A Head Start for the Whole Family: Assessing the Labor Supply Response of Mothers of Head Start Participants

Anna Russo*
Advised by: Joseph G. Altonji

April 2017

Abstract

Childcare programs have the potential to benefit both children, through cognitive development, and parents, by lessening the burden of childcare. This paper is one of the first to study the impact of Head Start, a preschool program serving nearly one million low-income children in the United States, on the parents of participants. Using discontinuities in the eligibility requirements of Head Start, I provide the first evidence that Head Start has a positive causal effect on the labor market participation of the mothers of Head Start children. I find that Head Start increases mothers' hours worked per week, probability of employment, and probability of an income increase. These effects are large in magnitude and particularly strong for black families who exhibit greater take-up rates of the program.

*I would like to express my gratitude to my adviser, Professor Altonji, for his patience, expertise, encouragement, and trust. I also thank Marianne Bruins and Robert Moffitt for helpful comments, and Jennifer Cassidy-Gilbert at the BLS for her help obtaining the geocoded NLSY79 dataset. Finally, I thank my professors at Yale who taught me to think critically and creatively; this research would not be possible without them. All errors are my own.

1 Introduction

"In today's economy, when having both parents in the workforce is an economic necessity for many families, we need affordable, high-quality childcare more than ever. It's not a nice-to-have – it's a must-have. So it's time we stop treating childcare as a side issue, or as a women's issue, and treat it like the national economic priority that it is for all of us."

– President Obama, State of the Union Address, January 20, 2015

Government-provided education is well-understood as a critical service to foster equality of opportunity and invest in the human capital of the future labor force. However, while public K-12 schooling has been offered since Massachusetts established a law granting the right to school attendance free of charge in 1827, the role of the US government in childcare services and preschool remains an unsettled debate. Discussions regarding childcare policy hinges on the parents' role in raising young children, making early childhood education policy unique as it involves, to a far greater extent than K-12, both parents and children. In this respect, childcare policy serves a dual purpose, investing in both the current and future generations of productive workers. Effective childcare policy gives children the developmental support necessary to launch a successful academic career and parents the opportunity to shift time spent from informal childcare to formal gainful employment.

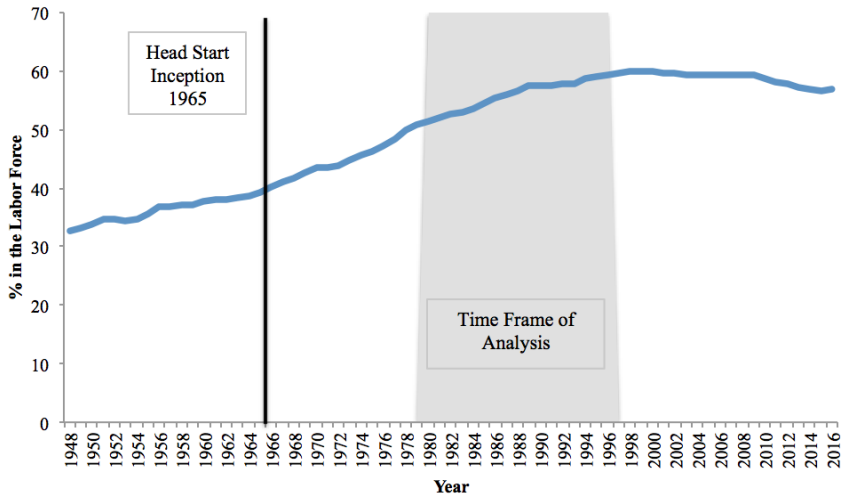
While no universal public pre-Kindergarten or childcare programs exist in the United States, a number of subsidies and smaller scale programs provide some services to families. These include the Child and Dependent Care Tax Credit (CDCTC), state-funded universal pre-K programs in select states,[1] and Head Start, a nationally-funded preschool program for low-income children. Head Start, which has served over 33 million families since 1965 and has a current operational budget of over 8 billion dollars, represents the nation's greatest effort to provide early childhood education services to American families.

Head Start's impact on the immediate and longer-term development of participating children has been the subject of much research and debate, but Head Start's additional, implicit purpose as a work-life balance policy for families has remained largely unstudied. As Figure 1(a) illustrates, the female labor force participation rate has increased dramatically since Head Start's inception. This demonstrates a major shift in labor force composition and the relevance of childcare prices and policy to the structure of the U.S. labor force as a lack of affordable childcare restricts parents' abilities to work. However, even with these restrictions, Figure 1(b) shows that the population of women with young children surpassed the average female in labor force participation in the early 1990s, demonstrating an increasing number of women balancing work with childcare. This paper seeks to understand the role of Head Start for parents in the context of this changing

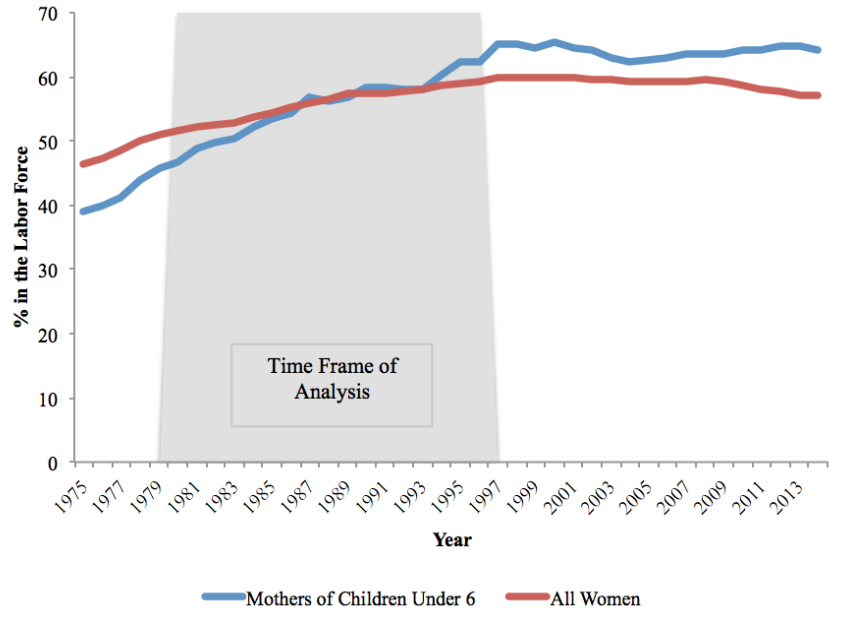
[1] Florida (since in 2005), Georgia (since in 1995), and Oklahoma (since 1998) all have state-funded universal pre-Kindergarten programs. Forty-one states (and Washington DC) have some form of state-funded childcare, but they only serve a small fraction of four-year-old children in the state.

Figure 1: Female Labor Force Participation Rates

(a) Total Female Labor Force Participation Rate



(b) Female Labor Force Participation Rate by Age of Child



Data in (a) from the Current Population Survey (CPS) administered by the U.S. Bureau of Labor Statistics. Data retrieved from FRED, Federal Reserve Bank of St. Louis. Data in (b) from the Bureau of Labor Statistics *Women in the Labor Force: A Databook 2015*.

dynamic of female labor force participation: to what extent does Head Start play a role in changing parents' labor market behavior?

Building upon Carneiro and Ginja's (2014) study of Head Start's impact on participating children, this paper uses their identification strategy to provide the first quasi-experimental study of Head Start's labor market effects and presents the first evidence that Head Start does indeed induce mothers to enter the work force. Using state-varying eligibility requirements for Head Start applied to the geocoded versions of the NLSY79 and CNLSY79, I find significant evidence of labor market changes as mothers increase the number of hours worked per week, move into the labor force, and experience an accompanying rise in earned income. These findings are a contribution to the child development-focused Head Start impact-evaluation as well as an addition to the literature studying mothers' labor supply responses to childcare prices as I provide new insight on the labor market responses of low income mothers to large subsidies.

The remainder of the paper will proceed as follows. Section 2 provides a brief background of Head Start and the state of early childhood education in the United States. Section 3 presents a simple theoretical framework to ground and motivate the empirical analysis, followed by Section 4, which contextualizes these arguments within the environment of existing empirical literature. Section 5 explains the identification strategy, Section 6 summarizes the data sources used in the analysis, and Section 7 presents results. I conclude in Section 8.

2 Head Start and American Childcare

In 2015, childcare costs averaged an annual $9,589 across the nation, 17% of the median household income of $56,516. This exceeds the benchmark 10% share of household income the Department of Health and Human Services recommends to be devoted to childcare expenses. Among working mothers with children below the poverty line, private childcare costs an average of 30% of working income (Shulte and Durana 2016). To address these high costs, the government has initiated two federal programs: the Child and Dependent Care Tax Credit and Head Start.

The Child and Dependent Care Tax Credit (CDCTC) is a tax credit available to all parents with an income tax liability and dependents. It reimburses up to 35%, with the percent reimbursement decreasing in family income, of childcare costs under $3,000 for one qualifying dependent and $6,000 for two or more.[2] The CDCTC can be claimed for any care, both in-home and out-of-home, as long as the guardian is working or seeking employment. Because the CDCTC is available to families with an income tax liability, the CDCTC serves everyone except the very-poor who pay no income tax. The CDCTC is a flexible credit, paired with

[2] $3000 is 70% of the childcare costs in Mississippi, the least expensive state, 31% of the national average, and 18% of the childcare costs in Washington, DC, the most expensive "state" (Child Care Aware of America 2016).

a working requirement, that is available to almost all families for almost all childcare purposes.

Head Start, the other major US childcare program, differs substantially from the CDCTC on a number of grounds. Head Start provides primarily center-based[3] free early childhood education programs to 3-5 year old children[4] for the amount of time the child is at the center. The exact number of hours provided is heterogeneous at the level of each Head Start center, but can range from 18 to over 35 hours per week (HHS 2015). Head Start is a means-tested program designed to serve low-income families: 90% of all spots in Head Start classrooms are reserved for children who meet a set of poverty guidelines.[5] To be eligible for participation, a family must either have a total income below a threshold or receive public assistance (eligibility requirements will be discussed in greater depth in Section 5 as they form the basis of the empirical strategy). The income metric that a family reports is flexible; they can report their income from the calendar year prior to a child's enrollment or an alternative income period that best illustrates a family's current situation. Not all children who are eligible participate, as there are are insufficient resources to provide services to all low-income children. In 2012, even after the large enrollment and funding increases exhibited in Figures 2(a) and 2(b), only 42% of all eligible children were served by Head Start programming (Schmidt et al. 2013)

Unlike many other government assistance programs like the CDCTC or TANF (Temporary Assistance for Needy Families), Head Start eligibility is not accompanied by a work requirement. While Head Start could perhaps serve as a nudge into the work force, the lack of a work requirement makes Head Start's labor market impact more uncertain and far more interesting to study. In addition, while many means-tested programs require monthly or even real-time re-evaluation of eligibility,[6] Head Start families remain eligible for two years following their initial determination, regardless of any changes in family income that could alter eligibility status (HHS 2007). Because this two-year grace period is longer than the time that most children stay in Head Start, Head Start parents are free to change their income earning behaviors, perhaps in the form of labor force movements, and still remain in the program.

Since its inception in 1965, Head Start has provided childcare services for over 33 million children with a mandate to improve children's educational, social, health and nutritional development. Its goals have also included parental assistance and support; in the 1965 White House announcement, one of five stated goals was "to assist the parents as well as the children" (Johnson 1965). Today, early learning, health, and family well-being continue to make up the three primary services Head Start offers. Head Start's funding and scope to achieve these goals have grown substantially: over the past 50 years, Head Start has expanded

[3]In the 1980s and 1990s, the time frame of the analysis in this paper, center-based, part-day programs made up the majority of Head Start services (HHS 2013). While home-based services have grown in popularity over time, center-based services still constitute 96% of all Head Start programs in 2015 (HHS 2015).

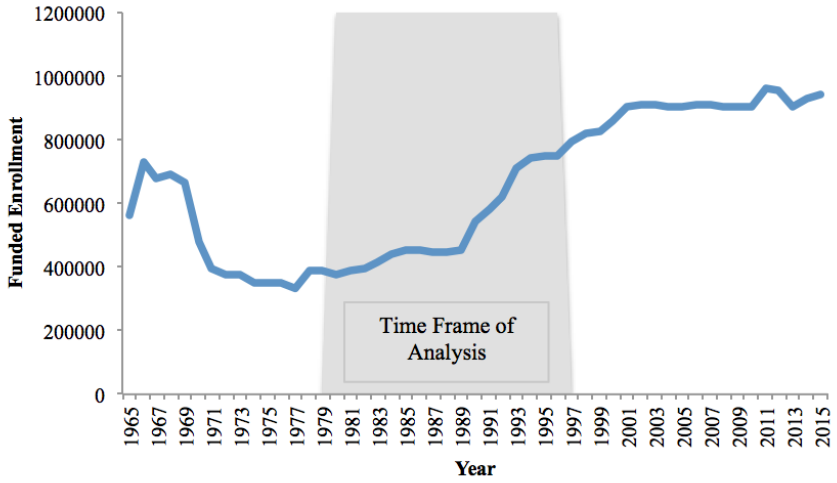
[4]Early Head Start began to offer programs to 0-3 year old children and pregnant women in 1994. However, 3-4 year old children still make up 80% of Head Start's participants today (HHS 2015)

[5]The remaining 10% of spots may go to children whose families are above these guidelines but below 135% of the income threshold.

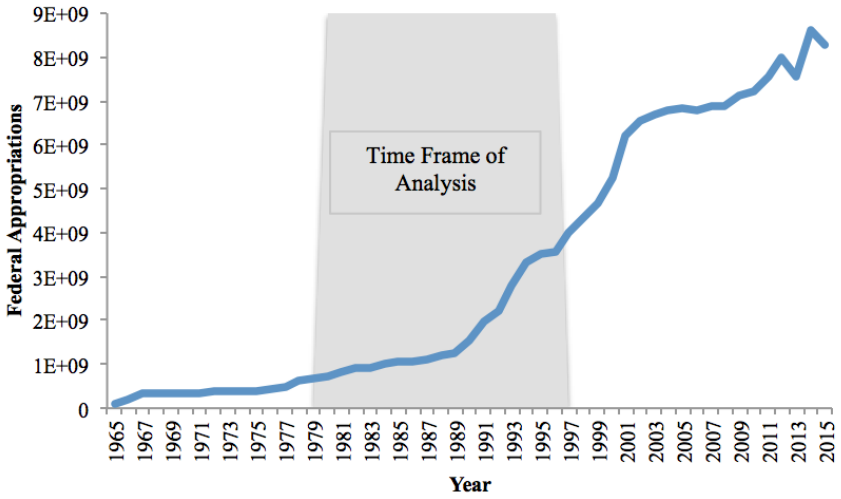
[6]TANF, for example, requires notification of any changes in earned income of magnitudes larger than $100.

Figure 2: Head Start Program Growth

(a) Head Start Funded Enrollment



(b) Head Start Appropriations



Data from the Department of Health and Human Services Head Start Fact Sheet 2015.

to serve children through 1,700 agencies in every state, with participation nearly doubling from 450,000 children in 1989 to over 850,000 in 2000. This increase in enrollment was accompanied by significant growth in federal appropriations; today, federal Head Start expenditures have reached over 8 billion dollars.[7] Such a large program certainly merits careful evaluation to understand the extent to which Head Start meets its responsibilities.

3 A Simple Model

Introducing a one-person labor-leisure choice model (Pencavel 1986) enriched to include childcare costs (Blau 2003) provides a framework with which to proceed in the analysis. Consider a mother, who we treat as the primary caregiver. The mother is endowed with one unit of time, which she can choose to allocate to wage-earning activities, h, or non-wage-earning activities, l. Time spent in non-wage-earning activities represents time spent both consuming leisure as well as time spent informally caring for a child. We assume further that a child cannot be left alone without care and that a child is cared for by the mother during time allocated to l but not cared for during any time allocated to h.

A mother generates income I from her wage-earning activities with wage w. Absent free childcare services or alternative arrangements, a mother must also pay the hourly price of childcare, p, during any hours that she chooses to work. A mother's income in the absence of any subsidized childcare is:

$$I = h(w - p) \tag{1}$$

The mother chooses h to maximize her utility of consumption and leisure:

$$\begin{aligned} \max_{h} \quad & u(c, l) \\ \text{s.t.} \quad & 0 \leq h \leq 1 - l \\ & c = I + y \end{aligned} \tag{2}$$

where c is consumption and y is income from sources other than the mother's wage-earning activities. Note that childcare costs necessarily reduce the marginal utility of each hour worked because each hour produces strictly less income to consume. In addition, by allowing l to represent both the leisure good and time spent caring for a child, we assume that a mother values caring for her child.

If we alter the price of the childcare by adding a constant s such that $|s| \leq p$, we change the constraint to

$$I = h(w - p + s) \tag{3}$$

[7]These funding increases have driven both an increase in child participation as well an increase in investment per child. Federal cost per year per child has increased from $5,100 (2015 dollars) in 1993 to roughly $7,000 in 2015 (HHS 2015).

where we can interpret $s > 0$ as a price decrease or government subsidy and $s < 0$ as an increase in the price of childcare. As s changes, the marginal utility of working with respect to leisure changes.

Figure 3: Labor Supply Shifts with a Childcare Subsidy

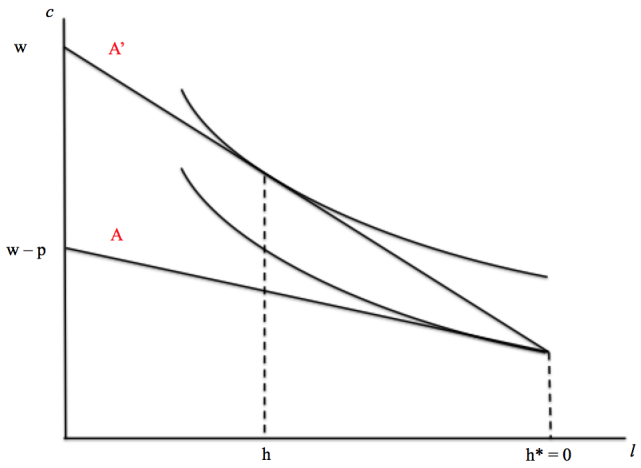


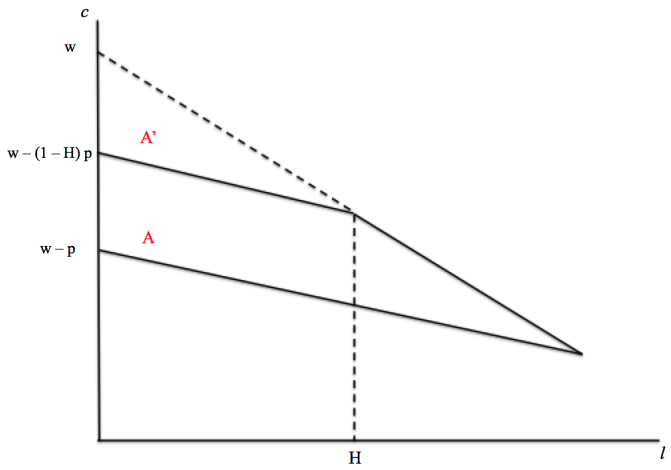
Figure 3 illustrates graphically how a simple subsidy can change the optimal choice of h. Budget constraint A represents a mothers budget set without any form of subsidy; if the mother spends all of her available time working, she earns a wage w minus the cost of childcare p. The mother's indifference curve, representing her preferences, is tangent to the unsubsidized budget set A at the point $h^* = 0$; she chooses not to enter the labor force. Under a full government subsidy, where $s = p$, the budget constraint shifts to A'. At this steeper budget constraint, the substitution effect unambiguously induces the mother to work $h \geq h^*$ hours.[8] For a mother that is already working, the effect of a childcare subsidy is ambiguous, as either the income or substitution effect may dominate. Much of the empirical literature on childcare subsidies focuses on the elasticity of labor supply with respect to changes in the cost of childcare.

Head Start can be interpreted as a childcare subsidy with $s = p$ for all hours that Head Start is provided. Because Head Start provides a set of hours free, rather than reducing the price for all hours a mother chooses to work, one must model the single-agent labor supply model slightly differently and introduce a kink into the budget set. This motivates empirical analysis of Head Start not only as an impact evaluation, but also as an opportunity to shed new light on a 100% subsidy for a fixed period of time, a deviation from the commonly studied simple price change (Gelbach 2002).

The budget set for Head Start is illustrated in Figure 4, where constraint A represents the budget set in the absence of Head Start, and A' the constraint with Head Start. For a given number of hours that Head Start programming is available, H, Head Start acts as a complete childcare subsidy, and the budget constraint has a slope of w. There is a kink at H: at hours worked beyond H, Head Start is no longer

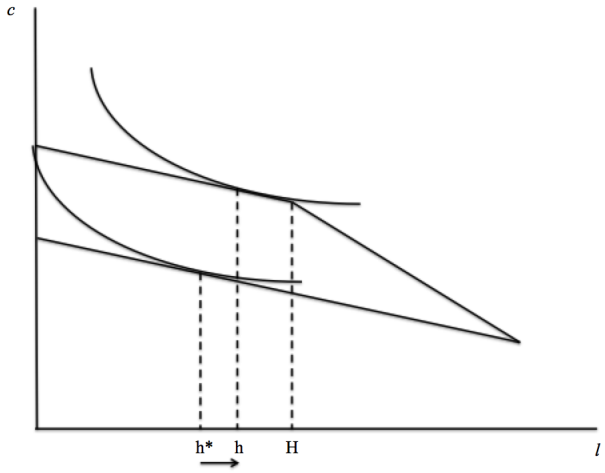
[8] In the illustration in Figure 3, the mother enters the work force. However, it may still be the case that $h = h^* = 0$.

Figure 4: The Budget Constraint Under Head Start



available and the slope of the budget set reverts to the slope of $w - p$, as the mother pays for childcare in excess of what Head Start can provide.

Figure 5: The Income Effect Dominates for Initial High Suppliers



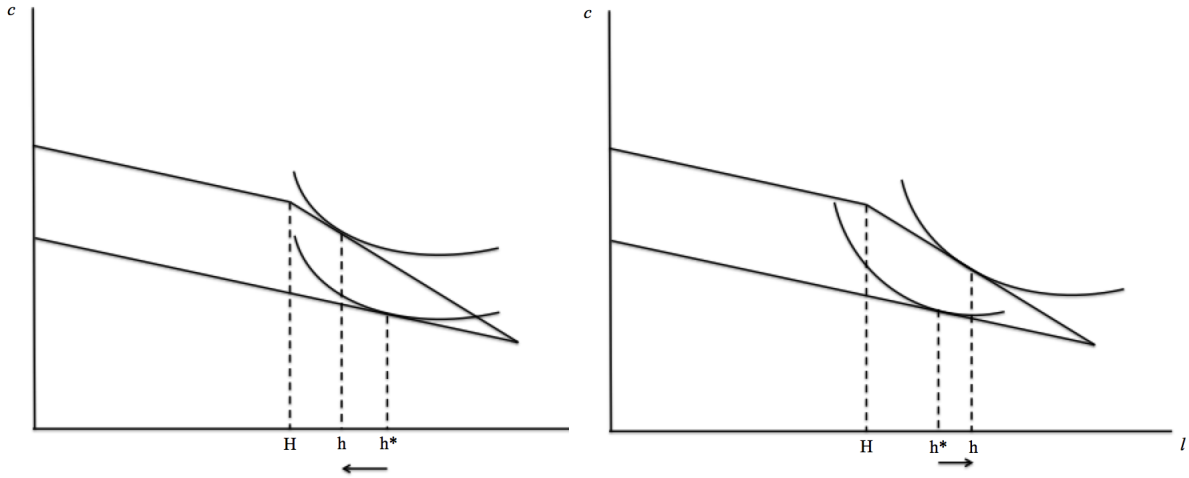
Let h^* be the number of hours that a mother would have worked taking the price of childcare as given with no alternative arrangements, earning an effective wage of $w - p$. If $H < h^*$, then the mother chooses h such that income becomes:

$$I = Hw + (h - H)(w - p) = h(w - p) + Hp \tag{4}$$

where the mother works for H hours at a full wage while receiving free childcare and chooses $h - H$ additional

hours[9] at an effectively diminished wage rate. In this representation, Head Start would provide no nudge into the labor force, as the program acts as a lump-sum transfer of Hp that should not change her marginal utility of labor with respect to leisure. The lump-sum transfer would act entirely as an income effect, reducing employment hours. This is illustrated in Figure 5, where the mother is a sufficiently high initial labor market supplier that the slope of the budget set does not change, it only shifts outward by Hp. It is important to note that this representation considers a case in which the mother has full access to childcare, was previously willing to pay for childcare at the given price, and can instantaneously and without cost move children between Head Start programming and some costly programming.

Figure 6: Competing Income and Substitution Effect for Low Suppliers



(a) Substitution Effect

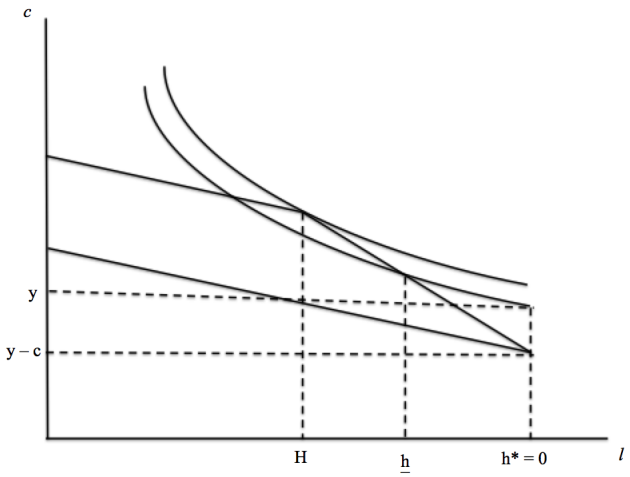
(b) Income Effect

Consider instead the case in which $H > h^*$, or the number of hours that Head Start provides free childcare exceeds the number of hours a parent would have worked and paid for childcare otherwise. Under a concave utility function, we know that h, the optimal number of hours worked with Head Start, should be less than or equal to H. However, in the case where $H > h^*$, the marginal utility of labor changes as the effective wage increases from $w - p$ to w, and Head Start may induce mothers to increase h up to the level of H if the substitution effect dominates the income effect. Figure 6 illustrates the ambiguity of this case. For low labor market suppliers, the analysis of a labor response to Head Start is analogous to the simple childcare price change model, but with a kink at $h = H$, where H represents an upper bound on the theoretical increase in the number of hours worked.[10] While we would expect only an income effect for mothers who were already willing to work many hours ($h^* > H$) while paying for childcare, a substitution effect may dominate and induce a mother to work $h > h^*$ if h^* is sufficiently low, and H and p are sufficiently high.

[9]Note that $h - H$ could be zero, where a mother chooses to work at the kink in the budget constraint.

[10]Even if the mother may be induced into working more hours to such an extent that she would work additional hours above H, the mother will not value consumption highly enough to choose to work at $w - p$ for $h > H$.

Figure 7: Labor Force Changes with A Fixed Cost of Employment



To enrich this model slightly, I introduce a simple fixed monetary cost of employment[11] (Cogan 1980). These fixed costs may be associated with transportation or incurred to search for and successfully find a job or childcare arrangements. Figure 7 illustrates the budget constraint and preferences incorporating a fixed monetary cost of employment. If a mother chooses to work zero hours, she has a consumption level of y from non-wage income. If she chooses to enter the labor force, she incurs a fixed cost of c. In this illustration, in the absence of Head Start, the mother chooses to work $h^* = 0$ hours, where Head Start may induce a mother to work up to H hours. However, with a fixed cost, working at all levels $h < \underline{h}$ is still inferior to remaining out of the labor force. Thus, under the more realistic assumptions of a fixed monetary cost of employment, any inducement into the labor force should be large in magnitude to offset the fixed costs of employment.

Finally, I must address an important complication: Head Start is a means-tested program. The analysis of many means-tested programs suffers from discontinuous constraints due to constant eligibility re-evaluation. However, the laws governing eligibility for Head Start rule out this possibility and allow for the empirical analysis in the remainder of the paper. To understand the key points in the Head Start legislation that allow for this analysis, I briefly present the Head Start model with regular means-testing.

Let $\overline{I}$ be the maximum income that a mother can earn while still qualifying for Head Start. Thus, if wage earning potential is sufficiently high, there exists some $\overline{h}$ such that:

$$\overline{I} = \overline{h}w + y \tag{5}$$

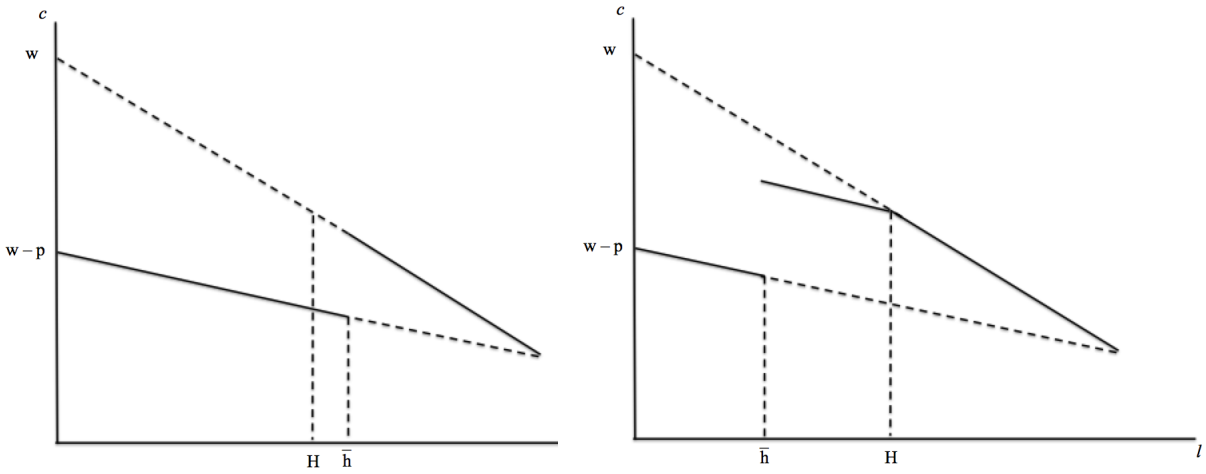
[11] A time cost of employment would have a similar effect: women who are induced to enter the labor force as a result of Head Start will reduce their leisure time to a greater degree than in the model without any fixed costs. However, incurring a time cost means that some of the additional movements out of leisure time are not being spent earning wages in the labor force, but are rather being spent on the fixed time cost of employment.

For lower wage earners:

$$\overline{I} = Hw + (\overline{h} - H)(w - p) + y \tag{6}$$

The discontinuity in the budget constraint caused by means-tested nature is illustrated in Figures 8(a) and 8(b), respectively.

Figure 8: A Means Tested Budget Constraint



(a) High Wage Earner

(b) Low Wage Earner

Consider a mother who initially works h^* such that $h^* < \overline{h}$ and $h^* < H$. Because $h^* < H$, the substitution effect may dominate, and the mother may work $h > h^*$ after receiving Head Start. However, under a standard means-tested program, if $h > \overline{h}$ in the next income report, the mother would become ineligible for Head Start, and her budget constraint would drop discontinuously to a line with slope $w - p$, inducing her to reduce her number of hours worked. While in this case it would be optimal to work $\overline{h}$ hours, the complicated nature of qualifying for means-testing makes it essentially impossible for a mother to manipulate her labor supply to work at exactly $\overline{h}$. Thus a woman may be constantly moving across the discontinuity in the budget constraint as Head Start changes her labor supply and eligibility status. This not only makes it difficult to measure the effects of means-tested programs on employment, but could also significantly limit Head Start's theoretical effectiveness in inducing parents to work.

However, two elements of Head Start's eligibility requirements in tandem play a key role in preventing this situation. First, Head Start eligibility is initially determined prior to enrollment, meaning that eligibility will always be determined based upon labor market choices of h^* under the $w-p$ budget constraint. Second, Head Start is only initially means tested; once a family is determined eligible, a child is able to stay in the program for two years. Because this paper focuses on Head Start eligibility at age four, this two year period is longer than the length of possible Head Start participation. Thus, Head Start is effectively only means-tested once on the $w - p$ budget constraint and children remain eligible for the remaining time in the

program regardless of changes in h with respect to $\overline{h}$. This implies that the budget set is in fact the simple budget set illustrated in Figures 4 - 7, and not the means-tested budget constraints illustrated in Figure 8.

However, the means-tested nature of Head Start still plays an important role in considering the potential labor market effects of Head Start. Because h^* must be less than $\overline{h}$ for a family to qualify, parents are far more likely to be low suppliers ($h^* < H$), and thus on the part of the budget constraint where the substitution effect may dominate. Because Head Start is targeting low labor market suppliers, the program may have a differentially large impact on mothers' labor supply decisions.

4 Empirical Literature Review

Head Start's success in promoting child cognitive development, health, and school readiness has been extensively studied. However, a consensus on the merits of Head Start on improving the lives of participants has not yet been reached. The most influential and rigorous of these studies is the Head Start Impact Study (HSIS), the only randomized evaluation of Head Start to date. The HSIS attempted to quantify the impact of Head Start on a cohort of entering three year old and four year old children, collecting data from the year of entrance, in 2002, through first grade, in 2006. The HSIS study found that children who were assigned to Head Start experienced significantly improved cognitive skills relative to their non-Head Start peers, but by the first grade, only a single cognitive impact remained. The modest results found in the HSIS on the long-term impacts of Head Start have sparked criticism of the program.

In addition to the HSIS, numerous studies using quasi-experimental methods have attempted to answer the question of Head Start's impact with mixed findings (Currie and Thomas 1995, 1999, Currie, Garces, and Thomas 2002, Ludwig and Miller 2007, Deming 2009). Carneiro and Ginja's (2014) analysis, which this paper builds upon, uses discontinuities in eligibility requirements to find long term reductions in childhood obesity and criminal behavior. These studies use convincing identification strategies to account for the selection problems that arise when attempting to quantify the impacts of Head Start, but they focus exclusively on the impact of child participants rather than any labor market outcomes of the parents.

However, a body of literature on labor supply responses to childcare subsidies motivates inquiry into extending Head Start's impact beyond child development. Many studies have estimated the elasticity of maternal labor supply with respect to childcare costs, although the extent of the change varies from 0.6 to magnitudes greater than -1.0 (Anderson and Levin 2000, Baum 2002, Blau and Robins 1988, 1991, Connelly 1992, Connelly and Kimmel 2003, Han and Waldfogel 2001, Blau and Hagy 1998, Cleveland et al. 1996, Fronstin and Wissoker 1995, Ribar 1995, Tenkin 2007, Chaplin et. al. 1999, Michalopoulos and Robins 2002, Michalopoulos et al. 1992, Powell 1997).

Two studies are of note because they focus on programs that provide free childcare for a predetermined amount of time rather than a simple subsidy. Gelbach (2002) and Cascio (2009) both analyze the effects of free kindergarten on mothers' labor supply, a type of 100% subsidy similar to Head Start. Gelbach finds that free public kindergarten causes a 6-24% increase in labor force participation among single women whose youngest child is of kindergarten age and a 6-15% increase in labor force participation among married women, regardless of the presence of younger children in the household. Cascio reports an even larger effect, finding a 45% increase in labor force participation of mothers with no younger children for every additional child enrolled in public kindergarten.

The child's age at the time of a subsidy may play a role in labor supply decisions, as women may be more likely to enter the workforce when a child is of kindergarten age than preschool age. Baker, Gruber, and Milligan (2008) study Quebec's "$5 per day childcare," a nearly 100% subsidy introduced in the 1990s to provide inexpensive childcare to all families in Quebec, regardless of need. They find that the employment rate of married women in Quebec relative to the rest of Canada increased by 14.5%, with suggestive evidence that women were moving from no work to full-time work as a result of the policy.

While kindergarten and Quebec's universal preschool program provide services similar to Head Start, Head Start differs due to its focus on low income women. Berger and Black (1992), study low-income single mothers who are eligible for two Kentucky childcare subsidy programs in 1989. Using random assignment from the program's waitlist, Berger and Black estimate a 12% to 23.5% increase in employment as a result of the Kentucky childcare subsidy programs. Kentucky programs do not provide a 100% subsidy for a given number of hours as is the case with Head Start, but rather allow for reimbursement of up to $40 to $50 per week for childcare costs.[12] Moreover, to be eligible for the subsidy, mothers must work at least 20 hours per week and must be recertified for the program every six months (where they demonstrate that they are both still eligible and are meeting the work requirement). The bundling of a work requirement with a childcare subsidy in the Berger and Black study makes it difficult to extrapolate how a childcare subsidy without work requirements may shift incentives to induce increased labor force participation. Moreover, it is uncertain as to whether Kentucky mothers can be treated as a representative sample.

Despite evidence regarding the relevance of childcare policy to parental outcomes, this paper is one of the first to examine the impacts of Head Start on participating mothers and the first to examine these mothers using quasi-experimental methods. Sabol and Chase-Lansdale (2014) (hereafter SCL) study Head Start's impact on parents, making their analysis the closest in content to this study. SCL use information collected on the parents that participated in the Head Start Impact Study. Using the HSIS randomized design, SCL find that the parents of children who were randomly assigned to participate in Head Start at

[12]The average daily cost of childcare in Berger and Black is reported to be 11.81 dollars per day (implying costs of $59 per week for a five day week). With an average subsidy amount of $45.62, this represents a 76% subsidy. Blau and Robins (1988) report findings that for a subsidy of similar magnitude, $45.15, labor force participation increases by 43.0%. Berger and Black attribute this difference to the fact that all the women in the Kentucky study are single.

age three were more likely to increase their educational attainment by the time the child turned six than those assigned to the control group. These effects were differentially evident for black parents. In fact, the effects of Head Start were only evident for black families and not for white or Hispanic/latino families.

SCL found no effects on employment changes in either the three-year-old or four-year-old cohort. However, the HSIS was designed to study children, not parents, and therefore SCL have very few labor market indicators to use in order to analyze changes in labor supply. In fact, they only have the response of a single annual question asking parents how they spent their time in the previous week.[13] Although the authors report no changes in employment status between the treatment and control groups, they mention that differential coding of ambiguous responses could lead to statistically significant increases in labor market outcomes, leaving readers without conclusive evidence of a lack of labor market effects. The report of no effect is quite surprising in the context of the related literature on childcare subsidies, and merits further inquiry in a quasi-experimental setting. In addition, although the randomized nature[14] of the HSIS provides a controlled environment to study Head Start's effects, it does not allow for any analysis of Head Start take-up because the study randomized among parents who had already applied for Head Start. When studying Head Start in the context of parents' time-use decisions, understanding take-up plays an important role. This paper seeks to explore Head Start's impact on participating parents in far greater detail than is possible using data from the HSIS.

5 Empirical Strategy

A naive comparison of the labor force behavior of mothers who enrolled their children in Head Start and those who did not present significant concerns of selection bias. It is plausible that parents who choose to enroll a child in Head Start differ on unobservable characteristics that may also be correlated with their labor-leisure choices. While a randomized experiment like the Head Start Impact Study provides the ideal environment in which to identify the causal effects of Head Start, exogenous sources of variation can provide a similar opportunity to identify causality. The eligibility requirements of Head Start provide such an opportunity; following the strategy of Carneiro and Ginja (2014), I exploit a discontinuity in Head Start eligibility requirements to identify the causal effect of Head Start on the labor market choices of the mothers of participating children.

Eligibility for Head Start is determined if either of the following conditions are met:

[13] They had response options of: (1) working full-time (35 hours or more per week), (2) working part-time, (3) looking for work, (4) laid off from work, (5) in school/training, (6) jail or prison, (7) military, (8) keeping house, or (9) something else.

[14] The design of the HSIS has been criticized on a number of grounds (Zigler 2010), including criticism that a substantial portion of control group families enrolled their children in Head Start (18% of the three year old cohort and 14% of the four year old cohort).

1. Family income is below a determined poverty line, a threshold set by the Department of Health and Human Services as a function of household income, family size, state of residence, and year of application. Income conditions are determined in the calendar year prior to the year in which a child enrolls in Head Start or an alternative metric that "most accurately reflects the needs of the family" (HHS 2007).[15]

2. A child is eligible for public assistance (AFDC/TANF[16] or SSI.[17])

The point of discontinuity is not a single number but a function of income, state of residence, family size, year of application, and public assistance eligibility. Most children qualify via the income eligibility requirement: in the sample used in this study, 98% of the children are eligible as a result of their placement on either side of the income cutoff. This facilitates interpretation that any discontinuous labor market effects are a result of Head Start, without the danger of confounding Head Start with the receipt of public assistance. To ensure that this is indeed the case, I control for receipt of public assistance in the regression discontinuity control function and provide further identification checks that Head Start is not confounded with any other programs (see Section 7.5).

Before proceeding to describe the empirical strategy in detail, I first extend the discussion in Section 3 of the two important pieces of legislation specific to Head Start that allow for the analysis of means-tested programs. In Section 3, I discussed how (1) eligibility determination prior to enrollment and (2) a two-year eligibility period removed the danger of a discontinuous budget constraint when studying Head Start. Though initially discussed in the context of theory, these arguments have clear analogues of critical importance to the empirical strategy.

First I address the importance of eligibility determination prior to enrollment. Studying the labor-force outcomes of means-tested programs can be particularly challenging as the eligibility requirements and outcome variables are often jointly determined. To be more precise, let income, INC, be a function:

$$INC = f(h, w(X), h_{sp}, w_{sp}(X_{sp}), y) \tag{7}$$

where h and h_{sp} are the hours worked of the mother and spouse respectively. w and w_{sp} are the wages of the mother and spouse, both dependent on vectors of characteristics that determine their wage-earning potential. y is outside income. Because income determines eligibility, and labor market outcomes including

[15]The idea here is that while most families would submit earnings report from the previous calendar year (as it is the simplest option), families are also allowed to, for example, report earnings for the preceding twelve months. Because the reported income must reflect the family's current needs status, any alternative income reported for periods prior to the previous calendar year would not be an acceptable income metric. Only one source of income report is required (i.e. a family does not need to provide both the earnings from the previous calendar year as well as a current pay stub).

[16]AFDC (Aid to Families with Dependent Children) was replaced by TANF (Temporary Assistance for Needy Families) under a 1996 welfare reform. Among other changes, TANF instituted a work requirement to eligibility, while AFDC was an entitlement program.

[17]Supplementary Security Income.

hours worked, wages, employment, and labor force participation are the outcome variables of interest, it is clear how one could easily run into a problem as eligibility and labor force outcomes are simultaneously determined.

However, just as the timed nature of eligibility requirements contributes to removing the threat of a discontinuous budget set, it also removes the danger of simultaneous determination in the empirical strategy. As stated in eligibility criterion (1), Head Start eligibility is determined by family income for the calendar year prior to the year in which a child becomes eligible (at age three, four, or five) or, alternatively, using a time frame to present information that most accurately reflects the needs of the family. Due to the constraints on what is available in NLSY79 data, this paper, following Carneiro and Ginja, uses income from the previous calendar year to determine eligibility. This amounts to restricting the analysis to children who determine their eligibility using income from the calendar year prior to the year a child enrolls. However, this restriction resolves the simultaneity problem, allowing measurement of outcomes for a child at ages t, $t+1$, and $t+2$ with eligibility determined at age $t-1$.

Even if the variables are time-indexed in such a way that prevents simultaneous determination, if Head Start causes changes in labor force behavior, movement around the discontinuity would pose a problem. In Section 3, movement around $\overline{h}$ in Figure 8 is analogous to movement around the eligibility indicator employed in the empirical strategy. For a family whose income is just below the cutoff, an inducement to work n extra hours to earn an income nw higher than it was originally could be sufficient to move the family above the cutoff, rendering the child ineligible to continue participating in Head Start. In this situation, no impact would be perceptible, as a family would bounce around the discontinuity. However, rule (2), a two-year eligibility period without re-determination allows us to measure eligibility prior to attending Head Start and determine labor market outcomes during and following enrollment.

Although in theory, eligibility requirements constitute a robust identification strategy, I must first determine empirically that eligibility does indeed predict program participation. I predict participation with eligibility determined for enrollment at ages three and four, as these ages constitute 90% of all Head Start enrollment.[18] The probability of Head Start participation should increase discontinuously with eligibility status. To test this hypothesis, I estimate the equation:

$$HS_{i,t} = \alpha + \beta ELIG_{i,t} + f(X_{i,t-1}, Z_{i,t-1}) + \epsilon_i \tag{8}$$

for $t = 3$ and $t = 4$, where $HS_{i,t}$ is an indicator variable that takes the value 1 if a child participated in Head Start and 0 if not, $ELIG_{i,t}$ is an indicator variable that takes the value 1 if a child is eligible for Head Start and 0 if not, f is a flexible function of $X_{i,t-1}$, a vector of the determinants of Head Start eligibility, and $Z_{i,t-1}$, a vector of variables describing the labor force activities of the parents in the year in which Head

[18] The remaining 10% of children enroll at age five.

Start eligibility was determined. The inclusion of $Z_{i,t-1}$ provides a baseline to evaluate whether or not Head Start causes any labor market movements from the time frame during which eligibility was determined. The coefficient of interest is β, the coefficient on $ELIG_{i,t}$, the indicator variable that determines if a child is eligible for Head Start. More specifically, $ELIG_{i,t}$ is defined by:

$$ELIG_{i,t} = \begin{cases} 1 & [INC_{i,t-1} \leq g(INC_{i,t-1}, FSIZE_{i,t-1}, STATE_{i,t-1}, YEAR_{i,t-1})] or [PA_{i,t-1} = 1] \\ 0 & [INC_{i,t-1} > g(INC_{i,t-1}, FSIZE_{i,t-1}, STATE_{i,t-1}, YEAR_{i,t-1})] and [PA_{i,t-1} = 0] \end{cases} \tag{9}$$

where g is a function determining the eligibility cutoff with arguments: income, family size, state, year, and $PA_{i,t-1}$, an indicator variable that takes the value 1 if the child's family is eligible for receipt of AFDC[19]. These variables (income, family size, state, year, and receipt of public assistance), make up the vector of controls $X_{i,t-1}$ in equation (8). The coefficient of interest in (8), β, determines if Head Start eligibility positively predicts Head Start participation; β determines if eligibility requirements form a valid instrument.

The unit of observation is the child, as each child's eligibility and enrollment in Head Start is determined separately (i.e. a child is not automatically eligible if a sibling is also eligible). Moreover, a mother's labor force responses are likely to be heterogeneous at the level of each participating child, so estimating at the level of the child preserves the opportunity for this heterogeneity, allowing each child to have a flexible effect in the RD specification.

A first stage estimate of the form presented in (8) is necessary because Head Start eligibility does not necessarily imply participation. Even with substantial resources, Head Start does not have enough funding to provide programming for all its students. In 2012, Head Start provided services to 42% of eligible children (Schmidt et al 2013), and in 1980, when the first children studied in this paper began entering preschool, funding provided for as few as 20% of all eligible children (Muenchow and Shays 1980). For any number of reasons — local centers may be full, parents may decide not to enroll their children — eligibility does not imply treatment into Head Start. Moreover, the guidelines determining Head Start eligibility include a provision that 10% of all available slots in the program may be made available for children whose family income is above the threshold.[20] The possibility that parents may use an alternative income time frame to determine their eligibility, and thus may be ineligible when coding eligibility using income from the previous calendar year could be an additional factor that reduces the strength of $ELIG_{i,t}$ as an instrument to predict Head Start participation. Thus, as with any instrument, we must check to ensure that eligibility predicts participation. In section 7, I establish that the eligibility requirements form a valid instrument for eligibility

[19]TANF had not yet replaced AFDC during the years of analysis of this paper, which span from 1980-1995. SSI (Supplementary Security Income) is not included because its eligibility requirements are strictly below those of Head Start, and thus are not binding in determining an eligibility cutoff. In addition, the medical requirements determining SSI eligibility are impossible to determine given the data available in the NLSY79.

[20]However, their income must be within 135% of the family income threshold.

determined at age four but not at age three.

The first equation of interest regarding labor force outcomes is the reduced form specification:

$$L_{i,t} = \alpha + \beta ELIG_{i,4} + f(X_{i,3}, Z_{i,3}) + \epsilon_i \tag{10}$$

All the right-hand variables are identical to those in (8), but the outcome variable $L_{i,t}$ represents a labor force outcome. The time-stamp on eligibility is specified to be age four. I restrict the analysis to short and medium term labor market changes, and estimate these equations with $t = \{4, 5, 6\}$, labor market behavior with children ages four, five, and six. The labor supply outcome variables include hours worked per week, an indicator variable for employment, and weeks out of the labor force. I also examine other outcome variables related to labor supply including hourly wage and earned income. I specify these labor market outcome variables in greater detail when discussing results in Section 7.

Because Head Start eligibility does not perfectly predict participation, equation (10) represents the labor force outcomes on the population of people Head Start is intended to treat, or the average treatment effect among both compliers, eligible mothers that enroll their children in Head Start, and noncompliers, eligible mothers that do not. We can also use eligibility to instrument for Head Start participation in a fuzzy regression discontinuity by estimating:

$$L_{i,t} = \alpha + \beta \widehat{HS}_{i,4} + f(X_{i,3}, Z_{i,3}) + \epsilon_i \tag{11}$$

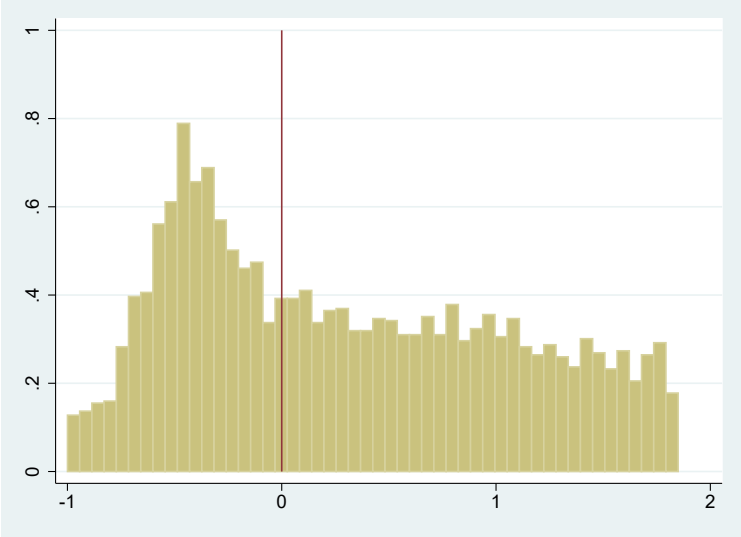
where $\widehat{HS}_{i,4}$ are the predicted values estimated in equation (8). Equation (11) scales up the estimates from equation (10) to measure the effect of Head Start on the average mother that enrolls a child in Head Start. However, problems can arise if the relationship between eligibility and Head Start participation is sufficiently weak.

Before proceeding to employ this strategy to evaluate the labor market effects of Head Start, I must establish the validity of the experimental design by addressing two concerns: bunching at the discontinuity and confounding effects of other programs. A regression discontinuity design is rendered invalid if agents have the ability to manipulate their positioning around the point of discontinuity (Lee and Lemieux 2010). In the case of Head Start, eligibility is a complex function of family size, income, year, and state poverty thresholds, therefore the exact threshold point where agents are incentivized to bunch is unlikely to be salient enough to make income manipulation around the discontinuity a possibility.[21] Figure 9, which plots the distribution of distances to the eligibility cutoff, confirms this. If there were evidence of bunching, there would be an over-representation of observations just to the left of the cutoff. There is no such bunching evident, which supports the validity of using a regression discontinuity empirical strategy. The right skew

[21] Recall that in Section 3, we discussed that it may be optimal for a mother to work $\overline{h}$ hours, but that it is unlikely that a mother would be able to effectively position herself at that point.

of this distribution is expected given the income distribution in the NLSY79 sample, which over-samples low-income respondents.

Figure 9: No Evidence of Bunching at the Discontinuity



A second concern when using an RD design is that the point of discontinuity confounds the effects of Head Start with other social programs. Although other forms of public assistance are also means-tested, the income measure and eligibility cutoffs are unique to determining Head Start eligibility and no other means-tested programs. However, Head Start eligibility is determined not only by the income criterion; families are also immediately eligible if they receive public assistance in the form of AFDC or SSI. However 98% of people in the NLSY79 sample were determined to be eligible based upon their placement below the income eligibility threshold, not because of their receipt (or eligibility to receive) public assistance. People receiving Head Start services did also receive public assistance, as programs to help lift people out of poverty often work in tandem. But because most people qualified as a result of their poverty levels using measures specific to Head Start, analysis at the margin of the discontinuity should reflect only the results attributable to the income cutoff. To be confident in the lack of confounding effects, I control for receipt of public assistance. I also provide additional identification checks (Section 7.5) that support attributing changes in labor force movements to Head Start.

6 Data

The regression discontinuity specifications are estimated using data from the geocoded version of the National Longitudinal Survey of Youth of 1979 (NLSY79) and the Children of the National Longitudinal Survey of Youth of 1979 (CNLSY79). The NLSY79 is a panel survey of 12,686 men and women, administered by

the U.S. Bureau of Labor Statistics, who were between the ages of 14 and 22 at inception of the survey in 1979. The respondents were then resurveyed in the years following (annually from 1979 until 1994 and biannually from 1994 until the present). The NLSY79 was launched with three sub-samples: a nationally representative sample (6,111 respondents), a Hispanic, Latino, black, or non-black, non-Hispanic, economically disadvantaged sample (5,295 respondents),[22] and a sample designed to represent the population of the U.S. military (1,280 respondents).[23] The NLSY79 surveys are administered throughout the year and ask questions about the prior calendar year. The survey contains a rich set of data regarding the labor market behavior, education, family life, and government program participation for each of the respondents.

The CNSLY79 is a derivative survey of the NLSY79 containing information about the children of the female respondents in the NSLY79 survey. The CNSLY79 survey began in 1986, with 5,255 children. As of 2012, the CNLSY79 respondent base has grown to 11,512 children born to NLSY79 mothers. Among many others, the CNLSY79 contains questions regarding children's preschool enrollment status and type, including their participation in Head Start.

The NSLY79 and CNLSY79 data was merged with a matrix, issued by the Department of Health and Human Services, that determine the income eligibility threshold as a function of family income, family size, state of residence, and year of eligibility determination. Similar data on the eligibility requirements for AFDC were also merged in order to determine public assistance eligibility.

I drop 2,304 children (from 814 mothers) with missing information about Head Start participation in all of the years of survey administration. I further drop 1,970 children (from 432 mothers) for whom I do not observe the necessary family size and income variables to determine eligibility at age 4, 1,415 children (from 472 mothers) without observations for those variables at age 3, and 220 children (from 60 mothers) whose eligibility I cannot determine at age 2. This brings the final sample size to 5,605 children of 3,154 mothers. Under the assumption of missing completely at random, dropping observations with missing values should not bias the results. Table A1 in the Appendix displays summary statistics with and without the dropped variables, and does not display any large discrepancies between the two populations.

Table 1 presents summary statistics of the mothers in the sample, including individual characteristics, labor market behavior, and children's Head Start eligibility status. The time-varying variables are calculated for mothers at the time that their first child is age three (the year that eligibility is determined). In the overall sample, 66% of women are employed, the average mother in the sample works 23 hours per week, and a substantial number of women are spending significant time out of the labor force, which also explains the very low average earned income of women. These labor supply rates are slightly higher but

[22]The economically disadvantaged non-black, non-Hispanic sub-sample suffered from significant attrition, and following the 1990 survey, none of the 1,643 members of the sample were eligible to interview. This may explain why the percentage of non-black, non-Hispanic children eligible for Head Start is low.

[23]After the 1984 interview, the U.S. military sample was dropped, with 201 randomly selected respondents kept as a representative sample.

Table 1: Summary Statistics: Mothers in Sample

	Sample		
Covariate	Obs	Mean	Std. Dev.
Hispanic	5605	0.202	0.401
Black	5605	0.299	0.458
White	5605	0.499	0.500
Married	5605	0.637	0.481
Highest Grade Completed	5604	11.97	2.12
Completed High School	5605	0.759	0.438
Hours Worked per Week	5532	23.13	19.26
Hours Worked per Week of Spouse	3636	43.73	12.51
Earned Income of Mother	5601	6268	9317
Employed	5605	0.663	0.473
Weeks Unemployed	5605	2.95	7.96
Weeks out of Labor Force	5464	23.74	22.37
Family Size	5605	4.33	1.55
Number of Children	5534	2.14	1.31
AFDC Eligible	5591	0.223	0.417
Age at Birth	5605	23.18	3.80
Child Eligible for Head Start at age 3	5166	0.341	0.474
Child Eligible for Head Start at age 4	5290	0.337	0.473

Variables that could vary over time (all variables except race, gender, and age at birth) are calculated during the calendar year in which a mother's first child turns three (the year during which eligibility at age four is determined).

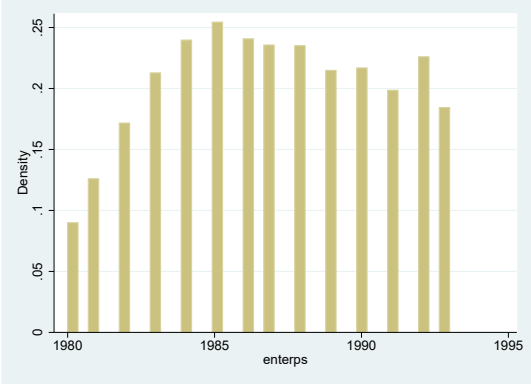
roughly consistent with an approximately 55% female labor force participation rate in the 80s and early 90s (Figure 1).

Figures 10 and 11 go into more detail painting a picture of the childcare arrangements of the mothers in the sample. Figure 10 illustrates the time frame of analysis: the mothers in this sample were enrolling their children in Head Start throughout the 80s and early 90s. Enrollment throughout the period was roughly uniform, with slightly fewer children of preschool age in the early 1980s.

Figure 11 goes a step further, showing the mode of childcare. Although roughly one third of all families in the sample are eligible for Head Start, consistent with insufficient funding and family choices to not enroll, Figure 11(a) displays that among all families, Head Start participation rates stayed constant at roughly 15-20% of the sample. Approximately 30-40% of those eligible participated in Head Start (Figure 11(b)). The uniformity over time in Figures 11(a) and 11(b) is somewhat surprising given that the funded enrollment of Head Start nearly doubled from 400,000 to just under 800,000 children during the same time frame.[24] Formal preschool is the dominant alternative to Head Start; only 10-20% of all children in the sample

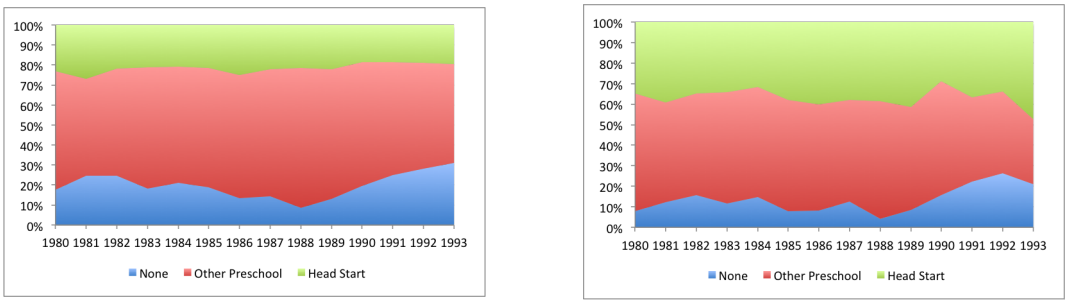
[24]One explanation for why Head Start enrollment in the sample remained constant in the face of a large expansion in enrollment at the national level may be that mothers in the early 80s were more likely to be young mothers (as mothers having

Figure 10: Children of Preschool Age



Histogram of the number of children age three in the the sample over time.

Figure 11: Preschool Enrollment by Type



(a) Whole Sample

(b) Eligible Families

are cared for informally. Although Figure 11 provides an annual cross-section of preschool enrollment, it does not give any indication of counterfactual preschool arrangement in the absence of Head Start eligibility.

The sample used to estimate the RD specification is restricted to a window around the discontinuity. Following Carneiro and Ginja, I restrict the window of observations to 15% to 185% of the eligibility cutoff specific to the state, year and family size of the observation. Table 2 presents summary statistics of the groups on either side of the cutoff within the window of estimation. Due to the large window, the eligible and ineligible populations differ on a number of covariates. Black mothers are more highly represented in the eligible group than the ineligible group. Eligible women are also more likely to receive public assistance. The labor market indicators of mothers upon eligibility determination also differ substantially. Eligible mothers work fewer hours, are less likely to be employed, and are more likely to be out of the labor force and earn a lower wage. Spouses work roughly the same number of hours, although eligible women are less likely to be married. In an ideal RD design, the window of observations would be narrow enough such that there would be no differences between those who were assigned treatment and those who were not. However, sample

children in 1980, for example, were only ages 15-23, whereas mothers having children in 1990 were ages 25-33) than those in the early 90s. Because young motherhood is associated with lower levels of socioeconomic status, women earlier in the study may have been more likely to enroll children in Head Start than those studied later in the time frame.

size limitations in the data require a wider window. To address these differences between groups, I included $Z_{i,t-1}$ in the control function of equation (8) so that $ELIG_{i,t}$ measures the effect of Head Start and not the pre-existing differences in labor market supply between groups. In particular, the difference in the number of families that receive public assistance could confound the effects of Head Start with other public assistance programs; I also control for receipt of public assistance in the year that eligibility was determined.

Table 2: Summary Statistics: Eligible and Ineligible

	Eligible for Head Start	Ineligible for Head Start
Covariate	Mean	Mean
Hispanic	0.210	0.222
Black	0.480	0.230
White	0.320	0.547
Married	0.320	0.759
Highest Grade Completed	11.00	12.08
Completed High School	0.583	0.823
Hours Worked per Week	15.87	26.87
Hours Worked per Week of Spouse	40.05	43.79
Weekly Wage	1604	6509
Employed	0.474	0.757
Weeks Unemployed	4.97	2.24
Weeks out of Labor Force	33.70	19.52
Family Size	4.70	4.13
Number of Children	2.67	1.99
AFDC Eligible	0.472	0.086
Age at Birth	21.75	23.39

Means are calculated using children as the unit of estimation. Variables that could vary over time (all variables except race, gender, and age at birth) are calculated during the calendar year in which a child turns three (the year during which eligibility at age four is determined). These values represent the differences between the groups when eligibility is determined for Head Start enrollment at child age four. Eligibility refers to the group of children whose family income is 15% to 100% of their specific income eligibility cutoff. Ineligibility refers to the group of children whose family income is 100% to 185% of their specific income eligibility cutoff.

The low labor market supply of women who are eligible for Head Start is intuitive. Because eligibility is determined by income, which is ultimately determined by wages, hours in the labor force, and the earnings of a partner or spouse, families whose household earnings are low enough to be below the threshold are likely to have low labor market supply ($h^* < \overline{h}$). The relatively low labor supply of eligible mothers also has implications for a means-tested preschool program like Head Start, as mothers must have a sufficiently low initial labor market supply ($h^* < H$) for Head Start to induce an increase in labor supply. Thus the low number of hours worked and low employment rate of eligible women further motivates a hypothesis that Head Start could increase mothers' labor market supply. The evidence to support this hypothesis is explored

in the following section.

7 Results

7.1 First Stage

I first test the validity of eligibility requirements as an instrument on the NLSY79 sample. The results of this first stage equation (equation (8)), estimated using a probit model, are presented in Table 3. These results show that eligibility at age four is a significant positive predictor of Head Start enrollment, associated with a 6% greater probability of Head Start participation in the overall sample and a 16% greater probability of program participation for black children. 6% and 16% increases in probability represent the average marginal effect of Head Start eligibility on participation for the whole sample and black mothers respectively, meaning that on average, an eligible family is 6% more likely to enroll a child in Head Start than an ineligible family.[25] These increases are from control means, the Head Start participation rates among "barley ineligible"[26] children, of 24% and 33% for the overall and black samples, respectively.[27] The control means of participation at age four are higher than the number of children that enroll at age three: 21% versus 12% for the overall sample and 31% versus 19% among black families. This may explain why age four eligibility better predicts participation than age three eligibility. The results are robust to changes in functional form, inclusion of additional covariates, and changes in the window of estimation. Estimates using a linear probability model are displayed in the Appendix in Table A2.

The instrument of eligibility requirements is not valid for eligibility at age three because there is not a strong enough relationship between eligibility requirements and participation. Carneiro and Ginja encounter similar results when using this instrument in their study. Thus, I restrict my analysis to children with eligibility determined at age four. The weakness of the instrument at age three may be due to a number of factors. Head Start does not have the funding available to allow all eligible students to enroll. Parents may be skeptical of Head Start or lack knowledge of the program. Finally, returning to the fact that we are using income from the previous calendar year as the eligibility-determining income (as outlined in the

[25] More formally, the average marginal effect is defined as

$$\frac{1}{N}\sum_{i=1}^{N}\{\Pr(HS_i = 1|ELIG_i = 1, Z_i, X_i) - \Pr(HS_i = 0|ELIG_i = 1, Z_i, X_i)\}$$

where N is the number of children in the regression sample. This is different from calculating the marginal effect for the average black child, which would be defined by calculating the probability of Head Start enrollment at the mean values of eligibility.

[26] Specifically, the control mean is the average Head Start participation rate for the population with incomes that place them between 100% and 125% of the eligibility threshold.

[27] The control means I chose to report in Table 3 are not control means specific to the time of Head Start enrollment but rather the control mean of all families with a child enrolled in Head Start. Thus one can think of Head Start eligibility at age four increasing enrollment upon an existing baseline of overall Head Start use. Eligibility at age four increased Head Start enrollment from the control by 23% for the overall sample and 47% for the black sample. The results of Table 3 can be thought of as "quantifying" the transition between Tables 11(a) and 11(b) at the margin of eligibility.

Table 3: First Stage Estimates

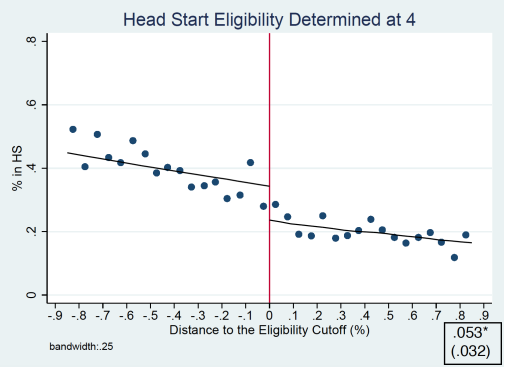
Head Start Participation		
	All Sample	Black
Eligible at Age 4	0.209**	0.462***
	(0.109)	(0.180)
Marginal Effect	*0.0554*	*0.155*
Observations	2759	1093
Eligible at Age 3	0.132	-0.0708
	(0.104)	(0.171)
Marginal Effect	*0.034*	*-0.025*
Observations	2702	1074
Control Mean	0.235	0.327
Standard Deviation	0.432	0.469

Estimated using a probit model, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, the gender of the child, and state and year dummies. All of these controls are time-stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, all time stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). The marginal effect refers to the average marginal effect of Head Start eligibility on the probability of enrollment. The control mean refers to the percentage of children that enrolled in Head Start just above the eligibility cutoff. Specifically, it is the mean of children whose family incomes were between 100% and 125% above of their state, family-size specific income cutoff. Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

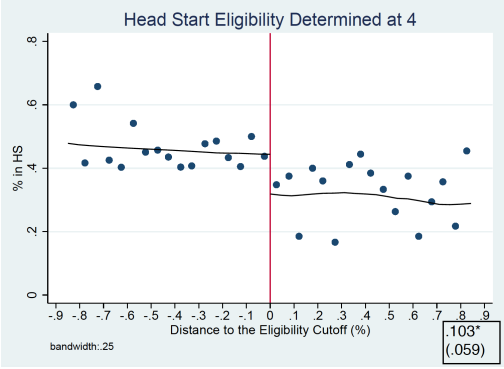
Head Start enrollment guidelines), because families have the opportunity to use an alternative, more recent, income metric that they believe better captures their financial situation, previous calendar year income may not always predict eligibility. It follows that it will also imperfectly predict enrollment. These eligibility requirements capture only a subset of possibly eligible families, making the eligibility indicator at age three too weak an instrument to predict enrollment.

Figure 12: Head Start Enrollment

(a) All Sample



(b) Black Families



Local linear regression using an Gaussian kernel with a bandwidth of 0.25. The individual points are binned averages, where the bins are determined by each child's distance from his/her individual eligibility cutoff, and each point represents the coordinate determined by the average distance to the eligibility cutoff and the percentage of children that participated in Head Start within the bin (with each bin representing a 0.05 point distance from the eligibility cutoff). The numbers reported in the box are an estimated magnitude of the discontinuity with bootstrapped standard errors using a bin size of double the optimal bandwidth, with optimality determined by an algorithm to minimize MSE, as fuzzy discontinuities often use double the optimal bandwidth (Imbens and Kalyanaraman, 2009). The bandwidth in the illustration and the bandwidth used to estimate the jump are therefore slightly different. The use of multiple bandwidths to both illustrate a jump and display a significant approximation of the size of the jump show robustness of the non-parametric specification to the size of the bandwidth. Bootstrapped standard errors (bootstrapped with 50 iterations) are displayed in parenthesis below the estimated magnitude of the effect. * significant at 10%; ** significant at 5%; *** significant at 1%

It is noteworthy that eligibility is a much stronger predictor of program participation for black families than in the overall sample.[28] Black families make up approximately half of all eligible families in the NLSY79 sample, so it is unsurprising that eligibility is a particularly strong predictor for that population.[29] Differentially stronger effects[30] for black mothers were also noted in Sabol and Chase-Lansdale's (2014) analysis of parental outcomes from the Head Start Impact Study. The differences in take-up rates could perhaps be due to a greater propensity for black women ages 20-60 to be a part of the labor force than white women prior to the 1980s (Juhn and Potter 2006). Head Start could therefore be a more attractive service for black mothers than for other populations with more women who stay at home. Due to black families' over-representation in Head Start eligibility, as well as the strength of the instrument, black mothers will prove to be a central component of the subsequent analysis. In their introduction of this instrument, Carneiro

[28]This is true for Head Start participation levels at age three and four.

[29]Black families do not seem to be overrepresented in Head Start participation on a national level, however. In 2015, black children comprised 29% of Head Start participants, while 38% identified as Hispanic/Latino and 43% as white (HHS 2015). In the Head Start Impact Study, only 17% of entering four-year-old children were black (HHS 2010).

[30]Though not in labor markets.

and Ginja also find differential strengths when estimating on different populations.[31]

Figure 12 graphically presents non-parametric illustrations of the impact of Head Start eligibility at age four on participation for both the entire sample and black families. The discontinuous decrease in the local linear regression at the point of the eligibility cutoff supports the results of the parametric regression discontinuity. In the non-parametric specification, the discontinuous jump is also larger in magnitude for black families, echoing the results of the parametric estimation. Non-parametric illustration of the discontinuity increases confidence in the instrument as the local linear regression presents only a flexible relationship between eligibility requirements and enrollment, a more agnostic approach that removes the potential of mispecification of the control function. The estimated magnitude of the jump in the non-parametric case, accompanied by bootstrapped standard errors, are displayed in the bottom right-hand corners of Figures 12(a) and 12(b). In the non-parametric case, the increase in the probability of participation upon eligibility is significant in both samples.

7.2 Alternative Arrangements

When studying the effect of Head Start on the labor market outcomes of mothers, understanding the counterfactual childcare option, the childcare type that would have been utilized in the absence of Head Start, is important. In Figure 11, a larger percentage of eligible families enroll their children in Head Start than ineligible children, but whether they are substituting from formal or informal care remains unanswered. Whether mothers are substituting from an alternative paid care or informal care has great implications for the labor market effects of Head Start. If families are substituting from alternative paid care, then Head Start acts as a lump-sum transfer for mothers who work and pay for care for hours $h^* > H$, resulting in only an income effect that reduces work incentives. Substitution from informal care, however, would make movements into the labor market more plausible. Table 4 presents estimates of how Head Start eligibility at age four impacts the use of two childcare alternatives: formal preschool and informal care. These estimates are derived from equation (8) with the outcome variable defined as a type of alternative care.

Head Start eligibility at age four is associated with negative, but statistically insignificantly different from zero, effects on enrollment in alternative preschools. A high-quality, free childcare service such as Head Start should induce some substitution from all alternatives. However, the only statistically significant change in alternative childcare arrangements as a result of Head Start eligibility is a 9% reduction in the probability of black children's use of informal childcare. This reduction in the use of informal childcare further points to a hypothesis that the labor market effects of Head Start may be particularly strong among black women as they substitute from informal childcare into formal employment.

[31] They find that the estimate is significant only for sons and not for daughters. They proceed by hinging their analysis on the long-term effects of Head Start exclusively on boys.

Table 4: Alternative Arrangements

	Alternative Preschool			
	Alternative Preschool		**Informal Care**	
	All Sample	Black	All Sample	Black
Head Start Eligible	-0.078	-0.126	-0.149	-0.521**
	(0.102)	(0.176)	(0.124)	(0.215)
Marginal Effect	*-0.0279*	*-0.0447*	*-0.0382*	*-0.089*
Control Mean	0.564	0.501	0.200	0.172
Standard Deviation	0.496	0.500	0.400	0.377
Observations	2761	1087	2749	1060

Estimated using a probit model, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The marginal effect refers to the average marginal effect of eligibility on the probability of use of the specific form of alternative care. The control mean refers to the percentage of children that used either alternative preschools or informal care just above the eligibility cutoff. Specifically, it is the mean of children whose family incomes were between 100% and 125% above of their state, family-size specific income cutoff. Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

7.3 Reduced Form Estimates

7.3.1 Labor Supply

With a valid instrument and evidence of substitution out of informal childcare, I proceed to explore the labor market effects of Head Start. I first present reduced form effects (estimates of the coefficients of equation (10)). Because of significant noncompliance, these reduced form estimates represent the intent to treat (ITT) estimate: the effect of Head Start programming on the subpopulation of eligible families. These estimates are "diluted" by noncompliers, eligible families who choose not to enroll their children or ineligible families who enroll their children in Head Start [32] To quantify the impact of Head Start participation on an average mother's labor supply, these estimates must be scaled up by the predicted probability that a child participates in Head Start, which I turn to in Section 7.4.[33]

[32]This is possible due to the provision that 10% of spots in Head Start centers are reserved for families who do not meet eligibility requirements.

[33]This is analogous to the interpretation of an instrumental variable strategy or a randomized study with non-compliers. The reduced form estimate gives a result for the entire affected population, while the local average treatment effect, or treatment on the treated, quantifies the impact on compliers only (Lee and Lemieux 2010).

One of the primary limitations of Sabol and Chase-Lansdale's (2014) analysis of Head Start is that they could only study the answer to one time-use question to make inferences about the employment effects of Head Start. Using the wide array of labor market questions in the NLSY79, labor supply can be looked at in a variety of dimensions, which in concert can provide a nuanced view of Head Start's impact. Table 5 presents estimates of the average impact of Head Start on the number of hours eligible mothers work per week. This measures a combined effect of both mothers entering the labor force, increasing the number of hours worked per week from a baseline of zero, as well as any increases for mothers who were already working at the time of eligibility determination.

As Table 5 demonstrates, Head Start eligibility significantly increases mothers' labor supply from baseline in both the overall sample and the sub-sample of black mothers. These estimates represents the average effect across all eligible families, including many who do not participate in Head Start, and present the first evidence to support the hypothesis that Head Start does indeed increase labor supply. Baseline labor market variables at eligibility determination are included in the control function. Therefore, these estimates can be interpreted as increases from baseline attributable to Head Start. The baseline means and medians are reported in Figure 5.[34] The increased hours in the labor market occur contemporaneously with Head Start enrollment at age four and persist to age five and age six. With a baseline labor supply among ineligible mothers of approximately 27 hours per week, by the time a child is age five, eligibility for Head Start reduces the initial 12 hour labor supply gap between the groups to 9.5 hours for the whole sample and to 5.5 for black mothers. The coefficients for black populations are larger in magnitude and more significant, consistent with both the increased strength of the instrument on black populations, as well as the greater propensity of black families to sort out of informal childcare.

The bottom panel of Table 5 presents estimates of Head Start's effect on the labor supply of the spouses or partners in eligible families. There is no evidence that Head Start causes any changes in the labor supply of spouses, supporting the evidence that mothers, as the traditional parent responsible for childcare, are increasing their labor supply as they substitute responsibilities from informal work at home to formal employment. This is consistent with evidence that husbands tend to have very weak wage and income effects (Hall 1973), while women are much more sensitive to changes in work-life balance policy.

I proceed to assess whether mothers are increasing the number of hours worked from a baseline of already working, for example, adding work hours to a schedule, adding a second job, or finding a new job with more available hours of employment, or are being induced to enter the labor force. Table 6 estimates the reduced form effect of Head Start eligibility on the probability of mothers' employment. The results of Table 6 indicate that Head Start causes a significant increase in the percentage of mothers working in

[34]Recall that due to data constraints, the window of estimation is large enough such that there are differences in labor supply between the two groups. The reason why a baseline mean rather than a control mean is reported is that it provides more information about how Head Start causes the eligible population to increase their labor supply above initial levels, which are lower than the eligible group.

Table 5: Weekly Hours Worked: Reduced Form

	Hours Worked per Week		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	2.35*	2.20*	2.08
	(1.21)	(1.32)	(1.61)
Observations	2591	2593	2376
Baseline Mean	15.09		
Baseline Std. Dev.	18.8		
Baseline Median	0		
Black			
Head Start Eligible	3.89**	6.28***	5.76***
	(1.91)	(2.01)	(2.43)
Observations	1028	1009	941
Baseline Mean	15.56		
Baseline Std. Dev.	18.7		
Baseline Median	0		
Spouses			
Head Start Eligible	0.99	0.276	0.276
	(1.51)	(1.44)	(1.59)
Observations	1174	1182	1138
Baseline Mean	40.05		
Baseline Std. Dev.	14.91		
Baseline Median	40.0		

Estimated using ordinary least squares, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The baseline mean and median refer to the average hours worked among eligible families when eligibility was determined (the hours worked when the child was age three). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

both the overall sample and among black women. This provides an additional layer to interpret the results of Table 5: since women are moving into the labor force, a significant component of the increase in hours supplied is driven by women increasing from a baseline of zero. These results are consistent with very low baseline labor supply indicators among eligible women. In both the overall sample and among black women, the median mother worked zero hours per week and was out of the labor force, providing evidence for the claim that for mothers with $h^* = 0$, Head Start should provide a nudge into labor force participation.

The average Head Start-eligible mother experiences a 6% and 7% increase in the probability of employment at age four for the overall sample and black sample, respectively. These estimates imply a 13% and 16%[35] increase from baseline employment levels[36] in the overall sample and black sample, respectively. The magnitude of these employment changes are similar to those found in Baker, Gruber, and Milligan (2008); they find that the employment rate of married women in Quebec, all of whom were eligible for $5 per day universal childcare, rose 14.5% from baseline relative to the rest of Canada.[37] However, because the Canadian example is not means-tested, there were unlikely to be large gaps in employment rates between Quebec and the remainder of Canada. In this study, Head Start closes the employment gap between the eligible and ineligible groups: at age four, Head Start decreases the employment gap between eligible and ineligible groups from 31%[38] to 25% in the whole sample and 24% among black women.[39]

The fixed cost model of employment suggests that women should be supplying labor at high levels; the estimates in Table 7 show that women are indeed moving predominantly into full-time work.[40] This is also consistent with the findings of Baker, Gruber, and Milligan (2008), who find that universal childcare induced women to differentially enter the labor force into full time rather than part time employment.

While the evidence seems to indicate that movement into the labor force is the primary driver of increases in the hours worked per week, women who are already in the labor force may be affected as well. As Figure 6 (Section 3) illustrates, Head Start has an ambiguous effect for existing labor market suppliers (mothers with $h^* > 0$). If the income effect dominates, then the estimates of the number of hours worked per week would be incorporating both movements into the labor force and decreases in labor supply for those who were already in the labor force. Alternatively, if the substitution effect dominates, the estimates

[35] These numbers are rough estimates calculated by dividing the marginal effects by baseline employment rates at age three.

[36] Note that the employment rate I refer is the employment rate for the population of mothers both in and out of the labor force. This is different than the employment rate often reported, which includes only those seeking work. Specifically, the employment rate discussed here is defined by:

$$\frac{1}{N}\sum_{i=1}^{N} EMP_i$$

where EMP is an indicator that a mother is working. In Table A3, when I study unemployment, I am referring to the standard metric where the denominator reflects only includes those that are actively seeking employment.

[37] This increase in employment is driven by increased childcare utilization of 14% percentage points, roughly comparable to the increase in utilization among black population.

[38] Ineligible mothers had a baseline employment rate of 76%.

[39] By age five, the gap between eligible and ineligible black women is reduced to 22%.

[40] This is interesting because it means that many women are induced to work $h > H$. This indicates that there exists a strong incentive to enter full-time work as women work beyond the theoretical upper bound of Head Start's nudge into the labor force.

Table 6: The Effect of Head Start on the Probability of a Mother Working: Reduced Form Estimates

	Mothers' Employment Dummy		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	.338**	0.123	0.073
	(.133)	(.127)	(.134)
Marginal Effect	*0.057*	*0.025*	*0.017*
Observations	2593	2551	2393
Baseline Mean	0.450		
Baseline Std. Dev.	0.497		
Baseline Median	0.00		
Black			
Head Start Eligible	.507*	.508**	0.344
	(.279)	(.227)	(.255)
Marginal Effect	*0.071*	*0.093*	*0.056*
Observations	1009	988	918
Baseline Mean	0.450		
Baseline Std. Dev.	0.498		
Baseline Median	0.00		

Estimated using a probit model, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The marginal effect refers to the average marginal effect of eligibility on the probability of employment. The baseline mean and median refer to the average number of mothers working among eligible families when eligibility was determined (at age three). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: The Effect of Head Start on the Probability of a Mother Working Part time versus Full Time: Reduced Form Estimates

	Part Time Work			Full Time Work		
	Child Age			Child Age		
	Age 4	Age 5	Age 6	Age 4	Age 5	Age 6
All Sample						
Head Start Eligible	0.338	0.017	-0.182	0.209*	0.119	0.223
	(0.133)	(0.122)	(0.125)	(0.122)	(0.107)	(0.126)
Marginal Effect	*0.057*	*0.004*	*-0.051*	*0.045*	*0.033*	*0.064*
Observations	2582	2557	2388	2591	2555	2393
Baseline Mean	0.193			0.267		
Baseline Std. Dev.	0.394			0.443		
Baseline Median	0.0			0.0		
Black						
Head Start Eligible	0.192	0.085	-0.128	0.221	0.343*	0.385*
	(-0.222)	(-0.245)	(-0.256)	(-0.218)	(-0.199)	(-0.229)
Marginal Effect	*0.04*	*0.02*	*-0.03*	*0.045*	*0.084*	*0.10*
Observations	996	996	922	1013	997	931
Baseline Mean	0.175			0.279		
Baseline Std. Dev.	0.380			0.449		
Baseline Median	0.0			0.0		

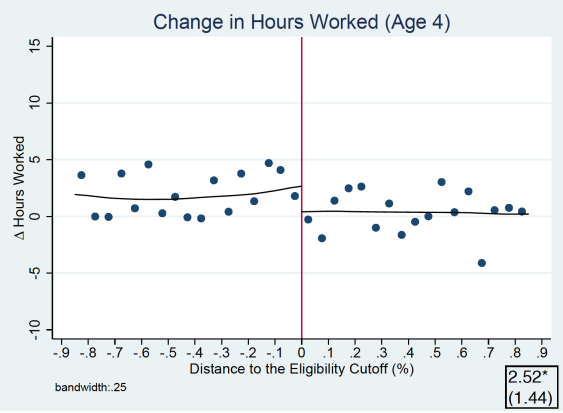
Estimated using a probit model, with children as the unit of observation. Full time work refers to employment while working more than 35 hours per week. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The marginal effect refers to the average marginal effect of eligibility on the probability of enrollment. The baseline mean refer to the average number of mothers working full and part time among eligible families when eligibility was determined (at age 3). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Weekly Hours Worked for Baseline Labor Suppliers: Reduced Form

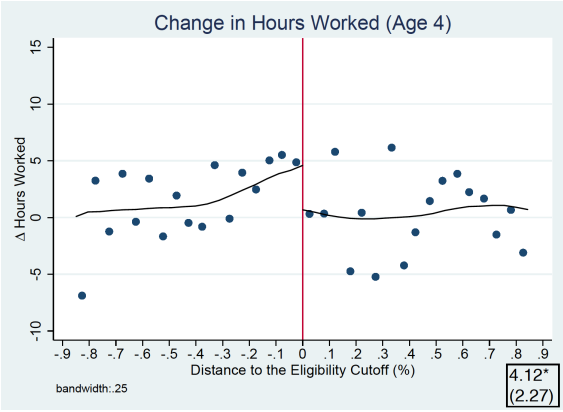
	Hours Worked per Week		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	1.10	1.86	2.16
	(1.35)	(1.80)	(1.90)
Observations	1436	1407	1307
Baseline Mean	33.97		
Baseline Std. Dev.	12.39		
Black			
Head Start Eligible	3.06*	1.63	3.30
	(1.84)	(2.70)	(2.99)
Observations	573	558	521
Baseline Mean	34.71		
Baseline Std. Dev.	10.63		

Estimated using ordinary least squares, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The baseline mean and median refer to the average hours worked among eligible families whose mothers were in the labor force when eligibility was determined (the hours worked when the child was age three). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

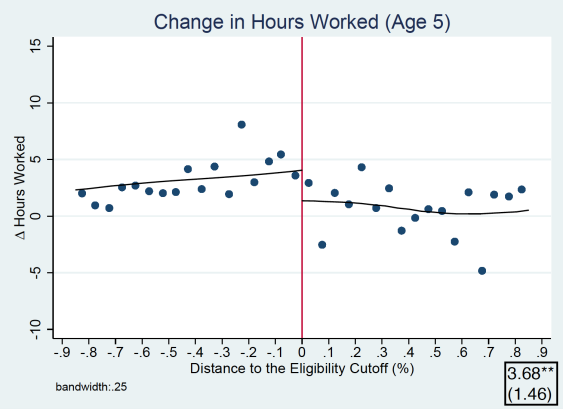
Figure 13: Hours Worked and Head Start Eligibility: Local Linear Regressions



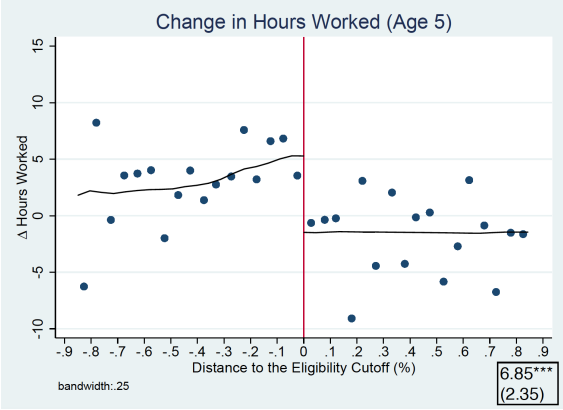
(a) All Sample, Age 4



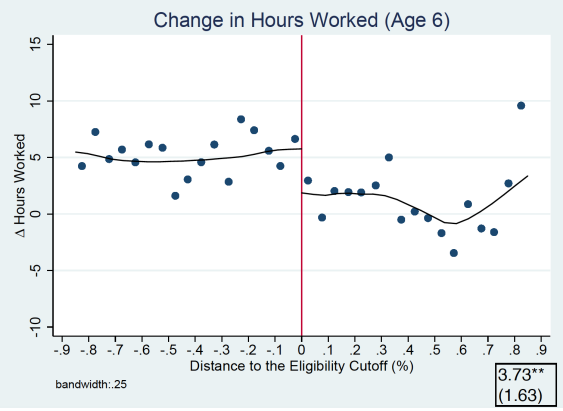
(b) Black Mothers, Age 4



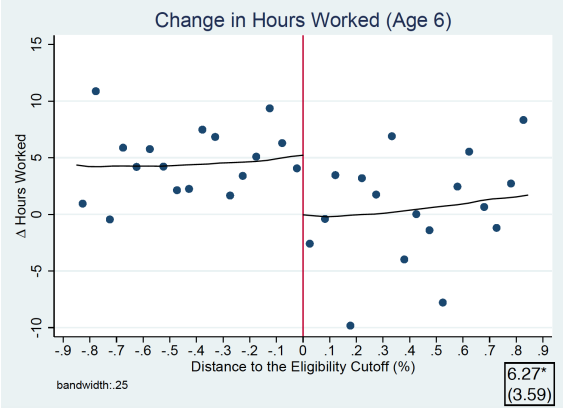
(c) All Sample, Age 5



(d) Black Mothers, Age 5



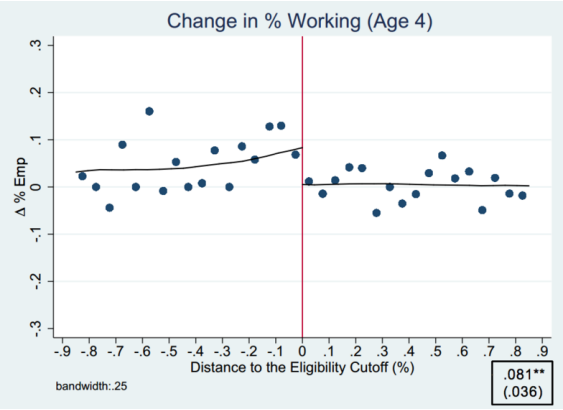
(e) All Sample, Age 6



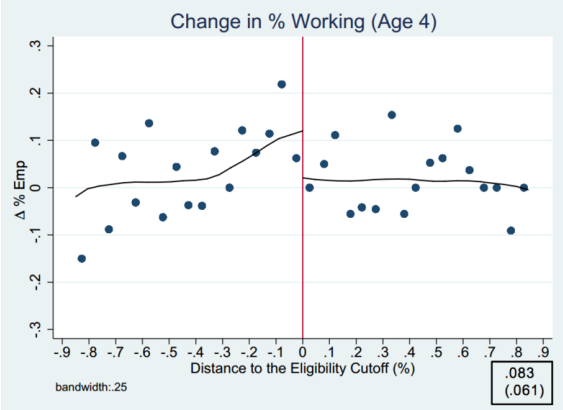
(f) Black Mothers, Age 6

Local linear regression using an Gaussian kernel with a bandwidth of 0.25. Negative numbers represent eligible individuals. The individual points are binned averages, where the bins are determined by each child's distance from his/her individual eligibility cutoff, and each point represents the coordinate determined by the average distance to the eligibility cutoff and the percentage of children that participated in Head Start within the bin (with each bin representing a 0.05 point distance from the eligibility cutoff). The numbers reported in the box are an estimated magnitude of the discontinuity with bootstrapped standard errors using a bin size of double the optimal bandwidth, with optimality determined by an algorithm to minimize MSE, as fuzzy discontinuities often use double the optimal bandwidth (Imbens and Kalyanaraman, 2009). The bandwidth in the illustration and the bandwidth used to estimate the jump are therefore slightly different. The use of multiple bandwidths to both illustrate a jump and display a significant approximation of the size of the jump show robustness of the non-parametric specification to the size of the bandwidth. Bootstrapped standard errors (bootstrapped with 50 iterations) are displayed in parenthesis below the estimated magnitude of the effect.. * significant at 10%; ** significant at 5%; *** significant at 1%

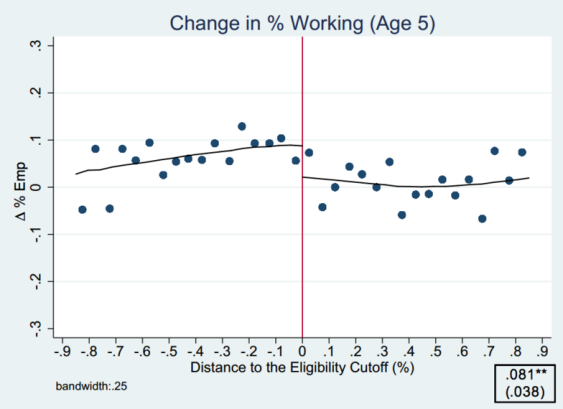
Figure 14: Employment and Head Start Eligibility: Local Linear Regressions



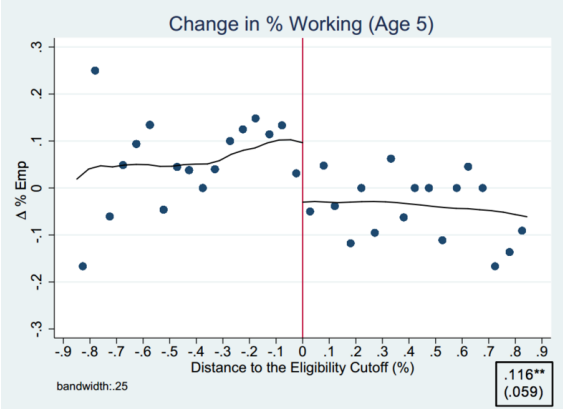
(a) All Sample, Age 4



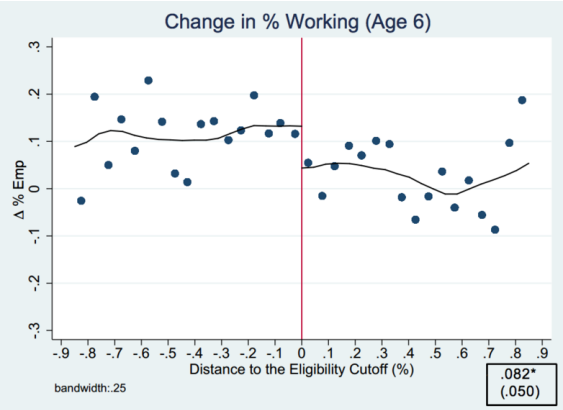
(b) Black Mothers, Age 4



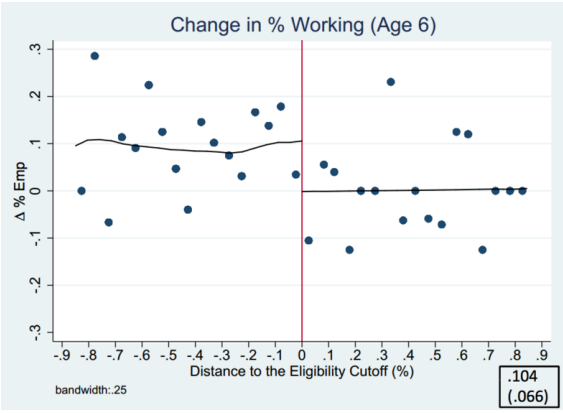
(c) All Sample, Age 5



(d) Black Mothers, Age 5



(e) All Sample, Age 6



(f) Black Mothers, Age 6

Local linear regression using an Gaussian kernel with a bandwidth of 0.25. Negative numbers represent eligible individuals. The individual points are binned averages, where the bins are determined by each child's distance from his/her individual eligibility cutoff, and each point represents the coordinate determined by the average distance to the eligibility cutoff and the percentage of children that participated in Head Start within the bin (with each bin representing a 0.05 point distance from the eligibility cutoff). The numbers reported in the box are an estimated magnitude of the discontinuity with bootstrapped standard errors using a bin size of double the optimal bandwidth, with optimality determined by an algorithm to minimize MSE, as fuzzy discontinuities often use double the optimal bandwidth (Imbens and Kalyanaraman, 2009). The bandwidth in the illustration and the bandwidth used to estimate the jump are therefore slightly different. The use of multiple bandwidths to both illustrate a jump and display a significant approximation of the size of the jump show robustness of the non-parametric specification to the size of the bandwidth. Bootstrapped standard errors (bootstrapped with 50 iterations) are displayed in parenthesis below the estimated magnitude of the effect. * significant at 10%; ** significant at 5%; *** significant at 1%

in Table 5 would be incorporating increases in the labor supply for existing suppliers. The estimates for the change in hours worked for mothers with a baseline labor supply of $h^* > 0$ are presented in Table 8. The labor market effects among initial suppliers, though positive, are small in magnitude and insignificant at almost all ages in both samples. This incorporates both the fact that women who increase the number of hours worked from an existing baseline will have a smaller change in labor supply than a woman who enters the workforce to work 40 hours per week, as well as Head Start's reduced impact influencing high-supplying mothers to change their labor market behavior. There is no significant difference in take-up between low and high suppliers: incorporating an interaction term between $ELIG_{i,t}$ and the hours worked at baseline in equation (8) yield an interaction coefficient extremely close to zero. Thus, these differences are not driven by heterogeneous take-up rates, but rather by stronger labor supply responses for women initially out of the labor force.

Figures 13 and 14 illustrate the main results presented in Tables 5 and 6, estimated in a more flexible manner. In these specifications, I use a local linear regression to estimate the change in hours worked per week (Figure 13) and the percentage of women working (Figure 14) from a baseline specified at eligibility determination. Just as non-parametric estimation of the effect of Head Start eligibility on enrollment provided estimates without danger of mispecification in Figure 9, Figures 13 and 14 do the same for labor market outcomes. Figures 13 and 14 echo the results of the parametric specification, but measure changes from a baseline in a slightly different way. Rather than including baseline labor force metrics in the control function, the local linear regression uses the deviation from the baseline as the outcome variable.[41] The results of Figures 13 and 14 are robust to variation in the bandwidth and kernel type.

The main results of Table 5 hold in Figure 13. There is a positive discontinuous effect on the change in hours worked per week from baseline after passing the eligibility threshold, implying that mothers that are eligible for Head Start are more likely to increase their labor supply than those that are not. The discontinuity in Figure 13 is larger for black families than in the overall sample, similar to the results in Table 5. These same trends are also evident in Figure 14, which illustrates the effects of Head Start on the change in the percentage of women working from baseline. Differentially larger positive discontinuities are also evident for black mothers in Figure 14. The magnitudes of the discontinuities and their standard errors, estimated via bootstrapping, are displayed in the bottom right hand corner of each graph and provide corroborating evidence of a discontinuity in labor force outcomes at the point of eligibility determination.

Alternative labor force metrics also support the hypothesis of the movement of mothers into the labor force as a result of Head Start. Total weeks out of the labor force as an alternative outcome variable provides additional suggestive evidence that women are moving into the labor force. However, there does not seem to be any changes in levels of unemployment,[42] supporting the hypothesis that women are substituting

[41] Specifically, the outcome variable is defined as $L_{i,t} - L_{i,3}$ for $t = 4, 5, 6$, with L defined as either hours worked per week or an indicator of employment.

[42] Defined by responses to the NLSY79 survey asking about the number of weeks a respondent was unemployed, as differen-

out of informal work or childcare, not recovering from instances of unemployment or negative labor market shocks. These results are presented in Appendix Table A3.

7.3.2 Wages and Income

Strong evidence that Head Start mothers are choosing to increase their labor supply motivates study of the subsequent effect on income growth. As mothers increase their labor supply, one would anticipate observed increases in mothers' earned income. Table 9 confirms this, providing estimates for the impact of Head Start eligibility on an indicator variable that a mother increases her earned income from baseline. By the time a child is age five, mothers are 8% and 15% more likely to have increased income in the overall sample and black sample, respectively. The probability of increased income at age five and six are similar in magnitude, but slightly above, the probability of increased employment, indicating that these increases in income are primarily driven by women entering the workforce.

Table 10 explicitly decomposes this growth in earned income into income growth driven by movement into the labor force and income growth from mothers that were already reporting earned income at baseline.[43] Table 10 illustrates that Head Start eligibility resulted in both income growth from zero (10(a)) and income growth among existing earners (10(b)) for black mothers. However, though positive, neither income growth from zero nor increased income showed statistically significant changes for the overall sample. The negative coefficients in both Tables 8 and 9 for earned income among black mothers with children aged four is surprising and somewhat perplexing. If mothers' labor supply is increasing, one would expect to see income increase. Perhaps some women leave jobs with fewer hours in an attempt to find a job with a higher level of employment, and this job separation causes an initial decrease in income.[44] However, these negative changes in earned income are statistically insignificant.

Table 11 explores whether the observed income growth is due exclusively to increases in labor supply, or whether the increase in labor supply is accompanied by rising wage rates. Increased labor supply may lead to rising hourly wage rates as increased experience creates highly valued employees, which in turn earn higher wages. Table 11 does not seem to indicate any statistically significant changes in hourly wages, either in a positive or negative direction. While a lack of wage growth coupled with an increase in labor supply seems counter-intuitive, this has been observed empirically in a number of other situations (Card and

tiated from being simply out of the labor force.

[43] Income was analyzed in this way because the distribution of income is log-normal with bunching at zero, making this the most intuitive way of studying changes. A simple linear model would be mispecified given that the distribution of the outcome variable is far from normally distributed, and a log model would fail to incorporate any of the initial zero values. A regression of the cubic root transformation of mothers' income yields similar positive results, incorporating both the results of 10(a) and 10(b).

[44] Table A3 in the Appendix provides suggestive evidence that this may indeed be the case. Among black mothers, age four is the only child age at which mothers experience an increase in the number of weeks of unemployment. The results in Table A3 are statistically insignificant, but so are the negative estimates of wage decreases among black mothers with children age four.

Table 9: Probability of Increasing Mothers' Earned Income: Reduced Form Effects

	Increase in Earned Income Dummy		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	0.128	0.22**	0.071
	(.108)	(.107)	(.121)
Marginal Effect	*0.044*	*0.076*	*0.026*
Observations	2601	2582	2434
Black			
Head Start Eligible	-0.085	.511***	0.280
	(.189)	(0.183)	(.195)
Marginal Effect	*-0.027*	*0.150*	*0.089*
Observations	1027	1015	936

Estimated using a probit model, with children as the unit of observation. The outcome variable is an indicator variable that takes the value 1 if a mother's income at a given child's age (4, 5, or 6) is higher than the baseline income determined at age 3. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The marginal effect refers to the average marginal effect of eligibility on a wage increase from baseline. Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: Mothers' Earned Income Decomposition: Reduced Form Effects

(a) Probability of Non-Zero Earned Income

	Non-Zero Income Dummy		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	0.169	0.176	0.129
	(.126)	(.116)	(.136)
Marginal Effect	*0.034*	*0.042*	*0.033*
Observations	2612	2585	2432
Black			
Head Start Eligible	-0.091	.493**	0.480**
	(.259)	(0.232)	(.226)
Marginal Effect	*-0.018*	*0.105*	*0.096*
Observations	1016	1015	932

(b) Log Wages

	Log Earned Income		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	-0.031	0.155	0.058
	(.118)	(.100)	(.117)
Observations	1390	1407	1379
Black			
Head Start Eligible	-0.085	.239*	0.313*
	(.184)	(0.145)	(.188)
Observations	548	547	536

Estimated using children as the unit of observation: panel (a) was estimated with a probit model, and panel (b) with ordinary least squares. Wages are measured in 2000 dollars. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The marginal effect (reported in (a)) refers to the average marginal effect of eligibility on a wage increase from zero. Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11: Hourly Wages of Mothers: Reduced Form Effects

	Log Hourly Wages		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	0.053	0.105	-0.142
	(.185)	(.199)	(.226)
Observations	1484	1558	1544
Black			
Head Start Eligible	0.037	-0.017	0.048
	(.085)	(.128)	(.072)
Observations	584	608	600

Estimated with ordinary least squares using children as the unit of observation. The hourly wage is the average hourly wage reported across all jobs, calculated as an average and not a weighted average of wage rates measured across all jobs because reporting error made it impossible to match the hours worked with the appropriate wage. The wages are in 2000 dollars. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year during which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12: Family Income: Reduced Form Effects

	Family Income		
	Child Age		
	Age 4	Age 5	Age 6
All Sample			
Head Start Eligible	541	-3167	-2000
	(3414)	(2496)	(2252)
Observations	2431	2415	2292
Black			
Head Start Eligible	-4396*	-2013	-1948
	(2547)	(2626)	(2697)
Observations	977	960	897

Estimated with ordinary least squares using children as the unit of observation. Family income is in 2000 dollars and is the same unit of measurement used to determine eligibility status. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year during which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Hyslop 2005 and Gladden and Taber 2009); increased time in the labor force does not necessarily translate into real wage growth. Mothers with a zero baseline hourly wage are excluded from this specification.

Even with increases in mothers' earned income, there is a surprising lack of growth in family income. Table 12 shows negative, though mostly insignificant, effects on total family income from baseline. Perhaps, although Head Start causes women to move into the labor force, it results in an income effect for the whole family. However, this hypothesis is still perplexing because the labor supply of spouses remains unchanged by Head Start participation (Table 5). These negative estimates could be a result of decreases in other forms of income, such as public assistance transfer programs. The women moving into the labor force as a result of Head Start are likely to be low wage earners: they have low levels of education and were perhaps out of the labor force for a significant period of time, making them less likely to earn a high wage. Because family income measures all income components, families may be receiving less non-wage income in the form of reduced public assistance, child support, or unemployment benefits as mothers move into the labor force. However, since mothers are supplying into low wage occupations, this decrease in other forms of income may dominate the effects of a mother's income increase. Although these estimates are negative, they are almost all insignificant. Untangling the income effects requires inquiry into the heterogeneous effects of Head Start, a concept I turn to in section 7.6.

These reduced form estimates provide evidence on the direction of mothers' movement into the labor force and insight to quantify the extent to which Head Start can close the employment gap between poorer eligible families and richer ineligible families. However, they do not give any indication of the magnitude of change for a given individual. These estimates represent an average treatment effect for all eligible families and are diluted by non-compliers. To quantify the effects, I use the eligibility requirements to instrument for Head Start participation.

7.4 Fuzzy RD Estimates

In this section I quantify the labor supply response to Head Start participation for a complying mother, rather than for the whole eligible population near the threshold. I focus exclusively on black mothers because the weakness of the instrument for the overall sample makes it difficult to scale up estimates.[45] Table 13 presents estimates of the treatment effect on compliers for three main variables of interest: hours worked per week (13(a)), the probability of employment (13(b)),[46] and the probability of a wage increase (13(c)).

[45] Carneiro and Ginja also experienced this problem when trying to scale up their estimates using this empirical strategy. In their original use of the instrument, Carneiro and Ginja remark that their estimates on child health and behavioral problems as a result of Head Start are implausibly large due to the weakness of the instrument, and a similar problem may persist in the following results. However there is an argument to be made that though these results are large, they are not necessarily implausibly large.

[46] Estimates in Table 13(b) and 13(c) use a probit model with both the predicted values and estimated residuals from equation (8) as predictors. Table A4 reports two stage least squares estimates for panel 13(b) as a comparison, but these estimates are too large in magnitude to be plausible. A bivariate probit, the ideal model in this situation, did not converge for values in panel

Table 13: Labor Supply Effects for Black Mothers: Fuzzy RD Estimates

(a) Hours Worked Per Week

	Hours Worked		
	Child Age		
	Age 4	Age 5	Age 6
Predicted Head Start Participation	27.21***	34.31***	30.85**
	(10.35)	(12.25)	(13.94)
Observations	1017	997	930
Baseline Mean	15.56		
Baseline Std. Dev.	18.7		
Baseline Median	0.00		

(b) Probability of Employment

	Mothers' Employment Dummy		
	Child Age		
	Age 4	Age 5	Age 6
Predicted Head Start Participation	3.86***	2.79**	2.17
	(1.45)	(1.19)	(1.39)
Marginal Effect	*0.451*	*0.461*	*0.335*
Observations	1003	985	930
Baseline Mean	0.450		
Baseline Std. Dev.	0.498		
Baseline Median	0.00		

(c) Probability of Earned Income Increase

	Increase in Earned Income Dummy		
	Child Age		
	Age 4	Age 5	Age 6
Predicted Head Start Participation	.374	2.33*	2.83**
	(1.46)	(1.30)	(1.31)
Marginal Effect	*0.071*	*0.466*	*0.474*
Observations	1010	1009	926

Estimated using children as the unit of observation: (a) estimated with ordinary least squares, (b) and (c) estimated with a probit model. In (b) and (c), the estimated residuals from the first stage estimate were included in the probit specification. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year during which the child was determined eligible (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Marginal effect refers to the average marginal effect of predicted Head Start enrollment at age four on (b) the probability of employment and (c) the probability of a wage increase. Robust standard errors (in parenthesis) are clustered at the state-year level. Baseline means in (a) and (b) refer to the average hours worked per week and average employment rate among eligible mothers when eligibility was determined. * significant at 10%; ** significant at 5%; *** significant at 1%.

The magnitude of the effect of Head Start on black mothers who take it up is large, but not implausible. Women whose children participate in Head Start increase their labor supply by 30 hours per week from baseline, have a 45% probability of entering the labor force, and a 45% probability of increasing their income from baseline. These large values are consistent with Cascio's (2009) study of free kindergarten expansion, which also finds a 45% increase in maternal employment for each youngest child that enrolls in kindergarten. Thus, studying Head Start provides supporting evidence that programs providing free childcare during school hours may have very large effects on maternal labor supply. A 1994 U.S. Government Accountability Office (GAO) simulation report predicted that a full childcare subsidy could increase the percentage of poor mothers working from 29% to 44%, and of near-poor mothers from 43% to 57% (GAO 1994). The estimates attributed to Head Start's impact as a childcare subsidy provide empirical support for this simulation.

Magnitudes of this size are plausible due to the low baseline labor supply of eligible mothers. Figure 15 shows the distribution of the labor supply of eligible women at baseline. The distribution in both the whole sample and in the sample of black mothers are similar: more than 50% of women in both samples do not work.[47] These women are indeed the low labor suppliers that were identified in Section 3 as those most likely to be induced to enter the labor force. As the reduced form effects illustrate, many women are moving into full-time work, consistent with a fixed costs model of labor supply. Thus the 30 hour per week increase accompanied by a 45% increased probability of employment is plausible as it is incorporating the effects of many women moving from zero hours to a full time job.[48] Because some initial suppliers also increased their labor supply, a 30 hour per week increase is incorporating both the effects of women moving into the labor force as well as some women increasing from an existing baseline, perhaps from part-time to full-time work. The labor supply movements flow into increased income, as black women who enroll their children in Head Start experience a greater than 45% probability of increasing their earned income by the time their child is five.

Although we cannot make claims about precise magnitudes on the overall population due to the weakness of the instrument, a back of the envelope scale-up calculation implies that in the overall sample, Head Start participation implies a 30-35 hour increase in labor supply. This is on the same order of magnitude as the black population, where greater take-up makes it easier to estimate the fuzzy RD.

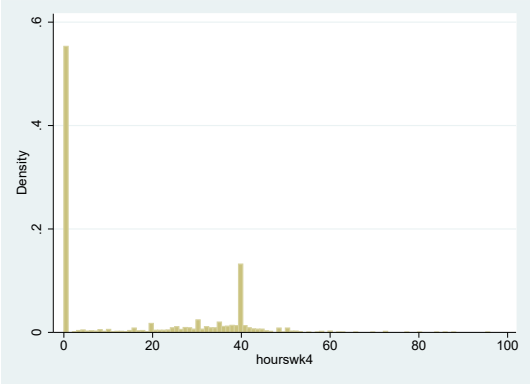
Another limitation to the weakness of the instrument as a result of significant noncompliance in Head Start is the fact that we cannot observe the labor market effects of mothers with eligibility determined at age three. Recall from the first stage estimates that eligibility requirements at age three failed to predict Head Start enrollment. However, there exists suggestive evidence (on educational investment) that the

(b) except for at age five, with a slightly lower estimate of 34%, another comparison.

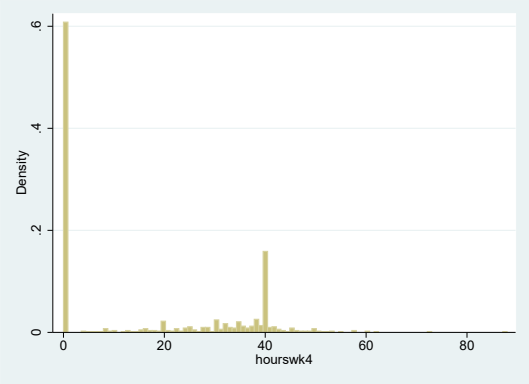
[47] This is also reflected by the zero baseline medians in Table 13.

[48] It is also relevant to note that the hours worked per week variable is susceptible to rounding error, where respondents may tend to round up or down their average hours worked.

Figure 15: Hours Worked at Time of Eligibility Determination: Eligible Mothers



(a) Whole Sample



(b) Black Mothers

impacts of Head Start are stronger for parents of three-year-old enrollees than four-year-old enrollees (Sabol and Chase-Lansdale 2014),[49] as they have the opportunity to experience more time in the program. If enrollment at age four has a lesser impact on families than enrollment at age three, one can make a very weak claim that the effects of enrolling at age four are perhaps generalizable to those at age three. However, this is simply an assertion that merits further study.

7.5 Identification Checks

Having established evidence of a positive effect on mothers' labor supply, I provide additional evidence to confidently attribute this observed increase to a causal impact of Head Start. The RD results are robust to different functional forms, inclusion of additional covariates, and changes in the size of the window of estimation. The local polynomial regressions are also robust to changes in the bandwidth.

I also conduct a series of "placebo tests" to provide additional evidence that the results presented previously are indeed the result of Head Start. This additional evidence is necessary due to the large window in the RD design. Because of the large window, there may be unobservable differences between groups, which would be correlated with $ELIG_{i,t}$, and may also be correlated with labor force outcomes. Consider two cases:

1. A case of short-term shifts: if more eligible mothers have suffered a labor market shock (that make them eligible) than ineligible mothers, $ELIG_{i,t}$ may be detecting simply a labor force recovery from a negative labor supply shock at baseline. This would result in an increase in labor supply 1-3 years

[49] In their analysis of the mothers of participants in the HSIS, Sabol and Chase-Landsale find that African American mothers whose children were assigned to Head Start were significantly more likely to invest in their education than those in the control group. However, this effect was only evident for the cohort of families who enrolled their children at age three, and not for families who enrolled their children at age four.

after eligibility is determined, regardless of Head Start participation.[50]

2. A case of child-age-dependent shifts: mothers choose to go back to work when their children are ages four to six because they require less attention than a toddler. Moreover, poor women (eligible mothers) are more likely to go to work because they are suffering from stronger pressures to earn money to provide for children and so differentially increase their hours worked with respect to less resource-constrained, ineligible mothers. This would result in an increase in labor supply at ages 4-6, regardless of Head Start participation.

Both of these scenerios are plausible, and would result in positive labor supply shifts for eligible mothers, which would incorrectly attribute these movements to Head Start when in reality, the program has no effect. To address this concern, I introduce a "placebo test." Consider estimating equation (10) at an age when a child could not possible be eligible for Head Start, regardless of their placement on either side of the income-eligibility discontinuity. At $t = 1, 2$ (child ages one and two) children are ineligible for Head Start as a result of their age, independent of their placement on either side of the eligibility cutoff.[51] We can estimate equation (10) in two ways. The first corresponds to case one:

$$L_{i,t+1} = \alpha + \beta ELIG_{i,t} + f(X_{i,t-1}, Z_{i,t-1}) + \epsilon_i \tag{12}$$

In equation (12), we are testing whether Head Start eligibility, determined at an age when a child could not possibly attend Head Start, causes some short to medium term labor market effects. The equation corresponding to case two is as follows:

$$L_{i,5} = \alpha + \beta ELIG_{i,t} + f(X_{i,t-1}, Z_{i,t-1}) + \epsilon_i \tag{13}$$

In equation (13), we are testing whether Head Start eligibility, determined at an age when a child could not possibly attend Head Start, causes an increase in labor supply during the time-frame we are studying (ages 4-6). If the coefficient β were positive and significant in either equation (12) or (13), it would provide evidence that Head Start may not be causing these labor market changes because the indicator for the discontinuity $ELIG_{i,t}$ is a meaningless indicator.

$ELIG_{i,t}$ for ineligible years ($t = 1, 2$) could only predict program participation for each child if $ELIG_{i,t} = ELIG_{i,4}$. However, given changes in family size, household income fluctuations, and small changes in the eligibility cutoff, precise placement on either side of the discontinuity over time is very unlikely.

[50] Perhaps many mothers are eligible for Head Start because they lost their jobs. This would cause eligible mothers to increase their labor supply even in the absence of Head Start programming, as they revert to their levels of employment before eligibility determination. This case assumes that low labor supply when determining eligibility for Head Start would not have been the steady state in the absence of Head Start.

[51] I do not estimate simulated eligibility at age zero because this would involve imputing information from prior to when the child was born.

Since $ELIG_{i,t}$ is not equal to $ELIG_{i,4}$, it is not instrumenting for program participation.[52] However, $ELIG_{i,t}$ will still measure any differences in unobservable characteristics (such as the examples given in the two cases above) between the eligible and ineligible groups, as the types of people who are eligible and ineligible are likely to remain the same over time. Eligible families will still be poorer, less educated, less likely to be married, less likely to work, and perhaps possess similar unobserved characteristics that accompany these observed characteristics.

Table 14 estimates equations (12) and (13) for both the labor market metrics of hours worked per week and an employment indicator. The columns describe the two cases. The first column for each outcome variable is checking for differences in short term labor market changes between eligible and ineligible groups (case 1), while the second checks for labor market changes at pre-K/kindergarten age specifically (case 2). For each sample, the equations are estimated for eligibility determined for placebo enrollment at age one and age two ($ELIG_{i,t}$ for $t = 1$ and $t = 2$) for both the whole sample and the sample of black families. Estimating these equations at both age one and two provide two points of evidence, but are both seeking to measure the same thing: is $ELIG_{i,t}$ measuring the discontinuity relating to Head Start or the differences between groups not included in the control function? Table 14 provides evidence that the indicator $ELIG_{i,4}$ used for the analysis of labor market outcomes indeed measures the effects of Head Start participation, as the estimates for the coefficient on the eligibility indicator at years that are independent of actual enrollment are close to zero (and sometimes insignificantly negative). This provides confidence in attributing the significant, positive labor market impacts to Head Start.

An important implication of Table 14 is that it provides strong evidence that this empirical strategy is not confounding the effects of Head Start with AFDC or any other means-tested public assistance program, a concern raised earlier in this paper. Head Start is an age-specific program, affecting families only at or after their children enroll at age three or four. If this RD confounded the effects of Head Start with other means-tested public assistance programs, one would observe similar effects regardless of the age of means-testing/eligibility determination. However, as Table 14 indicates, there do not exist similar effects at other ages that could indicate that the results of the observed labor market impacts could be due to any means-tested programs other than Head Start. Moreover, if we were indeed measuring a difference in groups related to the level of public assistance received, transfer-based welfare programs should disincentive work, biasing the results downward. Another final source of bias, unrelated to alternative means-tested programs, and instead resulting from unobservable differences between groups may be that mothers whose earnings fall below the eligibility cutoff work fewer hours because they have characteristics that make them less employable. Because we observe positive movements in labor market participation, these unobservable characteristics bias the result toward zero, and do not confound the effects of Head Start. Table 14 can be interpreted as a crude attempt to quantify these sources of bias.

[52] This is supported empirically; eligibility at ages one and two do significantly predict Head Start enrollment.

Table 14: Identification Check: Discontinuities at Ineligible Ages

	Hours Worked		Mothers' Employment Dummy	
	Case 1: t + 1	Case 2: t = 5	Case 1: t + 1	Case 2: t = 5
All Sample				
	(Age 2)	(Age 5)	(Age 2)	(Age 5)
Head Start Eligible at Age 1	-1.01	-0.754	-0.181	-0.079
	(1.34)	(1.523)	(.112)	(.111)
Observations	2336	2150	2340	2157
	(Age 3)	(Age 5)	(Age 3)	(Age 5)
Head Start Eligible at Age 2	-1.16	-0.705	-0.041	-0.080
	(1.32)	(1.51)	(.115)	(.120)
Observations	2544	2357	2563	2368
Black				
	(Age 2)	(Age 5)	(Age 2)	(Age 5)
Head Start Eligible at Age 1	1.43	-1.52	-0.208	-0.107
	(2.48)	(2.75)	(.239)	(.241)
Observations	886	808	869	784
	(Age 3)	(Age 5)	(Age 3)	(Age 5)
Head Start Eligible at Age 2	-0.93	0.065	0.018	0.054
	(2.15)	(2.34)	(.211)	(.219)
Observations	1023	943	1003	921

Estimated using children as the unit of observation. The specification with hours worked as an outcome variable was estimated using ordinary least squares, and the specification with mothers' employment dummy was estimated with a probit model. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year during which the child was determined eligible (age zero for eligible at age one and age one for eligibility at age two). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC. All of these controls are time-stamped for the year in which eligibility was determined (variables at age zero to determine eligibility at age one, and variables at age one to determine eligibility at age two). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

7.6 Heterogeneous Effects

Heterogeneous effects of both Head Start take-up and labor supply response can provide evidence to better understand the mechanisms behind the results of Head Start's labor supply impact. Although first explicitly discussed in this section, this paper has hinged on heterogeneous effects throughout its analysis: heterogeneity in take-up. The first stage estimates showed that eligibility requirements better predict program participation for black children than for the overall sample; an eligible black child has a 16% probability of enrolling in Head Start if eligible, three times more likely than a 6% probability of enrollment in the overall sample.[53] These heterogeneous effects in take-up have important implications for the strength of the instrument and make it difficult to evaluate the differences in labor force outcomes between groups. Low Head Start take-up within a group does not indicate differential labor market responses for those who do enroll their children, although the weakness of the instrument for the overall sample makes it difficult to confidently scale up these estimates. One of the strengths of this study is that it allows for measurement of the heterogeneous effects at both the level of enrollment and labor supply response.

The differentially large probability of Head Start take-up among black families while substituting from informal care is notable on its own, and merits a brief inquiry into the driving forces behind different levels of program enrollment. An initial hypothesis was that black women historically had higher labor force participation rates than other women, thus making Head Start a more attractive program to balance child-rearing and employment. This would indicate that baseline levels of labor force participation may influence the probability of take-up. However, the inclusion of a term in the estimation of equation (8) interacting eligibility and baseline hours worked yields a (positive) coefficient very close to zero, indicating that initial labor supply level does not seem to be an important determinant of take-up.

Beyond levels of take-up, it is clear that different family circumstances could easily lead to heterogeneous labor market responses to Head Start. Identifying these differences could help better understand the nuances of the labor market changes. Do married women respond differently than female heads of households? Do women respond differently for first-born children than others? What sources of heterogeneity can shed light on the fact that family incomes do not seem to increase, even with increased wage-earning time? Many of the studies discussed in Section 4 identify heterogeneity,[54] further motivating inquiry in this study. However, answering these questions with a fuzzy RD and NLSY79 data poses problems of statistical power due to small sample sizes. The NLSY79 was designed to be a nationally representative sample, so conducting analysis using state-determined eligibility significantly reduces the variability in the data that

[53] Carneiro and Ginja also found heterogeneous effects in their original use of this instrument, finding that Head Start eligibility is a strong instrument for male children but an insignificant predictor of enrollment for female children.

[54] Cascio (2009) and Gelbach (2002) both find differential effects between married mothers and single mothers as well as mothers with and without additional children younger than the enrolling child in their kindergarten studies. Sabol and Chase-Lansdale (2014) find race-based heterogeneous effects, finding that Head Start increases educational attainment only among black parents.

can be used to estimate the effects of Head Start. The requirement of restricting the window to estimate a regression discontinuity reduces the sample size further.

Unfortunately, due to these data constraints, attempts to measure the differential effects on female heads of households and unmarried women proved to be unfruitful. With respect to birth order, I found that among first-born children, eligibility does not predict take-up. For non-first born children, the instrument does predict Head Start enrollment, and I find evidence of similar patterns of labor supply increases for non-first born children, consistent with other studies that find effects for mothers without additional children (Gelbach 2002 and Cascio 2009). These results are presented in Appendix Table A5. While these results still fail to identify heterogeneity in labor supply response, as the labor supply response for first children is unobservable in the data, it does confirm that there do exist labor market effects for mothers who already have older children. Future research to evaluate the impact of Head Start should seek experimental designs that can focus on understanding heterogeneity in labor market response.

8 Conclusion

This study contributed a novel addition to the Head Start impact evaluation and childcare subsidy literature by evaluating Head Start's impact on the labor supply and income of mothers with participating children. Using individually-varying eligibility requirements to estimate a regression discontinuity, I provide the first quasi-experimental evidence of Head Start's impact on mothers' labor force participation and medium-term income growth. In the impact evaluation literature, I challenge and build upon the results of Sabol and Chase-Lansdale's (2014) brief exploration of the labor force outcomes of mothers in the Head Start Impact Study by providing the first evidence of statistically significant labor market effects and extending the scope of Head Start's impact to the entire family. Theoretically, I contribute by providing an argument and evidence of large labor market effects attributable to a 100% childcare subsidy to low-income women. Methodologically, this study is also interesting as an example of exploiting a time-stamped discontinuity to evaluate the impact of a means-tested program without the challenge of a discontinuous budget constraint.

Studying NLSY79 women and children in the 1980s and early 1990s, I find positive movements into the labor market, particularly among black women. In both the whole sample and among black women, various metrics of labor supply move in tandem to show increases in labor force participation of Head Start mothers, which flow into mothers' earnings. The fuzzy RD estimates for black women are large in magnitude, with predicted increases in labor supply of approximately 30 hours per week and an increased probability of employment at levels near 45% as women substitute out of informal care. I observe these large labor market effects in the context of very low levels of labor supply among eligible mothers, implying that Head Start's means-tested nature positions it to provide a strong nudge into the labor force. Future work may explore

heterogeneous responses to Head Start in a number of dimensions or look further into why we do not observe increases in family income in the context of mothers' rising earned income.

The evidence presented in this paper can re-frame our understanding of Head Start's impact and goals as a program. Head Start can strive to change not only children's school readiness in hopes of promoting intergenerational mobility, but can have a contemporaneous impact on the economic situation of participating families as well. This situates Head Start within the portfolio of policy options to assist America's poorest families enter the workforce and exit poverty. While many transfer programs disincentivize work, and many others have explicit work requirements, Head Start is a special form of public assistance that provides the transfer of a valuable service with an implicit incentive to work.

Because this paper studies women in the 1980s, it runs the risk of being outdated: Figure 1 illustrates that women are no longer moving into the labor force to a similar extent. To what extent are eligible women still poised to be impacted by Head Start? Today, the stay-at-home mom is seen as a luxury for the wealthy,[55] not a characteristic of the poor, but in fact, America's poor mothers are more likely to be impacted by Head Start than ever. Over one third of all women who stay at home to look after children are in poverty today. And this figure is growing: 14% of stay-at-home mothers were in poverty in 1970, compared to 34% in 2012 (Cohn and Caumont 2014). From 1970 until today, middle-class women were the driving forces behind large increases in female labor force participation while prohibitively high childcare costs kept poor women at home. If assisting the stay-at-home mother in poverty is a priority, the findings in this paper indicate that Head Start is well-poised to play an important role.

[55]See Goldin (2014) for more discussion on the labor supply of wealthy, highly-educated women.

References

[1] Anderson, Patricia M., and Philip B. Levine. 2000. "Child care and mothers' employment decisions." In *Finding jobs: Work and welfare reform*, ed. Rebecca M. Blank and David Card, 420–62. New York: Russell Sage Foundation.

[2] Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy*, 116(4): 709-745.

[3] Berger, Mark C. and Black, Dan A. 1992. "Child Care Subsidies, Quality of Care, and the Labor Supply of Low-Income, Single Mothers." *The Review of Economics and Statistics*, 74(4): 635-642.

[4] Blau, David M. 2003. "Do Child Care Regulations Affect The Child Care And Labor Markets?" *Journal of Policy Analysis and Management*, 22(3): 443-465.

[5] Blau, David M. and Alison P. Hagy. 1998. "The Demand for Quality in Childcare." *Journal of Political Economy*, 106(1): 104-146.

[6] Blau, David M. and Philip K. Robins. 1988. "Child care costs and family labor supply." *Review of Economics and Statistics*, 70(3): 374–81.

[7] Blau, David M. and Philip K. Robins. 1991. "Child care demand and labor supply of young mothers over time." *Demography*, 28(3): 333–52.

[8] Baum, Charles L. 2002. "Child care costs and work decisions of low-income mothers." *Demography* 39(1): 139–64.

[9] Bureau of Labor Statistics. 2005. "Women in the Labor Force: A Databook."

[10] Card, David, and Dean R Hyslop. 2005. "Estimating the effects of a time-limited earnings subsidy for welfare-leavers." *Econometrica*, 73: 1723–1770.

[11] Carneiro, Pedro and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy*, 6(4): 135-73.

[12] Cascio, Elizabeth. 2009. "Maternal labor supply and the introduction of kindergartens into American public schools." *Journal of Human Resources*, 44: 140-170.

[13] Chaplin, Duncan D., Philip K. Robins, Sandra L. Hofferth, Douglas A. Wissoker and Paul Fronstin. 1999. "The Price Sensitivity of Childcare Demand: A sensitivity Analysis," unpublished manuscript.

[14] Childcare Aware of America. "Parents and the High Cost of Child Care: 2016."

[15] Cleveland, Gordon, Morley Gunderson, and Douglas Hyatt. 1996. "Childcare Costs and the Employment Decision of Women: Canadian Evidence." *Canadian Journal of Economics*, 29(1): 132-148.

[16] Cohn, D'Vera and Andrea Caumont. 2014. "7 Key Findings about Stay at Home Moms." Pew Research Center.

[17] Cogan, John. 1980 "Labor Supply with Time and Money Costs of Participation," in *Female Labor Supply: Theory and Estimation*. Princeton: Princeton University Press.

[18] Connelly, Rachel. 1992. "The effects of child care costs on married women's labor force participation." *Review of Economics and Statistics*, 74(1): 83–90.

[19] Connelly, Rachel and Jean Kimmel. 2003. "Marital Status and Full-time/Part-time Work Status in Childcare Choices." *Applied Economics*, 35(7): 761-77.

[20] Currie, Janet, Eliana Garces and Duncan Thomas. 2002. "Longer Term Effects of Head Start." *The American Economic Review*, 92(4): 999-1012

[21] Currie, Janet and Duncan Thomas, 1995,"Does Head Start make a difference?" *American Economic Review*, 85(3): 341-364.

[22] Currie, Janet and Duncan Thomas, 1999, "Does Head Start help Hispanic children?" *The Journal of Public Economics*, 74: 235-262.

[23] Deming, David. 2009. "Early Childhood Intervention and Life-Cycle Skill Development: Evidence from Head Start," *American Economic Journal: Applied Economics*, 1(3), pages 111-34.

[24] Department of Health and Human Services. 1993. "Creating a 21st century Head Start: Final Report of the Advisory Committee on Head Start Quality and Expansion."

[25] Department of Health and Human Services. 2007. "Head Start: Eligibility, Recruitment, Selection, Enrollment and Attendance."

[26] Department of Health and Human Services. 2010. "Head Start Impact Study: Final Report." Administration for Children and Families, Washington, DC.

[27] Department of Health and Human Services. 2015. "Head Start Program Facts Fiscal Year 2015."

[28] Fronstin, Paul, and Doug Wissoker. 1995. "The effects of the availability of low-cost child care on the labor-supply of low-income women." Working Paper. Washington, D.C.: Urban Institute.

[29] Gelbach, Jonah B. 2002."Public Schooling for Young Children and Maternal Labor Supply." *Economic Review* 92(1): 307-322.

[30] Gladden, Tricia and Taber, Christopher. 2009. "The relationship between wage growth and wage levels." *Journal of Applied Economics*, 24: 914–932.

[31] Goldin, Claudia. 2014. "A Grand Gender Convergence: Its Last Chapter." *American Economic Review*, 104(4): 1091-1119.

[32] Government Accountability Office. 1994. "Child Care Subsidies Increase Likelihood That Low-Income Mothers Will Work." Report to the Congressional Caucus for Women's Issues, House of Representatives.

[33] Guido Imbens and Karthik Kalyanaraman. 2012. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator" *Review of Economic Studies*, 79(3): 933-959.

[34] Hall, R.E. 1973. "Wages, Income and Hours of Work in the U.S. Labor Force" in *Income Maintenance and Labor Supply* ed. Cain and Watts, Chicago: Markham.

[35] Han, Wen-Jui and Jane Waldfogel. 2001. "The effect of child care costs on the employment of single and married mothers." *Social Science Quarterly* 82(3): 552–68.

[36] Juhn, Chinhui and Simon Potter. 2006. "Changes in Labor Force Participation in the United States." *Journal of Economic Perspectives*, 20(3): 27-46.

[37] Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature*," 48: 281-355.

[38] Ludwig, Jens and Douglas Miller. 2007. "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *The Quarterly Journal of Economics* 122(1), 159-208.

[39] Lyndon B. Johnson. "Remarks on Project Head Start." May 18, 1965.

[40] Meunchow, Susan and Susan Shays. 1980. "Head Start in the 1980s. Review and Recommendations. A Report Requested by the President of the United States." Washington, DC: HHS/OHDS/ACYF, Head Start Bureau.

[41] Michalopoulos, Charles, and Philip K. Robins. 2002. "Employment and child care choices of single parent families in Canada and the United States." *Journal of Population Economics* 3: 465–95.

[42] Michalopoulos, Charles, Philip K. Robins, and Irwin Garfinkel 1992. "A Structural Model of Labor Supply and Childcare Demand." *Journal of Human Resources*, 27(1): 166-203.

[43] Pencavel, John. 1986. "Labor supply of men: A survey." In *Handbook of labor economics*, vol. 1, ed. Orley Ashenfelter and Richard Layard, 3–102. Amsterdam: North Holland.

[44] Powell, Lisa M. 1997. "The Impact of Childcare Costs on the Labor Supply of Married Mothers: Evidence from Canada." *Canadian Journal of Economics* 30(3): 577-594.

[45] Ribar, David. 1995." A structural model of child care and the labor supply of married women." *Journal of Labor Economics*, 13(3): 558–97.

[46] Sabol, Terri J. and P. Lindsay Chase-Lansdale. 2014. "The influence of low-income children's participation in Head Start on their parents' education and employment." *Journal of Policy Analysis and Management*, 34(1), 136–161.

[47] Schmit, S., H. Matthews, S. Smith, and T. Robbins, 2013, "Investing in Young Children: A Fact Sheet on Early Care and Education Participation, Access, and Quality. Fact Sheet." National Center for Children in Poverty and Center for Law and Social Policy.

[48] Schulte, Brigid and Alieza Durana. 2016 "The New America Care Report." New America.

[49] Tekin, Erdal. 2007. "Childcare Subsidies, Wages, and Employment of Single Mothers." *The Journal of Human Resources*, 42(2): 453-487.

[50] Zigler, Edward. 2010, "Putting the National Head Start Impact Study into a Proper Perspective." National Head Start Association.

A Appendix

Table A1: Sample vs. NLSY79

	Sample			NLSY79		
Covariate	Obs	Mean	Std. Dev.	Obs	Mean	Std. Dev.
Hispanic	5605	0.202	0.401	11514	0.192	0.394
Black	5605	0.299	0.458	11514	0.277	0.447
White	5605	0.499	0.500	11514	0.529	0.499
Married	5605	0.637	0.481	8860	0.626	0.484
Highest Grade Completed	5604	11.97	2.12	10774	12.29	2.43
Completed High School	5605	0.759	0.438	8776	0.730	0.444
Hours Worked per Week	5532	23.13	19.26	8748	22.14	19.63
Hours Worked per Week of Spouse	3636	43.73	12.51	5670	43.92	12.36
Earned Income of Mother	5601	6268	9317	8816	6951	12.983
Employed	5605	0.663	0.473	8860	0.631	0.482
Weeks Unemployed	5605	2.95	7.96	8596	3.06	8.40
Weeks out of Labor Force	5464	23.74	22.37	8596	24.67	22.65
Family Size	5605	4.33	1.55	8860	4.30	1.63
Number of Children	5534	2.14	1.31	7417	1.97	1.54
AFDC Eligible	5591	0.223	0.417	11258	0.169	0.375
Age at Birth	5605	23.18	3.80	11492	25.24	5.91
Child Eligible for Head Start at age 3	5166	0.341	0.474	8161	0.346	0.476
Child Eligible for Head Start at age 4	5290	0.337	0.473	8023	0.347	0.476

Table A2: First Stage Estimates: Linear Probability Model

Head Start Participation		
	All Sample	Black
Eligible at Age 4	0.064*	0.156**
	(0.035)	(0.062)
Observations	2759	1093
Eligible at Age 3	0.040	-0.013
	(0.034)	(0.062)
Observations	2702	1074
Control Mean	0.235	0.327
Standard Deviation	0.432	0.469

Estimated using ordinary least squares, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, the gender of the child, and state and year dummies. All of these controls are time-stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, all time stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). The control mean refers to the percentage of children that enrolled in Head Start just above the eligibility cutoff specifically, it is the mean of children whose family incomes were between 100% and 125% above of their state, family-size specific income cutoff. The control means are identical for eligibility determined in the panel with eligibility determined at age three and eligibility determined at age four because I report the mean number of children that ever enrolled in Head Start. Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A3: Alternative Labor Force Metrics

	Weeks out of the Labor Force			Weeks Unemployed		
	Child Age			Child Age		
	Age 4	Age 5	Age 6	Age 4	Age 5	Age 6
All Sample						
Head Start Eligible	-2.36**	-1.60	-1.16	0.894	0.791	-0.346
	(1.14)	(1.48)	(1.69)	(.598)	(.703)	(.793)
Observations	2546	2490	2330	2546	2490	2330
Black						
Head Start Eligible	-0.471	-2.72	-2.63	0.426	-.234	-1.77
	(1.98)	(2.50)	(2.86)	(1.20)	(1.40)	(1.71)
Observations	1008	990	920	1008	990	920

Estimated using ordinary least squares, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, the gender of the child, and state and year dummies. All of these controls are time-stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, all time stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A4: Probability of Employment: Two Stage Least Squares Estimates

	Mothers' Employment Dummy		
	Child Age		
	Age 4	Age 5	Age 6
Predicted Head Start Participation	.655**	0.869***	0.715*
	(.307)	(330)	(0.398)
Observations	1003	985	930
Baseline Mean	0.450		
Baseline Std. Dev.	0.498		
Baseline Median	0.00		

Estimated using a two stage least squares model, with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, the gender of the child, and state and year dummies. All of these controls are time-stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, all time stamped for the year prior to eligibility determination (age three for eligibility determined at age four, and age two for eligibility determined at age three). Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A5: Effects for Non First Born Children

(a) First Stage Estimates

	Head Start Participation
Head Start Eligible	.282*
	(.146)
Marginal Effect	*0.070*
Observations	1618
Baseline Mean	0.268
Baseline Std. Dev.	0.443
Baseline Median	0.00

(b) Hours per Week: Reduced Form

	Hours Worked		
	Child Age		
	Age 4	Age 5	Age 6
Head Start Eligible	3.33**	4.415**	4.27**
	(1.52)	(1.73)	(2.16)
Observations	1496	1455	1336
Baseline Mean	14.59		
Baseline Std. Dev.	19.04		
Baseline Median	0.00		

Estimated using a probit model in (a) and ordinary least squares in (b), with children as the unit of observation. Controls not included in the table include a cubic function of log family income, family size, an interaction between log family income and family size, a dummy for the presence of a father figure, the birth order of the child, and the gender of the child, and the state and year. All of these controls are time-stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). Additional labor market controls include the hours worked per week, weeks out of the labor force, weeks unemployed, the earned income of the mother, and a dummy for eligibility or receipt of AFDC, time stamped for the year in which eligibility was determined (variables at age three to determine eligibility at age four). The marginal effect in (a) refers to marginal effect of eligibility on the probability of enrollment in Head Start. The baseline mean and median refer to (a) the average number of children enrolled in Head Start and (b) the average hours worked per week among eligible families with more than one child when eligibility was determined. Robust standard errors (in parenthesis) are clustered at the state-year level. * significant at 10%; ** significant at 5%; *** significant at 1%.