021). The main differences between the estimates is the inclusion of school fixed effects within this paper. The estimations where I exclude them, and include other controls return similar estimates to those found in Schneider, Krajcik, Joseph, et al. (2021).

This paper has two main contributions. First, it contributes to the study of project-based learning, a method that is praised by experts but for which rigorous empirical evidence of its merits is limited. By accounting for program attrition at different levels and using alternative estimation methods that mitigates bias from school heterogeneity, we can confirm the robustness of the estimates found in Schneider, Krajcik, Joseph, et al. (2021).

Second, this paper contributes to the education of economics literature. I study a method of instruction using econometric tools and find the intervention has positive effect on student achievement and therefore, an impact on human capital. A key advantage of this paper is the use of data from a cluster randomized control trial (CRTC). Over the last 20 years, randomized controlled trials have become the gold standard in economics, as they facilitate identification and average treatment effects are perceived as more reliable than results from other empirical methods (Deaton and N. Cartwright, 2018).

Despite having a geographical component that might hinder the external validity and the project's potential for application across schools in the United States, I hope this study sheds light on the importance of evaluating different instruction methods and their effects on student learning and achievement. While other types of interventions and treatments (i.e. class sizes and technology employed in the class) are easier to quantify in terms of student outcomes, teachers and their teaching-learning philosophy have long-lasting effects on students (T. J. Cartwright and Hallar, 2018). Future research in the economics field should incorporate of instruction to impact evaluations related to student outcomes and earnings.

The remainder of the paper is organized as follows: Section 3.2 details the participating schools, student and teacher data, and section 3.3 discusses project based learning and the intervention. The model section (3.4) is followed by the results, discussion and paper conclusion.

3.2 Data

3.2.1 Participating schools

We have collected data from three large public urban school districts in addition to several urban and rural schools districts from California and Michigan. In addition to the geographic variation, the schools vary considerably with respect to language spoken at home and racial composition. Table 3.1 highlights the differences in student demographics across the three urban regions and the one sub-urban areas of Michigan. The resulting sample with matched pre- and post- tests covers 119 teachers at 61 different high schools, and over 4,000 students.

The school districts were recruited prior to school-level assignment of the treatment condition. When a principal or similar representative of the school agreed to participate in the study, they were then randomized with equal probabilities of being in either condition. Schneider, Krajcik, Joseph, et al. (2021) details the post-randomization balancing of the schools involved in this study.

3.2.2 Student data

The efficacy study collected a pretest, a background survey, a post-test, and an exit survey for the purposes of examining any effects of the PBL curriculum on either scientific achievement or attitudes and practices toward science.

3.2.2.1 Assessments

The students were given a pretest on general scientific knowledge. The assessment items were chosen by the PBL designers from a bank of publicly available science questions used by the National Assessment of Educational Progress (NAEP). A variety of middle- and elementary- school grade level items were used in the assessment, to differentiate students' skill levels. The pretest consisted of 12 items, seven were multiple choice and three were constructed response in the form of short answer. The remaining question was a combination of both. Statistics of these items are included in Table 3.3. Both chemistry and physics students received the same pretest.

The constructed response items were partially scored by college students majoring in a STEM. In an effort to reduce labor costs, the research team employed machine learning from a third-party to complete the scoring for one question. Questions sent to the machine had a sufficiently high human agreement score (>0.8 κ) for the machine to return results with fit statistics ranging between 0.64 and 0.81 (Maestrales et al., n.d.). As the scoring for that constructed response item was non-binary, the inclusion of the item allowed for the researchers to better differentiate the students for pretest ability.

Item response theory was used to construct a measure of scientific ability or skill, as consistent with (Walstad and Robson, 1997; Chan and Kennedy, 2002). This measure of ability is advantageous relative to a metric such as percentage of questions correctly, as this solution rewards students more for answering difficult questions correct and penalizes less for answering the difficult questions incorrectly. The process uses a Bayesian updating algorithm on a prior of both student abilities and question difficulty, until the question parameters converge and the student scores are calculated (Gwet, 2014).

The post-test was administered in the final month of the course, after the CESE units had concluded. Unlike the pretest, different chemistry and physics students received different assessments. Assessment items were chosen by the CESE team from a bank of constructed response questions provided to them by the Michigan Department of Education (MDE). The items in the bank were questions that may have been used in state standardized testing for the 11th grade. The team then chose the items based on which questions appropriately fit the respective subjects, as the test bank covered all high school courses, that is, included courses like biology and astronomy.

To make the post-tests comparable, the Schneider, Krajcik, Joseph, et al. (2021) paper used a measure of post-test ability from the process of "equivilation" that studied student performances from the control group for both pre- and post- tests. This process allows for a conversion of number of questions correct from one subject to the next, so that students with similar converted scores have similar abilities in science. This paper examines both the equivalent scores as well as scores standardized within a subject, and reports only the normalized scores, as the results are practically invariant to which measure of post-test ability is used.

3.2.2.2 Background Surveys

Students and teachers for both the control and treatment conditions were given background surveys to complete at the beginning of the year. They were asked questions about attitudes and aptitudes towards learning, as well as basic demographic questions.

These questions, while generally informative, are of limited use in most analysis because of the large number of missing student responses. The surveys were typically administered after the respective test, and the limited time in the class period, in addition to an unwillingness of students to provide answers, led to many of these missing responses. Where possible, administrative records on demographics were used to fill in missing information in the background survey and the student exit surveys.

Table 3.1 shows how some student characteristics vary across the regions in the study. One major difference between students in these regions is the racial composition of the student bodies. Nearly two-thirds of Michigan (non-Detroit) students self-reported that they are non-Hispanic white, the largest racial and ethic majority across the four regions. Roughly half of the students in California identified as being Latino.

Another stark contrast across the regions is educational attainment of a student's mother. In Michigan, 26% reported that their mother had a high school degree or less, and 47% had some degree beyond high school (the remainder were uncertain, did not know, or did not live with their mother). In Los Angeles, the percentage of students with mother's holding a high school diploma or less was slightly lower, however the proportion with a post-secondary degree was substantially less (10.6%).

Data from the background surveys were taken as official record for the data set. In the event that administrative data were supplied for race or sex, that data was only used to complement missing responses from the background survey. Lastly, I use United States Social Security Administration records of the frequencies of live births and sexes to compute a ratio of females to total births of persons with that name born between 2000 and 2005, the birth years corresponding to high-school aged students in the 2018-2019 school year. If the sex of the student was missing, I replaced it with the most likely sex depending on the student's first name.[2]

3.2.3 Teacher Data

Similar to the students, the teachers were given background and exit surveys at the beginning and end of the school year. The background survey asked questions about general attitudes toward science and learning, basic demographic information, and information about the typical work load and class sizes they work with. The summary statistics are found in Table 3.2.

As expected, there were differences between teachers in Michigan and those in California. A higher percentage of teachers in Michigan and Detroit (85% and 71%,respectively) had a masters degree compared to Los Angeles and San Diego (56% and 66%, respectively). Additionally, more teachers in the California cities had a English Language Learner (ELL) certification compared to Michigan and Detroit. This was primarily due to the higher percentage of Hispanic students in Los Angeles and San Diego.

Lastly, the teachers' familiarity with project-based learning did not vary much across the areas of study. On a scale of 1 to 3 (with 1 being least familiar with project based learning), the average familiarity ranged from 1.29 to 1.93, with the highest average in San Diego.

[2] The sex distribution of a name varies by state. For names with at least 5 live births in that state for the year (the Social Security Administration's minimum reporting threshold), I used the data for the state in which the student attends high school.

3.2.4 Attrition

There are various sources of attrition in this study, each with different implications for any analysis of the program. These happened at the school level, teacher level, and student level. A breakdown of the attrition by level and assigned treatment condition is found in Table 3.6.

Much of the school-level attrition cannot be disentangled from the teacher-level attrition, as at some schools the study only included one teacher participating for the entire school. When possible, the researchers attempted to continue the study. However, a large portion of the schools who withdrew from the study before any data were collected but after the randomization were control schools.

There were numerous reasons for individual teachers to leave the study. Overall, 18 teachers exited the survey. The majority (13 teachers) provided a statement that they were not continuing in the program and nine provided a reason. Treatment teachers left the study at higher rates than control teachers (12 vs 6), which could be a result of a combination of factors. One reason was a lapse in teaching days in the classroom, such as for an illness or pregnancy. Teachers in the treatment condition cited not wanting to have the substitute teacher carry out their plans or difficulty fitting a new curriculum back into the schedules while having a disruption in their time teaching. As the schools were randomized into treatment conditions, it is plausible that the control teachers experienced leaves of absence during the school year as well, but they had no reason to report such an absence, and their participation in the program only required their presence for the beginning and end of the course.

Another possible explanation for the higher propensity for the treatment teachers to drop from the program was that they did not believe that the program was effective. This indeed was a cited reason for leaving the program by at least two teachers, where one said the materials were too complicated for their students and another said the content was at an "inappropriately low" level. The teacher's belief that the program is not effective is a threat to the identification of the treatment effect, as the scores from the students in the treatment are conditioned on whether the teacher thought the program was working for their students.

Lastly, student-level attrition was possible from many different areas. If the students did not enjoy the PBL lessons and were able to disentangle those units from a science class generally, they may have opted to drop the class, switch to another section if the other teacher was not participating in the study, or be absent on the day of the post-test.

3.3 Intervention

3.3.1 Project Based Learning

The intervention for Crafting Engaging Science Environments (CESE) included multiple channels. First, teachers in the treatment received training on "project-based learning" (PBL) in addition to learning about the Next Generation Science Standards multiple times of the year. Second, the teachers received lesson plans for three units for their respective class (chemistry or physics) as well as training on the experiments and units, as well as the lab materials. Lastly, the teachers received remote support for conducting the units as needed. A more complete explanation can be found in (Schneider, Krajcik, Joseph, et al., 2021).

3.3.2 Recruitment

School districts were recruited to the study before treatment condition was randomly assigned. Once the school principal or other agent of the district agreed to participate, the individual schools were randomized into treatment or control. The randomization was performed at the school, as opposed to teacher, level to assuage concerns over spillovers from one teacher in the treatment group discussing the practices and content of PBL with their colleagues. Further, it allowed the option for schools to have all of their teachers participate in the study so that the course content and coverage was similar for all of their students. This will facilitate the instruction in future courses. The option was especially important for schools on the block schedule system, where students would not necessarily take Chemistry

B in the spring semester from the teacher who taught Chemistry A, where both of those courses were using PBL units.

Incentives were offered to teachers, and by extension, to the schools for participating in the study. The compensation varied between treatment and control. The teachers in the treatment group received monetary compensation and were provided the lab materials and the curriculum for the three PBL units. Additionally, they were allowed to remain in the treatment condition for the 2019-2020 school year.

The control teachers also received monetary compensation, though $500 less than the treatment condition. They were also offered the option to continue in the study receiving all of the treatment materials for the 2019-2020 school year during a continuation study where the results in their classrooms would be compared to the results of teachers who received the treatment for two consecutive years.

3.4 Model

3.4.1 CESE Treatment Effect

This paper seeks to identify any treatment effect of the CESE program using pooled OLS with school-level fixed effects. The outcome variables for identification are normalized measures of the post-test, so treatment effect is measured in standard deviations of performance on the post-test.

As the intervention was a cluster-randomized control trial, additional controls are only strictly required for when the observable characteristics are statistically different across treatment conditions. However, all models pooling chemistry and physics students include an identifier to indicate the subject, as the treatment was only similar, not the same across subjects.

The main estimation model is below, where the outcome of interest (y_{isb}) for student i at school s with teacher b is a measure of scientific ability based on the post-test at the end of the 2018-2019 study.

$$y_{isb} = \alpha + \gamma\ Treat_s + X_{isb}\beta + T_b\delta + \Sigma_s S_s + \epsilon_{isb} \tag{3.1}$$

The array of student-level controls (X_{isb}) include the estimated scientific ability based on the pretest, race, gender, self-reported GPA, and mother's educational attainment. Teacher-level controls (T_b) include years of teaching experience, squared teaching experience, and reported familiarity with PBL prior to the beginning of the intervention. School-level variables (S_s) include a school-specific intercept that captures socioeconomic factors that influence scientific achievement school-wide. For this model, standard errors are clustered at the school level.

3.4.2 Lee bounds

The presence of attrition in the study complicates the ability to recover the true treatment effects from the CESE intervention. One increasingly popular method for bounding the treatment effect is estimating Lee bounds (Lee, 2009). These bounds create an interval from the data and reasonable assumptions about how the attrition influences the means of the outcome variables conditional on treatment condition. The "true" treatment parameter is likely to be within those bounds.

The assumptions of Lee bounds are random assignment and monotonicty of the treatment condition's influence on attrition probabilities. For this study, the treatment condition was randomly assigned. The monotonicity assumption is likely to hold as well, as the treatment generally considered more work and effort to participate in the study. Therefore, the treatment condition likely did not decrease the probability of attrition from the study.

The Lee bounds can be useful in two different ways. First, the bounds are not subject to misspecification of any attrition equations. This is because they do not incorporate any covariates. Second, whether the bounds contain zero can give researchers more evidence that the real treatment effect is non-zero.

Lee's bounds estimates upper- and lower- limits on the treatment effects. The estimation of the bounds uses data on both the sample with an observed outcome (the standardized post-test) and the students without the post-test (defined in this paper as an attrited student). The sample means for control and treatment are then trimmed and compared to compute the bounds.

I estimate these bounds with different concepts to reflect different kinds of attrition in the program. First, I estimate them with all student records in the data set. This the most sensible form of the bounds, however the roster data for classrooms that did not produce any pretests or other data may contain duplicated students. Therefore, I also estimate the including students who produced some form of the pretest as well as the pre-test in addition to the background study.

Lastly, I follow Tauchmann (2014) to estimate additional Lee bounds that are "tightened" using the regional dummy indicators. The tightening involves separating the data into treatment and control within each region, and trimming the data at different rates across the region as appropriate. This allows for a sensible approach to addressing different rates of attrition across regions. This also allows for a tighter interval as the trimmed parts of the sample may be more geographically representative than the trimmed portions when not first partitioning the data by region.

3.4.3 Analysis of Attrition Decisions

For examining the attrition, I estimate the probability of a teacher or student exiting the study, by defining $attrit_t$ to be 1 if the teacher exited the study and 0 otherwise, and $attrit_s$ equal to 1 if the student attrited, conditional on the student's teacher remaining in the program.

To examine the teacher's probability of exit, I use both OLS and probit to estimate model 3.2 below. The control variables include teacher experience, its square, and the classroom[3] average on the student pretest interacted with the treatment condition. Region identifiers, indexed as r, are used in place of the school identifiers to conserve degrees of

[3]Some teachers taught multiple sections of the course. We pooled all students from all sections with the same teacher due to both data limitations and students moving between sections.

freedom while also capturing region invariant events that have occurred. For example, one region experienced a great budget crunch that led to the forced retirement and some school restructuring for the spring semester of the 2019 school year, and another urban area experienced a teacher strike in the middle of the school year. Standard errors are clustered at the region level.

$$Attrit_{sb} = \alpha + \gamma\ Treat_b + X_{sb}\beta + T_b\delta + \Sigma_r R_r + u_{sb} \tag{3.2}$$

Additionally, I also interact the treatment condition with the teacher's students' pretest average. The purpose of this is to examine whether any decision to exit the study is influenced by the performance of the students or the students' ability for which the pretest is a proxy. Equation 3.3 below captures the interaction of interest:

$$Attrit_{sb} = \alpha + \gamma_1\ Treat_b + Treat_b \text{ x } \bar{y_{sb}}\gamma_2 + X_{sb}\beta + T_b\delta + \Sigma_r R_r + \epsilon_{sb} \tag{3.3}$$

For estimating student-level attrition, I estimate model 3.4 using both OLS and probit methods. The student-level control variables are GPA, pretest-based ability score, and gender of the student. I use school identifiers and standard errors clustered at the school level.

$$Attrit_{isb} = \alpha + \gamma\ Treat_b + X_{isb}\beta + T_b\delta + \Sigma_r R_r + e_{isb} \tag{3.4}$$

3.5 Results

3.5.1 Main treatment effect

Overall, I find statistical evidence that the CESE PBL intervention has a positive effect on science comprehension, and the magnitude of roughly 0.3 standard deviations is substantial. However, the significance and magnitude of the estimate depends on the specifications of the model. The results with various specifications for the main effect model can be found in Table 3.4. Column 4, with school fixed effects and the student pretest as controls, is my preferred specification.

Column 1 shows that without other controls, the treatment effect is not statistically different from zero. However, the inclusion either school fixed effects (column 2) or the students' pretest (column 3) increases the point estimate sufficiently high so that the treatment effect is non-zero. Column 4 contains the controls of the two specifications before it, and contains neither student nor teacher demographics.

The inclusion of a student demographics result in a loss of some students (7% of the sample) who did not complete the background survey. The coefficients on the dummy variable for race (non-Hispanic white students are the reference group) suggests that minority students performed worse on the post-test than white students, even when conditioning on the pretest performance and school fixed effects. This finding does merit further investigation as to why the performance of those students was worse at the end of the year. However the interaction of race categories and the treatment effect suggests that the treatment effect itself does not depend on race.

Higher performing students entering the study, as measured by pretest performance and self-reported grade-point average, performed better on the post-test. This suggests that the two measures of student ability are measuring slightly different things.

3.5.2 Estimation of Lee bounds

The estimated bounds of the treatment effect are reported in Table 3.8. The bounds are reported in pairs of student sample definitions (all students, all students with a pretest, all students with a pretest and background survey) that reflect different definitions for what students participated in the study. The "tightened" specification trims the data and estimates the bounds within the regional identifiers consistent with Tauchmann (2014).

The various specifications contain two general takeaways. First, the bounds do not contain zero, so it is likely that the true treatment effect of the intervention is positive. However, the lower bound for the specification with the most liberal definition of a student participating in the study was statistically zero. Second, the Lee bounds are large. Therefore, the bounds are uninformative beyond providing an alternative source of evidence for the existence of a positive treatment effect.

3.5.3 Teacher-level attrition

The results on teacher-level attrition differ depending on whether OLS or probit is used in the estimation. The OLS results suggest the treatment status does not influence the decision to exit, instead the only significant variables are the region indicators. The probit results show that treatment teachers are more likely to exit, and the coefficient on the interaction of treatment is positive as well. This finding provides evidence that teachers with lower-performing students were not more likely to attrit from the study, which alleviates a concern of selective attrition biasing the treatment effect upward.

The average partial effect of the average pretest scores was statistically zero for the treatment group. That is, the measure of student scientific ability had no apparent influence on whether the teacher left the study.

3.5.4 Student-level attrition

The results for student-level attrition can be found in Table 3.7. I find no evidence that the decision to exit his or her classroom depending on whether the student is in the treatment condition.

3.6 Discussion

I found broad statistical support for the claim that the Crafting Engaging Science Environment's program as a whole increased scientific achievement as measured by the post-test administered at the end of the course. In this section, I discuss further venues for studying this sample as well as considerations for adapting a similar curriculum at a larger scale.

As the initial results seem promising, further investigation into these same student's outcomes may be more informative for this program translating into other real world outcomes. Some of these outcomes of interest in the near term would be number of science and other STEM courses taken through the rest of high school, attendance, college admittance exam scores (e.g. ACT and SAT), and high school graduation rates. With close partnerships between the states and school districts, the researchers may be able to obtain such data.

In this study, the teachers in the control group were not given any instructions on how to teach their classes and were not given information about the unit coverage in the treatment. Nor were the teachers instructed on how to teach. Therefore, teachers may have been teaching their students in PBL or in a manner similar to PBL. This would lead to a smaller estimated treatment effect in the study.

The treatment effects generalizability then depends on the current prevalence of PBL instruction in the new area to be treated. For example, if a university has already implemented PBL practices in their curriculum for future science teachers, and those teachers are geographically clustered into school districts, then the benefit of CESE program may be lower for those districts.

One consideration for rolling out these units at a larger scale is the cultural sensitivity of the both the assessment items and the units themselves. In the main effect, racial and ethnic minorities students performed worse than non-Hispanic students on the post test, even when conditioning on pretest performance, mother's education, and school fixed effects. While this is far from proof that the items may have some form of cultural bias in them, a further examination into these items could reveal potential areas of change in the program. For instance, the pretest contained a question about the behaviors of walking on ice as well a question about two vehicles travelling on a highway. Urban students from warm climates will likely have less life experience with those two scenarios than a rural student from colder climate.

Another major consideration for mass roll out is understanding how each aspect of the treatment influenced the overall treatment effect. For instance, in addition to the curriculum, the teachers received training and support in how to teach those specific units as well as the materials to conduct the experiments. Therefore, some of the treatment effect may be realized in schools without the curriculum but instead having more lab units with sufficient materials.

3.7 Conclusion

In this paper, I analyze data from the Crafting Engaging Science Environment's (CESE) Project-Based Learning (PBL) study for the 2018-2019 school year. The study was conducted in schools in four areas of the United States: sub-urban Michigan, Detroit, Los Angeles, CA, and San Diego, CA. The project intervention itself operates over multiple channels, with curriculum, teacher training, and equipment provisions for teaching high school chemistry and physics classes.

I find sound evidence to support the findings in Schneider, Krajcik, Joseph, et al., 2021 that the CESE intervention had a positive and significant effect on students' scientific learning. The statistically significant point estimates range from 0.24 to 0.34 standard deviations, with the preferred specification having an effect of +0.30 standard deviations.

I also considered the effects of multiple levels of attrition and how that may influence the findings. My auxiliary regressions do not provide evidence to suggest that teachers with lower-performing students were more likely to exit the study, nor was there evidence to suggest that students in the treatment group were more likely to exit the study because of their treatment condition. Additionally, I compute the Lee bounds for various definitions of attrition. The concepts differ by how much student data was collected by the researchers. In all cases, the estimated treatment effects fall within those bounds, and the bounds do not contain zero.

The CESE version of PBL has an optimistic outlook for increasing scientific achievement in American high schools, as this study took place in a mix of urban and rural high schools. However, more studies need to be performed before large roll-outs of the curriculum, as the project used a high degree of hands on training and assistance.

APPENDIX

Table 3.1: Student Summary Statistics

	Michigan		Detroit		Los Angeles		San Diego	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Demographic Characteristics								
White	0.63	0.48	0.01	0.10	0.03	0.16	0.12	0.32
Black	0.09	0.28	0.56	0.50	0.02	0.14	0.01	0.12
Hispanic	0.11	0.32	0.10	0.30	0.50	0.50	0.54	0.50
Female	0.50	0.50	0.63	0.48	0.53	0.50	0.55	0.50
Speaks Spanish	0.04	0.21	0.08	0.27	0.49	0.50	0.37	0.48
Mother Educ. (Less than HS)	0.26	0.44	0.34	0.47	0.30	0.46	0.33	0.47
Mother Educ. (Post HS Degree)	0.47	0.50	0.30	0.46	0.11	0.31	0.32	0.47
Students' Outcomes								
Pre-test score	0.08	0.82	-0.25	0.88	-0.29	0.85	-0.06	0.91
% of students with GPA 3.0+	0.57	0.50	0.46	0.50	0.25	0.43	0.47	0.50
In chemistry class (0=In physics)	0.68	0.47	0.67	0.47	0.72	0.45	0.72	0.45
Students' Perceptions of STEM								
Math importance (1-7)	5.07	1.76	5.68	1.59	5.45	1.62	5.19	1.80
Math usefulness (1-7)	5.08	1.70	5.58	1.58	5.46	1.61	5.17	1.74
Biology importance (1-7)	4.38	1.80	4.91	1.76	4.90	1.65	4.61	1.77
Biology usefulness (1-7)	4.24	1.79	4.82	1.74	4.75	1.68	4.46	1.77
N	2,155		619		2,287		1,659	

Students' perceptions of STEM questions were asked on a 7-point Likert scale, with 1 corresponding to "strongest disagreement" and 7 to "strongest agreement."

Table 3.2: Teacher Summary Statistics

	Michigan		Detroit		Los Angeles		San Diego	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Professional Characteristics								
Years of experience	17.24	7.09	19.64	8.79	18.58	9.80	13.83	9.09
Masters degree	0.85	0.37	0.71	0.47	0.56	0.50	0.66	0.48
Special ed credential	0.00	0.00	0.00	0.00	0.12	0.33	0.00	0.00
Familiarity with PBL (1-3 scale)	1.77	0.48	1.29	0.47	1.64	0.60	1.93	0.47
ELL credential	0.00	0.00	0.00	0.00	0.76	0.44	0.61	0.50
Demographic Characteristics								
Female	0.49	0.51	0.57	0.51	0.70	0.47	0.64	0.49
Hispanic	0.00	0.00	0.00	0.00	0.15	0.36	0.14	0.36
Black	0.00	0.00	0.50	0.52	0.00	0.00	0.04	0.19
Project Load								
Number of students (pretest)	55.26	32.05	36.41	30.78	55.78	38.03	51.84	38.00
Number of students (posttest)	49.18	30.09	30.06	37.20	49.61	47.19	47.31	45.09
Chemistry	0.59	0.50	0.59	0.51	0.71	0.46	0.66	0.48
N	39		17		41		32	

Table 3.3: Summary Statistics for Pretest Questions

	Michigan		Detroit		Los Angeles		San Diego	
Variable	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Question 1	0.56	0.50	0.42	0.49	0.42	0.49	0.42	0.49
Question 2	0.80	0.40	0.73	0.44	0.69	0.46	0.76	0.43
Question 3	0.41	0.49	0.42	0.49	0.42	0.49	0.49	0.50
Question 4	0.37	0.48	0.31	0.46	0.34	0.47	0.36	0.48
Question 5	0.68	0.46	0.59	0.49	0.59	0.49	0.70	0.46
Question 6	0.54	0.50	0.44	0.50	0.42	0.49	0.47	0.50
Question 7	0.70	0.46	0.56	0.50	0.53	0.50	0.62	0.49
Question 8	0.58	0.49	0.43	0.50	0.48	0.50	0.52	0.50
Question 9	0.76	0.43	0.63	0.48	0.62	0.49	0.70	0.46
Constructed Response Q1	1.72	0.70	1.55	0.75	1.51	0.74	1.66	0.78
Estimated Pretest Ability	0.22	0.94	-0.17	1.02	-0.22	0.98	0.07	1.03
N	2,114		616		2,170		1,644	

Note: The questions on the pretest are about scientific achievement. The estimated pretest ability is an estimate of students' scientific ability based on Questions 1-9.

Table 3.4: Main CESE Treatment Effect

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES						
Treatment	0.175	0.328**	0.238**	0.300*	0.287*	0.359**
	(0.110)	(0.115)	(0.0820)	(0.121)	(0.111)	(0.107)
Chemistry	0.0145	0.139	0.0887	0.178	0.198+	0.243*
	(0.0950)	(0.115)	(0.0872)	(0.120)	(0.112)	(0.102)
Pretest score			0.371**	0.293**	0.261**	0.267**
			(0.0300)	(0.0227)	(0.0218)	(0.0224)
Female					-0.0527+	-0.0525
					(0.0309)	(0.0330)
Hispanic					-0.219**	-0.222**
					(0.0567)	(0.0580)
Black					-0.214*	-0.258**
					(0.0946)	(0.0970)
Asian					0.0121	0.0553
					(0.0729)	(0.0615)
Other					-0.140**	-0.160**
					(0.0492)	(0.0539)
GPA 3.0+					0.239**	0.249**
					(0.0343)	(0.0352)
Mother Educ (HS or less)					0.0241	0.0336
					(0.0380)	(0.0371)
Mother Educ (>HS Degree)					0.0618	0.0531
					(0.0560)	(0.0558)
Teacher experience						0.00768
						(0.00634)
Female teacher						-0.0562
						(0.129)
Constant	-0.0995	-0.351**	-0.128	-0.282*	-0.241+	-0.251
	(0.0985)	(0.115)	(0.0905)	(0.119)	(0.139)	(0.187)
School fixed effects		x		x	x	x
Observations	5,968	5,968	4,236	4,236	4,182	3,942
R-squared	0.008	0.143	0.144	0.221	0.240	0.247

Note: **, *, and + correspond to significance at the .01, .05, and .1 level. Data from the CESE 2018-2019 school year. Dependent variable is normalized post-test performance, with mean 0 and standard deviation 1 for all scored post-tests. Standard errors are clustered at the school level.

Table 3.5: Teacher-Level Attrition Results

	(1)	(2)
VARIABLES	LPM	Probit
Treatment	0.172	1.892***
	(0.0775)	(0.551)
Average pretest score	0.0449	-1.012**
	(0.0394)	(0.425)
Treatment x average pretest score	0.119	1.854***
	(0.183)	(0.636)
Chemistry	0.0531	0.288
	(0.0783)	(0.487)
Teacher experience	-0.00422	-0.0157*
	(0.00204)	(0.00881)
Constant	-0.0301	-3.467***
	(0.0845)	(0.787)
Average partial effect of treatment		0.178***
		(0.0369)
Observations	118	118
R-squared	0.228	

Note: **, *, and + correspond to significance at the .01, .05, and .1 level. Standard errors are clustered at the region (MI, DET, LA, SD) level.

Table 3.6: CESE Intervention Attrition Statistics

Panel A: School Level

	Overall	Treatment	Control	Differential
Initial Schools	70	36	34	
Final Schools	61	30	31	
Attrition	12.86%	16.67%	8.82%	7.84%

Panel B: Student Level

	Overall	Treatment	Control	Differential
Initial Students	6211	3325	2886	
Final Students	4238	2127	2111	
Attrition	31.77%	36.03%	26.85%	9.18%

Source: Schneider, Krajcik, Joseph, et al. (2021).

Table 3.7: Student-Level Attrition Results

	(1)	(2)	(3)	(4)
VARIABLES				
Treatment	0.0415	0.0609	0.168	0.168
	(0.0435)	(0.0474)	(0.136)	(0.136)
Pretest score	-0.0266	-0.00966	-0.0394	-0.0394
	(0.0177)	(0.0156)	(0.0496)	(0.0496)
Treatment x pretest score	-0.00421	-0.0115	-0.0316	-0.0316
	(0.0261)	(0.0225)	(0.0680)	(0.0680)
Chemistry	0.0654	0.0131	0.0540	0.0540
	(0.0534)	(0.0417)	(0.131)	(0.131)
Female		0.00422	0.0103	0.0103
		(0.0119)	(0.0342)	(0.0342)
Hispanic		-0.0110	0.120	0.120
		(0.0527)	(0.105)	(0.105)
Black		0.144*	0.311*	0.311*
		(0.0835)	(0.167)	(0.167)
Asian		0.00809	0.0964	0.0964
		(0.0747)	(0.188)	(0.188)
Other		0.0216	0.0728	0.0728
		(0.0581)	(0.123)	(0.123)
GPA (3.0+)		-0.0554+	-0.214+	-0.214+
		(0.0206)	(0.0636)	(0.0636)
Mother's educ (less than HS)		-0.0394*	-0.177+	-0.177+
		(0.0211)	(0.0598)	(0.0598)
Mother's educ (HS+)		0.00281	-0.0391	-0.0391
		(0.0231)	(0.0669)	(0.0669)
cteachExp		-0.00117	-0.00472	-0.00472
		(0.00241)	(0.00761)	(0.00761)
Female teacher		0.0575	0.189	0.189
		(0.0383)	(0.120)	(0.120)
Constant	0.204+	0.231+	-0.643+	-0.643+
	(0.0523)	(0.0598)	(0.182)	(0.182)
APE				0.0554
				(0.0429)
Observations	5,771	5,408	5,408	5,408
R-squared	0.016	0.027		

Note: **, *, and + correspond to significance at the .01, .05, and .1 level. Standard errors are clustered at the region (MI, DET, LA, SD) level.

Table 3.8: Lee Bounds

	All students		With pretest		With pretest and survey	
Bound on Treatment effect						
Lower	0.029	0.064	0.154**	0.172**	0.113*	0.114+
	(0.035)	(0.044)	(0.056)	(0.061)	(0.050)	(0.063)
Upper	0.368**	0.343**	0.240**	0.305**	0.449**	0.366**
	(0.041)	(0.050)	(0.062)	(0.065)	(0.056)	(0.064)
"Tightened"		X		X		X
N total	11,804	11,804	9,320	9,320	9,925	9,925
N with posttest	5,968	5,968	4,236	4,236	3,602	3,602

Note: **, *, and + correspond to significance at the .01, .05, and .1 level."Tightened" bounds trim means of control and treatment students at the regional level. The bounds will only be tighter if the control group attrits less in all regions relative to the treatment group.

REFERENCES

REFERENCES

Aaronson, Daniel, Lisa Barrow, and William Sander (Jan. 2007). "Teachers and Student Achievement in the Chicago Public High Schools". en. In: *Journal of Labor Economics* 25.1, pp. 95–135. ISSN: 0734-306X, 1537-5307.

Altonji, Joseph G. (1995). "The effects of high school curriculum on education and labor market outcomes". English. In: *The Journal of Human Resources; Madison* 30.3. Num Pages: 30 Place: Madison, United States, Madison Publisher: University of Wisconsin Press, p. 409. ISSN: 0022166X.

Bell, Stephanie (Jan. 2010). "Project-Based Learning for the 21st Century: Skills for the Future". en. In: *The Clearing House: A Journal of Educational Strategies, Issues and Ideas* 83.2, pp. 39–43. ISSN: 0009-8655, 1939-912X.

Blumenfeld, Phyllis C. et al. (June 1991). "Motivating Project-Based Learning: Sustaining the Doing, Supporting the Learning". en. In: *Educational Psychologist* 26.3-4, pp. 369–398. ISSN: 0046-1520, 1532-6985.

Cartwright, T. J. and B. Hallar (Feb. 2018). "Taking risks with a growth mindset: long-term influence of an elementary pre-service after school science practicum". en. In: *International Journal of Science Education* 40.3, pp. 348–370. ISSN: 0950-0693, 1464-5289.

Chan, Nixon and Peter E. Kennedy (2002). "Are Multiple-Choice Exams Easier for Economics Students? A Comparison of Multiple-Choice and "Equivalent" Constructed-Response Exam Questions". In: *Southern Economic Journal* 68.4. Publisher: Southern Economic Association, pp. 957–971. ISSN: 0038-4038.

Chen, Cheng-Huan and Yong-Cih Yang (Feb. 2019). "Revisiting the effects of project-based learning on students' academic achievement: A meta-analysis investigating moderators". en. In: *Educational Research Review* 26, pp. 71–81. ISSN: 1747-938X.

Cutler, David and Adriana Lleras-Muney (July 2006). *Education and Health: Evaluating Theories and Evidence*. en. Tech. rep. w12352. Cambridge, MA: National Bureau of Economic Research, w12352.

Deaton, Angus and Nancy Cartwright (Aug. 2018). "Understanding and misunderstanding randomized controlled trials". en. In: *Social Science & Medicine*. Randomized Controlled Trials and Evidence-based Policy: A Multidisciplinary Dialogue 210, pp. 2–21. ISSN: 0277-9536.

Dee, Thomas S. (Aug. 2004). "Are there civic returns to education?" en. In: *Journal of Public Economics* 88.9-10, pp. 1697–1720. ISSN: 00472727.

Gwet, Kilem L. (Sept. 2014). *Handbook of Inter-Rater Reliability, 4th Edition: The Definitive Guide to Measuring The Extent of Agreement Among Raters*. en. Advanced Analytics, LLC. ISBN: 978-0-9708062-8-4.

Krueger, Alan B. and Diane M. Whitmore (Jan. 2001). "The Effect of Attending a Small Class in the Early Grades on College-test Taking and Middle School Test Results: Evidence from Project Star". en. In: *The Economic Journal* 111.468, pp. 1–28. ISSN: 0013-0133, 1468-0297.

Lee, David S (2009). "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects". en. In: *Review of Economic Studies*, p. 32.

Machin, Stephen, Sandra McNally, and Olmo Silva (July 2007). "New Technology in Schools: Is There a Payoff?" en. In: *The Economic Journal* 117.522, pp. 1145–1167. ISSN: 0013-0133, 1468-0297.

Maestrales, Sarah et al. (n.d.). "Using Machine Learning to Score Multi-dimensional Assessments of Chemistry and Physics". In: *Under review* ().

Nye, Barbara, Spyros Konstantopoulos, and Larry V. Hedges (Sept. 2004). "How Large Are Teacher Effects?" In: *Educational Evaluation and Policy Analysis* 26.3. Publisher: American Educational Research Association, pp. 237–257. ISSN: 0162-3737.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain (2005). "Teachers, Schools, and Academic Achievement". en. In: *Econometrica* 73.2, pp. 417–458. ISSN: 1468-0262.

Rockoff, Jonah E (Apr. 2004). "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data". en. In: *American Economic Review* 94.2, pp. 247–252. ISSN: 0002-8282.

Schneider, Barbara, Joseph Krajcik, et al. (2020). *Learning Science - the Value of Crafting Engagement in Science Environments*. en. Yale University Press. ISBN: 978-0-300-22738-3.

Schneider, Barbara, Krajcik, Joseph, et al. (2021). "Improving Science Achievement - Is it possible? Evaluating the Efficacy of a High School Chemistry and Physics Project-Based Learning Intervention: Crafting Engaging Science Environments". In: *Under review.*

Tauchmann, Harald (Dec. 2014). "Lee (2009) Treatment-Effect Bounds for Nonrandom Sample Selection". In: *The Stata Journal* 14.4, pp. 884–894. ISSN: 1536-867X.

Walstad, William B. and Denise Robson (Jan. 1997). "Differential Item Functioning and Male-Female Differences on Multiple-Choice Tests in Economics". en. In: *The Journal of Economic Education* 28.2, pp. 155–171. ISSN: 0022-0485, 2152-4068.