DISCUSSION PAPER SERIES

IZA DP No. 6315

Long Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start

Pedro Carneiro
Rita Ginja

January 2012

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Long Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start

Pedro Carneiro
University College London, Cemmap, IFS, Georgetown University and IZA

Rita Ginja
Uppsala University and IZA

Discussion Paper No. 6315
January 2012

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 6315
January 2012

ABSTRACT

Long Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start*

This paper provides new estimates of the medium and long-term impacts of Head Start on the health and behavioral problems of its participants. We identify these impacts using discontinuities in the probability of participation induced by program eligibility rules. Our strategy allows us to identify the effect of Head Start for the set of individuals in the neighborhoods of multiple discontinuities, which vary with family size, state and year (as opposed to a smaller set of individuals neighboring a single discontinuity). Participation in the program reduces the incidence of behavioral problems, serious health problems and obesity of male children at ages 12 and 13. It also lowers depression and obesity among adolescents, and reduces engagement in criminal activities for young adults.

JEL Classification: C21, I28, I38

Keywords: regression discontinuity design, early childhood development, non-cognitive skills, Head Start

Corresponding author:

Rita Ginja
Department of Economics
Uppsala University
Box 513
751 20 Uppsala
Sweden
E-mail: rita.ginja@nek.uu.se

* We thank Joe Altonji, Sami Berlinski, Richard Blundell, Janet Currie, Julie Cullen, Michael Greenstone, Jeff Grogger, James Heckman, Isabel Horta Correia, Hilary Hoynes, Mikael Lindhal, Jens Ludwig, Costas Meghir, Robert Michael, Kevin Milligan, Lars Nesheim, Jesse Rothstein, Chris Taber, Frank Wjindmeier, and seminar participants at IFS, 2007 EEA Meetings, Universidade Catolica Portuguesa, Banco de Portugal, 2008 RES Conference, the 2008 SOLE meetings, 2008 ESPE Conference, 2008 Annual Meeting of the Portuguese Economic Journal, 2009 Winter Meetings of the Econometric Society, 2011 Nordic Labor Meetings and the 2011 Workshop in Labor Markets, Families and Children (Stavanger) for valuable comments. Pedro Carneiro gratefully acknowledges the financial support from the Leverhulme Trust and the Economic and Social Research Council (grant reference RES-589-28-0001) through the Centre for Microdata Methods and Practice, and the support of the European Research Council through ERC-2009-StG-240910-ROMETA and Orazio Attanasio's ERC-2009 Advanced Grant 249612 "Exiting Long Run Poverty: The Determinants of Asset Accumulation in Developing Countries". Rita Ginja acknowledges the support of Fundacao para a Ciencia e Tecnologia and the Royal Economic Society.

1 Introduction

The need to cut public spending, together with recent disappointing evaluations of Head Start and Sure Start, have put severe pressure on compensatory preschool programs both in the US and the UK. Opponents call for the outright termination of these programs, while supporters argue that they are needed now more than never, as increasing numbers of families fall into poverty. Others propose maintaining the programs, as long as they are subject to comprehensive reform.

The Head Start Impact Study (HSIS) gained prominence in this debate. It evaluates Head Start, the main compensatory preschool program in the US, and it is the first experimental study of a large scale program of this kind in the world. The study shows that Head Start has short term impacts on the cognitive and socio-emotional development of its participants, which disappear by first grade. While there are grounds on which this study can be criticized (e.g., Zigler, 2010), its main findings are compelling because of its transparent design.[1] In parallel, an evaluation of Sure Start in the UK, although non-experimental and less influential than the Head Start Impact Study, finds that Sure Start also has limited impacts on the development of poor children.

Our paper shows that in spite of the lack of program impacts by first grade, there are important longer term impacts of Head Start on the health and criminal behavior of adolescents and young adults. Relatively to comparable non-participants, Head Start participants are 24% less likely to suffer from a chronic condition that requires the use of special equipment (such as a brace, crutches, a wheelchair, special shoes, a helmet, a special bed, a breathing mask, an air filter, or a catheter), 29% less likely to be obese at 12-13 years of age, less likely to show symptoms of depression at ages 16-17, and 31% less likely to have been in a correctional facility by ages 20-21. Our results are in line with the growing literature on the effectiveness of early childhood interventions, which shows that these programs have large long term impacts on behavioral problems even when they have limited short term impacts on cognitive development. Short term evaluations of early childhood programs miss most of their potential impacts.

We identify the causal effects of Head Start using a (fuzzy) regression discontinuity design which explores the eligibility rules to the program. We determine the eligibility status for each child aged 3 to 5, by examining whether her family's income is above or below an income eligibility cutoff, which varies with year, state, family size, and family structure. This is a new empirical strategy to study the effects of the program, which allows us to identify the effect of marginally relaxing the thresholds determining eligibility into the program. In our data, the marginal entrant affected by a relaxation of eligibility criteria is a boy, aged 4, who is more likely to be African-American than white or Hispanic.

In contrast with standard applications of regression discontinuity, there are multiple discontinuity points in our setup, which vary across families because they depend on year, state, family size and family structure. Therefore, our estimates are not limited to individuals located in the neighborhood of a single

[1] Another experimental evaluation of Early Head Start (DHHS, 2006), a program for children ages 0-3, also shows small program impacts.

discontinuity, but they are applicable to a more general population.

Beyond the HSIS (DHHS, 2010), described above as showing little or no effect of the program, there exist several non-experimental evaluations of Head Start which are also important, and it is worthwhile mentioning some of the most recent ones. Currie and Thomas (1995, 1999, 2000) compare siblings in families where at least one sibling attends Head Start and one does not. In contrast to HSIS, they find strong impacts of the program on a cognitive test (which fade-out for blacks, but not whites) and grade repetition. They use the Children of the National Longitudinal Survey of Youth (CNLSY), which is also used in our paper. Currie, Garces and Thomas (2002) apply a similar strategy in the Panel Study of Income Dynamics (PSID), and show that the program has long lasting impacts on schooling achievement of adults, earnings, and crime. Also, relying on within family comparisons and using the CNLSY, Deming (2009) finds no effects on crime but positive effects on a summary measure of children's test scores and adult outcomes.

Ludwig and Miller (2007) explore a discontinuity in Head Start funding across US counties, at the time the program was launched (1965). They show that Head Start has positive impacts on adolescents' and adults' health and schooling.[2]

Relatively to all these studies, we evaluate a more recent variant of the program (and employ a novel empirical strategy). This is relevant because Head Start has changed over the years and its costs have dramatically increased, closely approaching the costs of model interventions such as Perry Pre-School or Abecedarian. Furthermore, it means that, relatively to the studies mentioned above, ours is more comparable to the recent Head Start Impact Study, which examines children who applied for Head Start in 2002. Ludwig and Miller (2007), and Garces, Currie and Thomas (2002), study the effects of attendance between the mid-1960s and the 1970s. Currie and Thomas (1995), and Deming (2009), analyze effects of Head Start for those who attended the program during the 1980s. Individuals in our sample enrol in the program from the 1980s to the late 1990s.[3]

This paper proceeds as follows. In the next section we describe Head Start in more detail. We present the data in Section 4. We discuss the identification strategy in Section 3. Results are presented in Section 5. Section 6 presents a simple cost-benefit calculation. Section 7 concludes.

[2]In addition, Currie and Neidell (2007) use the CNLSY to study the quality of Head Start centers and find a positive association between scores in cognitive tests and county spending in the program. They also find that children in programs that devote higher shares of the budget to education and health have fewer behavioral problems and are less likely to have repeated a grade. Frisvold and Lumeng (2007) explore an unexpected reduction in Head Start funding in Michigan to show strong effects of the program on obesity. Neidell and Waldfogel (2006) argue that ignoring spillover effects resulting from interactions between Head Start and non-Head Start children and/or parents underestimates the effects of the program in cognitive scores and grade repetition. Finally, Anderson, Foster and Frisvold (2010) find that Head Start is associated with a reduction in the probability that young adults smoke.

[3]There exist a few studies in the literature examining the long term impact of universal pre-school (Cascio, 2009, Magnuson et al, 2007, Berlinski et al, 2008, 2009, Havnes and Mogstad, 2011). They concern programs that affect a much larger fraction of the population, and generally show long term impacts of preschool availability.

2 Background: The Head Start Program

Head Start started in 1965 as part of President Johnson's War on Poverty and currently provides comprehensive education, health, nutrition, and parent involvement services to around 900,000 low-income children 0 to 5 years of age (of which 90% were 3-5 years old in the FY of 2009[4]) and their families. Recognizing the importance of the earliest years of life, in 1994, the Early Head Start program was established to serve children from birth to two years of age.

Head Start programs are primarily funded federally but grantees must provide at least 20 percent of the funding, which may include in-kind contributions, such as facilities to hold classes. These programs are administered locally which leads to a considerable degree of heterogeneity in service delivery. Program costs, which include teacher salaries, vary considerably since some grantees may receive donations, such as low-cost space. Different grantees may also have widely different costs of personnel and space depending on many factors, such as geographic location (urban or rural), and type of sponsoring agency (school system or private nonprofit)[5]. Salaries generally comprise most of Head Start grantees' budgets, and grantees' teacher salary levels differ based on factors such as location and staff qualifications (GAO, 2003b).

Although there is a large degree of heterogeneity across programs, each Head Start center must comply with publicly known standards which are described in the Head Start Act (the last version dates of December 12, 2007 and re-authorizes the program through September 30, 2012). For example, centers may offer one or more out of three program options: center-based option, home-based option, or combination program option. The chosen program option must meet the needs of the children and families served, considering factors as the child's age, developmental level, disabilities, health or learning problems, previous preschool experiences, family situation, and parents' concerns and wishes. Each of the three options above must comply with the following rules. *Center-based* programs operate four/five days per week, between a minimum of 3.5 and 6 hours per day (full day programs operate 6 to 12 hours per day). Programs that operate for four (five) days per week must provide at least 128 (160) days per year of planned class operations. The program should offer a minimum of 32 weeks of scheduled days of class operations over an eight or nine months period. The *home-based* option should provide one home visit per week (a minimum of 32 home visits per year) lasting for a minimum of 1.5 hours and they should be conducted with the presence of the parents. At least two group socialization activities per month should be provided for each child (a minimum of 16 group socialization activities each year)[6]. During the home visits, the visitor should work with parents to help them provide learning opportunities that enhance their childs growth and development, whereas the socialization activities should be focused on both the chil-

[4]According to the Head Start Office, in 2009, among those 3-5 years old, 36% of children were 3 years, 51% were 4 and 3% were 5 years old.

[5]By 1989, just over one third of grantees and delegate agencies were primarily community action agencies; 28% were run by private, nonprofit organizations and 19% by public schools. The other grantees and delegate agencies were state or local government agencies, religious organizations, and other organizations, particularly tribal organizations (GAO, 1989).

[6]According to the performance standards, the average caseload should be of 10 to 12 families per home visitor with a maximum of 12 families for any individual home visitor.

dren and the parents. Finally, the *combination* option should offer an equivalent to the services provided through the center-based program option or the home-based program option, over a period of 8 to 12 months, with the acceptable combination of home visits class sessions clearly stated in the performance standards.

Head Start has been regarded as a part-day, part-year program (see GAO, 1989, Blau and Currie, 2006). However, during the 1990s it shifted towards a full day program, and by 2003 27% of all those enrolled were served by full-day programs with length of 6 to 8 hours a day, 20% were enrolled in full-day centers for 8 hours or more per day, 44% were enrolled in part-day center-based programs for less than 6 hours a day, and the remainder 9% were enrolled in the home-based option (GAO, 2003). In the data used in our analysis we are unable to identify which of these options a child attended, however, based on these numbers and also on numbers for previous years (GAO, 1981 and GAO, 1998) it is likely that most of the children participated in the center-based option.

Staff employed by centers must also comply with minimum standards. As of 1996, basic standards require that center-based programs employ two paid staff persons for each class,[7] and the directions point to children-staff ratio of about 8:1. Zigler (2010) points out that a typical Head Start classroom consists of approximately 17 children with one BA-level teacher trained in early childhood education and one assistant teacher[8]. To put these numbers in context, the Perry Preschool Program had a teacher-student ratio of one teacher for every 5.7 students, the Carolina Abecedarian Project ranged from 3 per teacher for infants to 6:1 at age 5 (Cunha, Heckman, Lochner and Masterov, 2006), and there were between 8-12 children per teacher in the Chicago Child Parent Center and Expansion Program (Fuerst and Fuerst, 1993).

Teacher qualifications are an important dimension of heterogeneity across programs, and Congress mandated in 1998 that 50% of all Head Start teachers should have a degree by September, 2003. Although this goal was achieved on average, there is no information about whether there exists a teacher with the minimum credentials in every classroom (GAO, 2003b). One impediment to the improvement of qualifications of Head Start's teachers is the difficulty of grantees to compete for teachers with degrees, since grantees are unable to match the salaries of other preschool teachers (in the late 1990, Head Start salaries were about half of what kindergarten teachers earned nationally). Additionally, Head Start teacher salaries varied by credentials and type of grantee administering the program. Teachers in programs administered by school systems (which included about 12% of the teachers in 2003) have on average a higher level of education, higher salaries, and a lower turnover than those in programs administered by other types of agencies.

Finally, the criteria governing the selection of children into Head Start are well defined and advertised regularly by the Head Start Office through Program Instructions. These instructions also alert to possible

[7] The two staff members would be a teacher and a teacher aide, or two teachers. Whenever possible, there should be a third person in the classroom who is a volunteer.

[8] The exact recommendations regarding the class size vary with the age of children. The standards of operation suggest that class of 4-5 years old should have between 17-20 children and 3-years old classes should have between 15-17 children.

frauds, either on the side of parents or centers (who may be tempted to serve less troublesome children).[9] Children are eligible to participate if they are of preschool or kindergarten age and if they live in poverty. In addition, at least 10% of the children served per center must have some type of disability. Since these selection criteria are explored in our identification strategy, we defer the explanation of details to Section 3. Eligibility criteria have been mostly unchanged since the 1971, covering the entire period we analyze (1981-2004). This is documented in Table A.1 in the Appendix, which also shows the main pieces of relevant legislation.

Figure 1 displays enrollment in Head Start since it was launched in the 1960s. It shows a steady and slow increase during the 1970s and 1980s, and a sharp increase in the early 1990s. Between 1991 and 1995 there was an increase of about 25% in the number of children served by the program (from 583,000 to 750,000). Simultaneously, there was an increase in the funding per child. In the early 1990s the program was reaching about 500,000 children per year, at a federal cost equivalent to $US5,400/child (in $US2009), whereas in the FY of 2009, Head Start operated 49,200 classrooms serving almost 1 million children at federal cost of $US7,600 per child (the staff included 212,000 paid workers and 1.2 million volunteers). While it is generally assumed that Head Start is funded at much lower levels than Perry Preschool or Carolina Abecedarian (see Blau and Currie, 2006, Deming, 2009), recent data shows that this is no longer true. The effort to expand and improve the program means that today its costs per child are reaching those of Perry Preschool. According to Heckman et al. (2010) the estimates of initial costs of Perry Preschool (presented in Barnett, 1996), which include both operating costs (teacher salaries and administrative costs) and capital costs (classrooms and facilities) reached $17,759/child over its two years (in 2006 $US), so that the current version of Head Start costs about 85% of Perry Preschool.

3 Empirical Strategy

Our goal is to estimate β from the following equation:

$$Y_i = \alpha + \beta HS_i + f(X_i) + \varepsilon_i \tag{1}$$

where Y_i is the outcome of interest for child i, which in our paper is measured between ages 6 and 21, HS_i is a dummy variable indicating whether the child ever participated in Head Start, X_i is a vector of controls (entering through function $f(X)$), and ε_i is an unobservable. β is the impact of Head Start on Y which, in principle, can vary across individuals. Even if β does not vary across individuals, the estimation of this equation by ordinary least squares (OLS) is problematic. On one end, since Head Start participants are poor, they are likely to have low levels of ε_i, inducing a negative correlation between HS_i and ε_i. On the other end, not all poor children participate in the program, and perhaps only the most motivated mothers enrol their children in it, which would create a positive correlation between HS_i and ε_i.

[9] This is advertised on the Head Start's Office web site: http://www.acf.hhs.gov/programs/ohs (consulted in September 14, 2011).

Program eligibility rules In order to address these problems we explore discontinuities in program participation (as a function of income) that result from program eligibility rules. Children ages 3 to 5 are eligible if their family income is below the federal poverty guidelines, or if their family is eligible for public assistance: AFDC (TANF, after 1996) and SSI (DHHS, 2011).[10] Once a family becomes eligible in one program-year, it is also considered eligible for the subsequent program-year.

We construct each child's income eligibility status in the following way (a detailed description can be found in Appendix C). First, the poverty status is imputed by comparing family income with the relevant federal poverty line, which varies with family size and year (Social Security Administration, 2011). Second, eligibility for AFDC/TANF requires satisfying two income tests, and additional categorical requirements, all of which are state specific. In particular, the *gross income test* requires that total family income must be below a multiple of the state specific threshold, that is set annually and by family size at the state level.[11] The second income test to be verified by applicants (but not by current recipients) is the *countable income test*, that requires total family income minus some disregards to be below the state threshold for eligibility (U.S. Congress, 1994). In addition, AFDC families must obey a particular structure: either they are female-headed families, or they are families where the main earner is unemployed. This means that children in two-parents households may still be eligible for AFDC under the AFDC-Unemployed Parent program. In turn, eligibility for AFDC-UP is limited to those families in which the principal wage earner is unemployed but has a past work history, so we consider eligible those whose father (or step-father) worked on average less than 100 hours per month in the previous calendar year.[12]

Finally, we do not impute SSI eligibility for two reasons. First, imputing SSI eligibility would require the imputation of categorical requirements which are complex to determine (e.g., Daly and Burkhauser, 2002), some of which we are unable to observe in the data.[13] The literature has showed that classification errors are likely to happen (see Benitez-Silva, Buchinsky and Rust, 2003). Second, SSI thresholds are below Poverty Guidelines and therefore these thresholds will not be binding (see U.S. Congress, 2004).

When using regression discontinuity it is only possible to identify program impacts in the neighborhood of the cutoff. Since we explore multiple discontinuities, it is helpful to know the range of neighborhoods of income over which we can identify program impacts. Figure 2 displays the distribution of cutoff values for households income over which there is variation in our data, which corresponds to the support

[10] AFDC denotes Aid to Families with Dependent Children, TANF denotes Temporary Assistance for Needy Families, and SSI denotes Supplemental Security Income.

[11] When this test was established in 1981 the multiple was set to 1.5. The Deficit Reduction Act of 1984 raised this limit to 1.85 of the state need standard.

[12] Since 1971, Federal regulations have specified that an AFDC parent must work fewer than 100 hours in a month to be classified as unemployed, unless hours are of a temporary nature for intermittent work and the individual met the 100-hour rule in the two preceding months and is expected to meet it the following month. Attachment to the labor force is one condition of eligibility for AFDC-UP. See U.S. Congress, 1994, for the specific requirements.

[13] There are five stages to assess the categorical requirements to receive SSI through disability. For instance, in the third stage, it is required that the applicant has any impairment that meets the medical listings, conditional on the fact that he/she is not engaging in a substantial gainful activity and has an impairment expected to last for more than 12 months. We do not have accurate information to impute this using NLSY79 (there are variables on whether health limits amount and kind of work an individual can perform, but not to which extent they fulfill medical listings).

of income values for which we are able to identify the effects of Head Start. Income cutoffs also vary across different family sizes, and in Figure D.1 in the Appendix B we plot the joint support of household income and family size over which we are able to estimate the relevant program effect.[14] Regarding the two possible sources of eligibility to the program (via federal poverty line, or AFDC/TANF), about 98% of the children in our sample have eligibility determined by the federal poverty line criterion.[15]

It is important to mention that eligibility rules for Head Start are not perfectly enforced (some ineligible children are able to enrol), and that take up rates among those eligible are far below 100%. There are several factors that influence the take up of social programs, such as shortage of funding to serve all eligible, barriers to enrollment and social stigma associated with participation (e.g., Currie, 2006, Moffitt, 1983). Due to limited funding, Head Start enrolls less than 60 percent of all 3 and 4 years old children in poverty. This is visible both in the CNLSY and in the CPS.[16]

The number of eligible individuals is also different from the number of actual participants because of lack of perfect enforcement of eligibility rules, and of other factors affecting participation. In the case of Head Start, centers may enrol up to 10 percent of children from families whose income is above the threshold (without any cap on the income of these families). Thus, the discontinuity in the probability of take-up of Head Start around the income eligibility threshold is not sharp, but ”fuzzy” (see Hahn, Todd and van der Klauww, 2001, Battistin and Rettore, 2007, and Imbens and Lemieux, 2007).

In summary, a child can enrol in Head Start at ages 3, 4, or 5 and it is possible to construct eligibility status at each of these ages. As we show in Section 5, eligibility at age 4 is a better predictor of program participation than either eligibility at 3 or at 5, and in our data (and in the administrative records from the Head Start Office) 50-60% children enrol in Head Start when they are 4 (Head Start Office, 2011). Therefore we focus on eligibility at age 4 in our main specification, but we also present results with eligibility at other ages. In our modeling of the outcome as a function of the forcing variable we rely on series estimation (widely used in other applications of this empirical strategy), restricting the sample to values of the forcing variable that are close to the highest and the lowest cutoff points.

We start by estimating the following reduced form model:

$$Y_i = \phi + \gamma E_i + f(Z_i, X_i) + u_i \tag{2}$$

[14]Table A.2 in the Appendix A illustrates how cutoffs vary with family size. It includes the number of observations per cell year-family size and the average cutoff values (as of US$2000).

[15]The states where some children are eligible via AFDC/TANF over poverty line are: Arizona, Arkansas, Illinois, Michigan, Mississippi, Nebraska, New York, Ohio, Texas, Washington and Wisconsin. The last two have the highest proportion of children eligible via AFDC/TANF as opposed to poverty: 81% and 87.5%, respectively.

[16]The problem of imperfect compliance is not unique to Head Start, but common across social programs. Only 2/3 of eligible single mothers used AFDC (Blank and Ruggles, 1996); 69 percent of eligible households for the Food Stamps program participated in 1994 (Currie, 2006); of the 31 percent of all American children eligible for Medicaid in 1996, only 22.6 percent were enrolled (Gruber, 2003); EITC has an exceptionally high take-up rate of over 80 percent among eligible taxpayers (Scholz, 1994); in 1998, participation in WIC (the Special Supplemental Nutrition Program for Women, Infants and Children) among those eligible was 73 percent for infants, 2/3 among pregnant women and 38 percent for children (Bitler, Currie and Scholz, 2003).

where E_i is an indicator of eligibility for Head Start, X_i is a set of all determinants of eligibility for each child except for family income (year, state, family size, family structure, measured at age 4), Z_i is family income (at age 4), and u_i is the unobservable.[17] We include state effects in our models not only because the criteria for eligibility are state-dependent but also to account for cross-state unobserved differences in generosity and services provided. The equation for E_i is:

$$E_i = 1\left[Z_i \leq \bar{Z}(X_i)\right], \tag{3}$$

where $1[.]$ denotes the indicator function.

$f(Z_i, X_i)$ is specified as a parametric but flexible function of its arguments, and $\bar{Z}(X_i)$ is a deterministic (and known) function that returns the income eligibility cutoff for a family with characteristics X_i (constructed from the eligibility rules). In section 5 we study the sensitivity of our results to the choice of different functional forms for $f(Z_i, X_i)$. We use probit models whenever the outcome of interest is binary (the linear probability model is especially inadequate when mean outcomes are far from 50 percent, which occurs frequently in our data; see Table 1).

Three conditions need to hold for γ to be informative about the effects of Head Start on children outcomes. First, after controlling flexibly for all the determinants of eligibility, E_i must predict participation in the program, which we show to be true.

Second, families are not able to manipulate household income around the eligibility cutoff. This is the main assumption behind any regression discontinuity design. It is likely to hold in our case because the formulas for determining eligibility cutoffs are complex, and depend on family size, family structure, state and year, making it difficult for a family to position itself just above or just below the cutoff. Still, in order to guard against the possibility of income manipulation, there are standard ways to test for violations of this assumption (e.g., Imbens and Lemieux, 2007), and below we discuss them in detail.

Third, eligibility to Head Start should not be correlated with eligibility to other programs that also affect child outcomes. This assumption is potentially more likely to be violated than the first two, because there are other means tested programs which have eligibility criteria similar to those of Head Start (e.g., AFDC, SSI, or Food Stamps). In order to assess the importance of this problem we implement the following test. While most welfare programs exist throughout the child's life, Head Start only exists when the child is between the ages of 3 and 5. If other programs affect outcomes of children, then eligibility to those programs in ages other than 3 to 5 should also affect children's outcomes. In contrast, if eligibility is correlated with children's outcomes only when measured between ages 3 and 5, then it probably reflects the effect of Head Start alone. Although we cannot definitely rule out the possibility that other programs confound the effects of Head Start (by operating exactly at the same ages), the results we present below are highly suggestive that this is not the case.

An additional problem is that, at first sight, the control group is not clearly defined. We consider two alternatives to Head Start: other preschool, and home (or informal) care. We show that individuals

[17] $f(Z_i, X_i)$ can be a different function in each side of the discontinuity. We empirically examine this case below.

induced to enter into Head Start because of a shift in eligibility status come almost exclusively out of home (or other informal) care if they are white or Hispanic, or from other formal preschool arrangements if they are African-American.

Notice that γ does not correspond to the impact of Head Start on the outcome of interest, because eligibility does not fully predict participation (imperfect compliance). In order to determine the program impact, we estimate the following system, for the case where Y_i is continuous:

$$Y_i = \alpha + \beta HS_i + g(Z_i, X_i) + \varepsilon_i \quad (4)$$

$$HS_i = 1[\eta + \tau E_i + h(Z_i, X_i) + v_i > 0], \quad (5)$$

where equation (5) is estimated using a probit model (van der Klauww, 2002). $1[.]$ denotes the indicator function. $P_i = \Pr(HS_i = 1|E_i, Z_i, X_i)$ is estimated in a first stage regression, and used to instrument for HS_i in a second stage instrumental variable regression (van der Klauww, 2002, Hahn, Todd and van der Klauww, 2001). If Y_i is binary we use a bivariate probit. $g(.)$ and $h(.)$ are flexible functions of (Z_i, X_i). In Appendix D we also discuss how we can identify heterogeneous effects of Head Start. Unfortunately, even though our estimates of heterogeneous effects are interesting they are also imprecise.

4 Data

We use data from the Children of the National Longitudinal Survey of Youth of 1979 (CNLSY), which is a survey derived from the National Longitudinal Survey of Youth (NLSY79). The NLSY79 is a panel of individuals whose age was between 14 and 21 by December 31, 1978 (of whom approximately 50 percent are women). The survey has been carried out since 1979 and we use data up to 2008 (interviews were annual up to 1994, and have been carried out every two years after that). The CNLSY is a biennial survey which began in 1986 and contains information about cognitive, social and behavioral development of the children of the women surveyed in the NSLY79 (assembled through a battery of age specific instruments), from birth to early adulthood.

Children 3 to 5 years of age are eligible to participate in the program if their family income is below an income threshold, which varies with household characteristics, state of residence, and year. Among the variables available in CNLSY there are those that determine income eligibility (total family income[18], family size, state of residence, Head Start cohort and an indicator of the presence of a father-figure in the child's household[19]) along with outcomes at different ages. For reasons explained in Section 3, we will focus mainly in the outcomes of children eligible for the program at age 4. In our data, the earliest year in which we can construct eligibility at age four is 1979 (for children born in 1975), since this is the first

[18] Monetary variables are measured in 2000 values using the CPI-U from the Economic Report of the President (2006).

[19] Although father's (or stepfather) employment is also a condition that determines Head Start eligibility, we did not consider it, because the variable "number of weeks mother's spouse worked" has missing values in half of the observations. Inclusion of this variable and an indicator for missing values does not change the results.

year in which income is measured in the survey (eligibility each year is determined by last year's income, which is precisely what is asked in the survey). Since we take outcomes measured at ages 6 and older, the youngest child in the sample is born in 2000 (after imposing additional sample restrictions). Therefore, we study the effects of participating in Head Start throughout the 1980s and 1990s.

Out of the 11,495 children surveyed by 2008 (corresponding to 137,940 observations between 1986-2008), we drop 2285 children for whom we cannot identify participation in Head Start between ages 3 to 5. This is our treatment variable and program attendance is constructed using information collected since 1988. The survey asks whether the child currently attends nursery school or a preschool program, or whether she has ever been enrolled in preschool, day care, or Head Start.[20] For participants we use the age at which the child first attended Head Start and the length of time attending the program to construct an indicator of Head Start attendance between ages 3 to 5.[21]

We recover information about preschool attendance from the question "Ever enrolled in preschool?". The data allows us to consider three alternative child care arrangements between ages 3 to 5: $HS_i = 1[Ever\ in\ HS]$, $OP_i = 1[Ever\ enrolled\ in\ other\ preschool]$, and $Home_i = 1[Never\ in\ HS\ or\ other\ preschool]$, where $1[.]$ is the indicator function (HS_i is the indicator for Head Start participation, OP_i denotes "Other preschool" and $Home_i$ denotes some other child care arrangement).

We further drop 1917 children for whom we are unable to assess income-eligibility status at age 4 because of lack of information on family income, family size, state of residence or regarding mother's co-habitational status.

Finally, we drop 858 children without information on income and family size before age 3 and birth weight. These variables are used as controls and we show in Section 5 that our results are not sensitive to the exclusion of these pre-determined control variables. We end up with 6372 children which are observed at least once between the ages of 6 and 21 (the relevant age group for our analysis) and born between 1977 and 2000. They could have had eligibility assessed for Head Start between 1981 and 2004. Of these, 3247 children are boys.

As mentioned, we distinguish three possible preschool arrangements: Head Start, other preschool programs, or neither of the previous two (informal care at home or elsewhere). In the raw data, 90 percent of mothers who report that their child was enrolled in Head Start also report that their child was enrolled in preschool, possibly confounding the two child care arrangements. Therefore, as in Currie and Thomas (1995, 2000), we recode the preschool variable so that whenever a mother reports both Head Start and preschool participation, we assume enrollment in Head Start alone. After recoding this variable, almost 20 percent (1299 children) of the children in the sample ever enrolled in Head Start, 40 percent

[20]The children to whom we do not observe participation in Head Start are surveyed on average less times than those without missing information (4.79 times out of 12 vs 10.64, respectively), they have lower average family income between the years of 1979-2008 ($32844.96 and $41243.36 for those with missing and non-missing Head Start, respectively), they are less likely to be Black (20% among those missing information vs 30% among children with Head Start information) and they have on average 1.6 siblings (whereas those with non-missing information have 2 siblings).

[21]The specific questions used to construct the indicator of Head Start attendance are: "Child ever enrolled in Head Start program?", "Child's age when first attended Head Start?" and "How long was child in Head Start?".

(3770 children) attended other type of preschool, and the remaining attended neither. In our data, about 40% of participants enter Head Start at age 3, and 50% enter at age 4. In the CNLSY, 90% of Head Start participants attend the program for at most one year.[22]

Since we rely on a discontinuity in the probability of participation around a threshold, it is good practice to restrict the sample to children whose family income at eligible age was near the income eligibility cutoff for the program since points away from the discontinuity should have no weight in the estimation of program impacts (see e.g., Black, Galdo, and Smith, 2005, Lee and Lemieux, 2010). Therefore, we focus on the sample of children whose income was between 15% and 185% the relevant income cutoff (we also present estimates using alternative intervals for income).

Table 1 summarizes the data. The sample we use consists of 1676 males for whom at least one of the measured outcomes is available and all the control variables used in the regressions are not missing (child care arrangement at ages 3 to 5, eligibility to Head Start at age 4, family log income and family size at age 4 and at ages 0 to 2, presence of a father or stepfather in the household, state of residence at age 4, and birth weight). We also discuss some results for females, but for reasons that become clear below, the bulk of our paper focuses on males. The table presents the variables in groups according to whether they are family or child variables, and according to the age at which outcomes are measured. It shows means, standard deviations and the number of available observations for each variable.

It is clear that the children in our sample come from fairly disadvantaged backgrounds. 40% of their mothers are high school dropouts, and only 10% ever enrolled in college (although not presented in table, these figures are 50% and 22%, respectively, when we use all children in the CNLSY). Their average annual family income is only slightly above $18000 (deflated to 2000; as opposed to $40586 for the whole sample), 9.7% of children are reported to have been of low birth weight, 30% of these children were enrolled in Head Start, 50% were in other types of preschool, and 20% were in neither[23]. 18.8% of the children at ages 12-13 were overweight, 27% were in special education, 33.6% had repeated a grade, and 5.4% need to use special equipment because of a health limitation. At ages 16-17 the proportion of children who were overweight was 17.8%, 13% had been drunk at least once, and 14% were ever arrested or convicted at least once up to that age. At ages 20-21, 14% were overweight, 14% had been drunk once

[22] A back-of-envelope calculation, suggests that based on official numbers we would expect the Head Start participation rate to be around 5% in the 1980s and early 1990s, but 8% in 2000. That is, according to the US CENSUS for 1980, 1990 and 2000, about 20% of children ages 3 to 5 in the US are poor, which amounts to 1,663,440 (out of 9,207,040), 2,021,299 (out of 10,275,120) and 1,836,383 (out of 10,601,578) children for the years of 1980, 1990 and 2000, respectively (although the definition of poverty in CENSUS is based on the poverty thresholds, whereas eligibility to Head Start is determined by the poverty guidelines), and for these years the number of children enrolled in the program is 376,300, 540,930 and 857,664. We have a larger estimate in our data, possibly because of two characteristics to the sampling of NLSY: (1) about 50% of our sample is an oversample of minorities and poor whites available in data and (2) the CNLSY contains an overestimate of children from young mothers. This explains why our number is comparable to the 19.4% Figure (in Currie and Thomas, 1995, who use the same data source. Currie, Garces and Thomas, 2002, estimate Head Start participation at 10% in the PSID, and Ludwig and Miller, 2007, have participation rates of 20 to 40% in the counties close to their relevant discontinuity (based on data from the National Educational Longitudinal Study). As a further note, the NLSY79 also includes a subsample of members of the military, which we exclude from our work.

[23] In the whole sample of children 8% of children are reported to have been of low birth weight, 21% of these children were enrolled in Head Start, 60% were in other types of preschool, and 19% were in neither

before, 25.7% had been arrested or convicted at least once, and only 59.3% had a high school diploma.[24]

5 Results

5.1 First Stage Estimates

We start by checking whether the discontinuity in eligibility status also induces a discontinuity in the probability of Head Start participation, by estimating equation (5) (participation equation). We present estimates for the three main samples we analyze (children ages 12-13, adolescents 16-17 and young adults 20-21) and by gender. Table 2 presents estimates of τ in equation (5) for the sample of children 12-13 years old[25] when eligibility is measured at age four, the age at which most children first enrol in Head Start.[26] The marginal effect included is the average marginal change in participation as a result of a change in the eligibility status, which is defined by:

$$\frac{1}{N}\sum_{i=1}^{N}\left\{\Pr\left(HS_i=1|E_i=1,Z_i,X_i\right)-\Pr\left(HS_i=1|E_i=0,Z_i,X_i\right)\right\}=$$
$$\frac{1}{N}\sum_{i=1}^{N}\left[\Phi(\eta+\tau+h\left(Z_i,X_i\right))-\Phi(\eta+h\left(Z_i,X_i\right))\right]$$

where N is the number of children in the regression sample, and Φ is the standard normal distribution function (we get similar results if take the average of marginal effects using observations only in a small neighborhood of each cutoff). Function $h\left(Z_i,X_i\right)$ consists of a cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure (father or step-father) in the household at age 4, indicators for gender, race and age, and indicators for year and state of residence at age 4. All standard errors in the paper are clustered at the level of the state-year, since eligibility rules are determined at this level and clustering accounts for the correlation between outcomes of children within each eligibility cell (below we also present estimates where clustering is done only at the state level).

It is clear from Table 2, that across age groups, eligibility at age four is a strong predictor of program participation for males, although the estimated effect is well below 100%. This is primarily an indication of weak take-up of the program at the margin of eligibility (common to many social programs), which could be partially driven by the fact some children start the program at either ages three or five when they

[24]Some of the indices reported in the table but not discussed in the text are the AFQT (Armed Forces Qualifying Test, which is a good measure of cognitive ability), the BPI, and the CESD. BPI is the Behavior Problems Index and it measures the frequency, range, and type of childhood behavior problems for children age four and over (Peterson and Zill, 1986). The Behavior Problems total score is based on responses from the mothers to 28 questions that intent to measure (1) antisocial behavior, (2) anxiety and depression, (3) headstrongness, (4) hyperactivity, (5) immaturity, (6) dependency, and (7) peer conflict/social withdrawal. The CESD (Center for Epidemiological Studies Depression) scale measures symptoms of depression and it discriminates between clinically depressed individuals and others.

[25]Estimates for the remaining samples ages 16-17 and 20-21 and by race, are included in Appendix A in table A.4.

[26]See Office of Head Start, fact sheet for 2010.

are also eligible, but it is likely to be mainly the result of several other factors, such as lack of available funds to cover all eligible children (since Head Start was never fully funded), stigma associated with program participation (Moffitt, 1983), or the fact that most of the centers are only part-day programs, and thus unable to satisfy the needs of working families (Currie, 2006). Our paper is novel in obtaining estimates of how the take-up of Head Start changes for individuals near the eligibility threshold as the eligibility status change. This can be interpreted as the increase in participation when thresholds are relaxed by a very small amount.[27]

Table 3 shows why the remainder of the paper focuses on eligibility at age four as the main determinant of participation in Head Start: eligibility at age four is a better predictor of participation than either eligibility at age 3 or eligibility at age 5. Therefore, the population of children for whom we are able to estimate the impact of Head Start are those at the margin of eligibility at age four and it is likely to consist of children who suffer income shocks between the ages of 3 and 5 (we account fully for these shocks through our set of controls). We are not able to estimate the impact of Head Start on those who are permanently and substantially below the poverty line.

Interestingly, changes in eligibility status are not associated with changes in participation in Head Start for females. This result holds across races, as reported in table A.4 in Appendix A. It is difficult to understand why there is such a gender discrepancy, and we come back to this below. The fact that the change in eligibility status is only associated with changes in participation for boys and not for girls suggests that the marginal entrant in Head Start is a boy. It also implies that we cannot provide an assessment of whether the effects of Head Start are important for girls. In the appendix we also report that the discontinuity in the probability of participation is larger for Black boys than for non-Blacks, so the marginal entrant is more likely to be Black. The size of the discontinuity in the probability of participation among males is robust to exclusion of the oversample of minorities and poor whites from the analysis.

When using a RD setup it is standard practice to present a graphical analysis of the problem. Relatively to the standard setting which has a single discontinuity, our setup makes use of a range of discontinuities. One graphical representation of the problem which does not correspond exactly to the specification of our model takes a measure of family income relative to each family's income eligibility cutoff, and defines this variable as "distance to the eligibility cutoff". Figure 3 plots Head Start participation at age 4 for males and females entering our analysis of outcomes at ages 12-13, 16-17 and 20-21, against the relative distance of family income to the income eligibility cutoff (at age 4). We divide the sample

[27]Most of the evidence of how newly eligible to social programs respond in terms of participation comes from Medicaid expansions throughout the 1980s and early 1990s. Cutler and Gruber (1996) and Currie and Gruber (1996) estimate that only 23 and 34 percent of newly eligible children and women of childbearing age take-up Medicaid coverage, as many were already covered by other insurance. In our sample, 40% of eligible four year olds not attending Head Start were enrolled in another preschool program. Card and Shore-Sheppard (2002) find that expansion of Medicaid eligibility to children whose family income was below 133 percent of the poverty line had no effects on the decision of take-up, whereas the expansion of eligibility to all poor children led to an increase of nearly 10 percent in Medicaid coverage. LoSasso and Buchmueller (2002) estimate that take-up rates among newly eligible children for SCHIP (State Children's Health Insurance Program) ranged between 8 and 14 percent.

into bins of this variable (of size 0.05) and compute cell means for participation. We draw a vertical line at zero (point of discontinuity), and we run local linear regressions of each variable on the distance to cutoff on either side of the discontinuity (bandwidth = 0.3; Appendix A includes the same picture for bandwidths 0.2 and 0.4 in Figures B.1 and B.2)[28]. These figures suggest discontinuities of about 15% in program participation at the eligibility cutoff for the sample of boys, but no jump in the probability of participation for females. This is exactly what our regressions show in Table 2.

It is puzzling that the first stage relationship for girls is so much weaker than for boys. Figure 3 shows that for all age groups, the average participation in Head Start is a steeper function of distance to the cutoff for girls than for boys, and that this function does not jump at the cutoff for girls. It would seem that parents are less sensitive to child care costs for girls than for boys, since they do not jump at the opportunity of enrolling the girl in a quality preschool program when it suddenly becomes free.[29] On the supply side, it is possible that there could be discretion on the part of centers to enrol boys and girls based on how easy they are to care for. However, this explanation is not plausible unless we believe boys are easier to educate than girls. In fact, precisely the opposite may be taking place on the demand side. If boys are more difficult to care for, parents may more jump more eagerly at the opportunity to enrol them in child care when it becomes freely available than they would have for girls.

Defining the control group In order to be able to interpret our results it is central to understand in which type of child care would children enrol in the absence of the program. This defines the ”control group” in our study. As we explained in Section 4, we consider three possible child care arrangements between ages 3 to 5: ”Head Start”, ”Other Preschool”, ”Informal care”. Table 4 shows how participation in these three alternative child care arrangements responds to eligibility. We regress the dummy variables indicating participation in each type of child care on eligibility and the remaining control variables. There are three panels in the table, corresponding to three different populations: those for whom we have outcomes at ages 12-13 (the youngest cohort), those with outcomes at 16-17, and those with outcomes at ages 20-21 (the oldest cohort). Columns 2 and 3 show that, for the youngest cohort, when an individual becomes Head Start eligible there is a statistically significant movement out of ”Other Preschool”. In contrast, columns 4-6 show instead that children in slightly older cohorts are more likely to leave ”Informal Care” when they become eligible for Head Start. Finally, for the oldest cohort of children (columns 7-9), there is movement out of both ”Other Preschool” and ”Informal Care” in response to a change in eligibility status, but movement out of the ”Informal Care” seems to be relatively more important. We

[28]Figure B.3 in Appendix B complements this picture and it presents the number of children within each interval of 0.05 to the distance to the threshold.

[29]Following the same approach that we use below to define the control group, we find that, unlike what happens for boys, when girls become eligible for Head Start there are no shifts out of informal care or preschool into the program. This holds for Black and non-Black children for all samples analyzed (ages 12-13, 16-17 and 20-21), with exception of the sample of Black girls aged 12-13 years old who leave informal care when become eligible. Within families we do not find that when eligibility status changes, parents act react differently towards marginally eligible boys and girls in terms of enrolment in the program (see table F.4), but this could be due to lack of power when we restrict the sample to families with at least a boy and a girl who were around the income eligibility cutoff at age 4.

show in the table A.5 in Appendix that when the analysis is separated by child's race, then Black children (columns 1-3) that become eligible are more likely to leave "Other Preschool" than non-Black children.[30] It is useful to contrast our control groups with those used in previous studies. Currie and Thomas (1995), Currie, Garces and Thomas (2002), and Deming (2009) compare siblings that attended Head Start vs. either "Other Preschool" or "Other type of care". In contrast, the HSIS, 2010, compares Head Start children with children in the waiting lists of about 80 centers, who attended a mixture of alternative care settings (around 60% of children in the control group participated in some type of child care or early education programs during the first year of the study, with 13.8% and 17.8% of the 4 and 3-year-old in the control group, respectively, participating in Head Start itself).

5.2 Validity of the Procedure

Our identifying assumption is that children *just* above the income eligibility cutoff are similar to those *just* below it in all dimensions except program participation.[31] A priori this is a plausible assumption, but there exist incentives for a family to try to manipulate eligibility. For example, a family just above the income cutoff could try to underreport income in order to become just eligible. Similarly, Head Start providers who know the eligibility rules well, and who have a desire to serve children who are easy to care for, may try to game the system in order to accept a large proportion of those children who are just ineligible. Fortunately, there are several sources of information on which we can draw on to understand (and ultimately dismiss) the importance of these concerns.

Eligibility and pre-determined variables We start this section with a standard check of the validity of our identifying assumptions. We take a set of pre-program variables that should not be affected by participation in the program, and we use them as dependent variables in equation (2). If our procedure is valid then the estimate of γ should be equal to zero. These variables are: the child's average MOTOR score before she turned three (a measure of the physical and social development for very young children), mother's education, birth weight, maternal grandmother's education, marital status of the mother before the child turned 3, mother's AFQT score, average log family income and family size between the ages of 0 and 2, and several variables related to the mother's family environment when she was 14 years old (whether the mother lived in a Southern state, whether she lived with her parents, how many siblings she had, and whether she lived in a rural area). Eligibility is measured at age 4, as explained above. The results are presented in Table 5, which focuses on boys for whom we observe outcomes at ages 12-13. Results for other older age groups and for difference race groups are similar (they are shown in the Appendix A, table A.6)).

Table 5 shows that our procedure is valid. Most estimates of γ are small (compared with the mean and standard deviation of each variable also included in table), and almost all of them are statistically in-

[30] The estimates for the marginal change in the take-up of the three child care alternatives do not change if a multinomial logit model is estimated instead of separate probit models for each choice.

[31] We thank Jens Ludwig for detailed comments and valuable suggestions on this section.

significant.[32] If the p-value is adjusted for multiple hypothesis testing, following the procedure suggested in Romano and Wolf (2005) then we cannot reject the hypothesis that there is no significant relationship between any of these variables and eligibility, even in the case of the two statistically significant coefficients in the table (birth weight and family income before age 3).[33] Figure B.4 in the Appendix shows local linear regression estimates similar to those in figure 3, but using variables taken before child turned three as dependent variables. Visual inspection of these figures yields similar conclusions to those in table 5.

Throughout the rest of the paper we augment our basic specification of $f(Z_i, X_i)$ with some of these variables as additional covariates, since they are useful to reduce sampling error and small sample bias (e.g., Lee and Lemieux, 2010). In particular, we add a cubic on log of average family income and average family size between ages 0 and 2, an interaction between the two, and a cubic on the child's birth weight.

Bunching of observations around cutoff An alternative and more direct test of manipulation, developed by McCrary (2007), checks whether there is bunching of individuals just before the discontinuity. This test is not practical with multiple discontinuities unless we have a large sample size. However, when we implement it using a single discontinuity (using percentage distance to the eligibility cutoff as the running variable) we find no evidence of income manipulation, as shown in figure B.5 in Appendix.

5.3 Estimates from the Reduced Form Equations

Tables 6-8 are the central tables of our paper. They present estimates of equation (2) (a regression of each outcome on eligibility) for the main set of outcomes. These tables include an indication of whether the null of no effect is rejected when we account for multiple hypothesis testing. Throughout our discussion we consider that the program has a statistically significant effect on a specific outcome only in the cases where we can reject the null that the effect is zero using the procedure of Romano and Wolf (2005).

We focus on several measures of social outcomes, health assessments and school achievement for children ages 12-13 (see Table 6). These include drug use, being overweight, grade retention, participation in special education, alcohol and tobacco use, damage of school property, and serious health limitations. For each outcome we report (in each line of the table) the coefficient on eligibility, its stan-

[32]In order to better understand the magnitude of these estimates we conducted the following exercise. Take a few of our main outcomes of interest, such as BPI at ages 12-13, and CESD by ages 16-17. Then regress each outcome on each of the variables in table (5), and compute predicted values for each regression. We can now rerun the regressions on table (5) using these predicted values instead of the variables that generated them, allowing us to translate the coefficients in table (5) into magnitudes of the outcomes of interest. We do not report this in a table, but describe the results briefly in the text (for all boys): in terms of BPI, all the coefficients in table (5) are between -0.0035 and 0.014 (expressed as a fraction of a standard deviation), and for CESD up to ages 16 to 17 they are between -0.0068 and 0.007 (expressed as a fraction of a standard deviation). All these figures are very small.

[33]Since we are examining the impact of a program on multiple variables (as opposed to a single variable) we need to account for that when doing hypothesis testing. Several multiple hypothesis testing procedures exist, but the most recent one is developed in Romano and Wolf (2005), which accounts for non-independence across outcomes, and has more power than most of its predecessors (namely Westfall and Young, 1993). We apply their procedure which is described in detail in Appendix E.

dard error, whether the hypothesis that the coefficient is equal to zero is rejected at the 10% level of significance using the algorithm of Romano and Wolf,[34] the number of observations in each regression, the mean of the outcome for those individuals below the cutoff (control mean), the standard error of the outcome for those below the cutoff, and the average marginal effect of eligibility on the outcome being analyzed in each column. We find that Head Start eligibility leads to a reduction in the probability of having a health condition that requires the use of special equipment (such as a brace, crutches, a wheelchair, special shoes, a helmet, a special bed, a breathing mask, an air filter, or a catheter) and having a condition that requires medical attention.[35] In addition, it leads to a reduction in the probability of being overweight, and on behavior problems as measured by the BPI.[36]

Table 7 shows estimates of the impact of eligibility to Head Start on outcomes for adolescents ages 16-17. For this age group we examine enrollment in high school, measures of engagement in criminal activity, obesity, sexual behavior and use of birth control, alcohol use, self-reported health status, mental health indicators (which include an index of symptoms of depression called the Center for Epidemiologic Studies Depression Scale - CESD, and some of its individual components), and expectations about adult outcomes (such as marital status, family stability and employment by age 35). We find that eligibility to Head Start is associated with a decrease in the probability of being overweight, and with less symptoms of depression, as measured by the CESD.

For young adults ages 20-21 years we study self-reported health status, use of birth control, sexual activity, obesity, alcohol abuse, criminal activities, college attendance and high school graduation. Table 8 shows that HS eligibility leads to a decrease in the probability of ever being arrested or convicted for a crime by ages 20-21.[37]

Sensitivity to functional form and sample size Table 9 shows (for a selected set of outcomes) that these results are robust to changes in the set of controls we use, and changes in the degree of the poly-

[34] We also report results for standard tests of the hypothesis that each coefficient is zero that do not take into account the fact that we are examining multiple outcomes, using the familiar procedure of including star symbols next to each coefficient to indicate statistical significance at different levels.

[35] According to Zigler (2010), about 12% of the children enrolled have some type of disability.

[36] One extra health related outcome we analyzed was the frequency of dental check-ups. This is an important outcome as one of the services provided to Head Start children, which we would expect to become part of family's habits. Additionally, this is one outcome where the Head Start Impact Study, 2010, found effect sustained until the end of kindergarten. We did not find any effects on whether the child has had any dental check either the last 12 or 24 months at ages 6-7, 9-10 or 12-13. We do not report these outcomes in our main tables as information on dental check-ups is only available since 1992, and the sample size in estimations is about 75% of that used for the other outcomes for these age groups. Another set of outcomes which are not reported in the main tables but are available in the appendix are test scores (see table A.7). Our results are imprecise, but overall we do not find evidence of impacts of Head Start participation on test scores. Although this is not consistent with the findings of Currie and Thomas (1995) and Deming (2009), it is somewhat consistent with the findings of HSIS. The children in our analysis are a mixture of the older cohorts studied in Currie and Thomas (1995) and Deming (2009) and the younger cohorts in the HSIS, so our results could be close to either of these sets of studies. In contrast with Currie and Thomas (1995) our focus is not on the widely analyzed PPVT because this test is administered fairly infrequently when compared to the other tests we study. Nevertheless, our results are essentially the same when we analyze the PPVT.

[37] In the Appendix F we discuss whether parental investments in children react to participation in Head Start, not only because this program has a parental component, but also because Head Start can be perceived to be either substitute or complement of parental investments.

nomial in income and family size at age 4. Column (1) presents our preferred specification. In column (2) we exclude several control variables from the model, namely those corresponding to pre-age 4 characteristics, while in column (3) we expand the set of pre-age 4 variables we include in the model (see the note to the table). The changes in the results relatively to column (1) are fairly minor. In columns (4) and (5) we change the order of the polynomial in income and family size, from cubic, to either quadratic or quartic. Again, our basic results barely change.

Table 10 shows the sensitivity of our results to the size of the window of data used around the discontinuity. We construct these intervals based on values of family income as a proportion of the household specific cutoff. Column (4) is the benchmark displaying our main results. The other columns present different window sizes going from very small (column (1)) to very wide (column (6)). If the window is very small so is the sample size, and the estimates potentially become more noisy. If the window is too large we are using large amounts of data that are not very relevant for the parameter of interest and which can instead make the problem of misspecified polynomial much worse. What we see in this table is that results are robust to reasonable changes in window size.

Confounding effects of other programs As mentioned in section 3, eligibility to Head Start is correlated with eligibility to other programs, such as AFDC, Medicaid, or SSI (see Table A.8 in the Appendix, which shows that our eligibility variable is also a good predictor of participation in these other programs). It is therefore possible that the estimates in tables 6-8 confound the effects of Head Start with those of other programs. However, while most of these programs exist during several years of the child's life, Head Start is only available when the child is between ages 3 and 5. This fact allows us to assess whether confounding effects from other programs are likely to be important. Our reasoning is as follows. Suppose that we estimate equation (2) using eligibility (as well as the covariates) measured at different ages of the child. If participation in other programs is driving our results, E_i should have a strong coefficient even when measured at ages other than 3 to 5. Otherwise, we can be confident that our estimates reflect the impact of Head Start, since it is (possible but) unlikely that other programs affect child development only if the child enrols at ages 3 to 5, but have no effect if she enrols either at ages 0, 1, 2, 6 or 7.[38]

Table 11 presents estimates of the impact of eligibility to Head Start at different ages on a representative set of outcomes, one for each panel. Each column represents a different regression, where the age of eligibility (and the corresponding controls) varies from 0 to 7. Across panels, the largest and strongest estimates occur consistently at age 4, and sometimes 5 (BPI and overweight in panels A.1, A.2 and B.1), while for all other ages the coefficients are generally small and insignificant (with only these exceptions). We take this as evidence that by using our procedure we are capturing the impact of Head Start and not of other programs.

[38]This reasoning will work if the set of individuals who are at the margin of eligibility at ages 3 to 5, are different from those who are at the margin of eligibility at ages 0, 1, 2, 6 and 7. If they were all the same individuals it would be impossible to distinguish eligibility to Head Start (only at ages 3 to 5) from eligibility to other programs (at all ages). Furthermore, it is not possible to rule out that other programs have most of its influence at ages 3-5.

Graphical Analysis We now present figures analogous to figure 3, but instead of focusing on participation in Head Start, in the vertical axis we represent our main outcomes. These are shown in figure 4. As before, we use a bandwidth equal to 0.3, and in the appendix we present results for bandwidths equal to 0.2 and 0.4 (figures B.6-B.7). The figures suggest that there are discontinuities in the level of the outcomes we represent at the eligibility cutoffs, and they have the same sign as those reported in the tables above. However, for all the outcomes we consider there is a fair amount of oscillations in both sides of the discontinuity.[39]

Discontinuities in outcomes at non-discontinuity points Table A.10 in the Appendix presents results of a test of whether our method does not spuriously detect non-zero impacts where impacts do not exist. In columns (1) and (2) we present estimates of size of the discontinuity when we set the cutoff to be 50% above and below, respectively, of the original cutoff. There should be no jump in outcomes at these points. We estimate that only one out of the 12 coefficients that should be zero is significant, and it has the opposite sign to the original estimate.

However, the sample used columns (1) and (2) includes the real cutoff and we are smoothing the regression function at a point where it is known to have a discontinuity. Therefore, as suggested in Imbens and Lemieux, 2007, we include in columns (3) and (4) the fake and the real discontinuities. Only two of the 12 coefficients that should be zero are found to be statistically different from zero.

Sensitivity to alternative clustering of the standard errors To account for the fact that the variation in eligibility to Head Start is set at the state-year level, our main results cluster standard errors at the state-year level (by year we mean the year at which the child was age 4). This accounts for common shocks to children who were 4 in the same year and lived in the same state. We also re-estimate our main specifications of tables (6)-(8) clustering the standard errors only by the state of residence at age four, to account for serial correlation in state specific shocks. These estimates are presented in table A.11. They show estimates of the standard errors of the coefficients which are very similar to those in our main results, although the number of groups is very different under both specifications (46 groups under the assumption of within state serial correlations at state level and 715 under the assumption of no-serial correlation).

Keeping the same sample across age groups Out of the 1873 boys included in our main analysis (presented in tables 6-8) only 601 were surveyed at all ages (12-13, 16-17 and 20-21). Therefore, when analyzing the evolution of results across ages our estimates confound age and cohort effects. It is useful to see how our results would change if we were to use exactly the same sample of children to estimate

[39]When we redo the regressions in tables 6-8 using distance to the eligibility cutoff as the running variable instead of family income and family size we get estimates which are slightly smaller than the ones we report as our basic specification, but always of the same sign and similar magnitude. Although they do not remain statistically significant for the outcomes where we saw the strongest effects at ages 12-13, they remain statistically significant at ages 16-17 and ages 20-21. These results are shown in table A.9 in Appendix.

the impacts of Head Start at different ages, since we would then be able to isolate age effects (assuming time effects are not important), but only for a limited set of cohorts. Our estimates for this exercise are presented in Table A.12, and show that our main conclusions hold in this smaller sample, although they become more imprecise. By restricting our sample to children present in the three main age groups which we study, we consider only those children born between 1977 and 1988, and who could have attended the program between 1981 and 1992 (the same cohort analyzed in Currie and Thomas, 1995, and Deming, 2009).

5.4 Measurement Error in Income

Our estimates of the size of discontinuity in participation and effects of Head Start at the cutoff might be underestimated due to measurement error in the running variable. This is especially problematic here because we are measuring income using survey data. Therefore, we assess the robustness of our results to alternative measures of family income constructed in the CNLSY.

The family income measure we use in our main specification is the Total Net Family Income for the years of 1979-2004. This variable includes the following components if all are not missing: (1) respondent's and spouse's military income, (2) respondent's and spouse's wages/salaries, (3) respondent's and spouse's farm/business, (4) respondent's and spouse's unemployment insurance, (5) alimony, (6) child support (own and spouse's), (7) AFDC, (8) other public assistance, (9) respondent's and spouse's educational benefits, which may include VEAP/educational benefits from military (educational benefits for veterans under the G.I. Bill or VEAP) and other kinds of scholarships, fellowships, or grants (own and spouse's), (10) other veteran benefits, (11) worker compensation or disability payments received by respondent, (12) total amount of money received by the respondent (or wife/husband) from any source excluding the previous categories, such as savings, payments from Social Security, net rental income or any other regular or periodic sources of income, (13) total income received by the respondent from adults that also live in the household and are related to her (excluding spouse and children)[40], (14) total amount of money received by the respondent (or wife/husband) from persons living outside household (if living in the same dwelling), outside her home in city of permanent residence (if living in Dorm, fraternity or sorority) or not living with respondent (if in military), and (15) food stamps.

We consider two additional income measures in our sensitivity analysis. The first one is the same as the one above subtracting Food Stamps, which is perhaps a better approximation to the measure of income officially used to determine eligibility than the one we use in the baseline specification. The second one is the sum of the non-missing income components described in the previous paragraph. This is very much like the baseline measure of income used in the paper, but with the inclusion of non-missing components as opposed to requiring that each of them is not missing to allow for a larger sample size. The three income measures analyzed in this paper are highly correlated but they are not exactly the same.

[40]Including AFDC, SSI, other public assistance or welfare from local, state or federal government, unemployment compensation, worker's compensation, or veteran benefits.

Currie and Cole (1994) document inconsistencies in the income measures reported in NLSY79. Therefore, we follow their recommendations to minimize these inconsistencies. First, we check if income categories reported by siblings living with parents are consistent among them. If not, we choose the mode among the values reported by the siblings. If there is no unique mode, we use the value reported by the oldest child that lives with parents (this consistency check is only performed between 1979-1986, which are the years with information for whether the member of NLSY79 is living with the parents). Second, whenever we use individual components of income, we identify those cases in which both spouses are present in the NLSY79, and if the reported earnings for one spouse are not consistent with the own report, we use the own report.

Estimates of the equation describing participation in Head Start when we use these two alternative measures of income are included in Tables A.13. These are very similar to those reported in Table 2. The impact of Head Start eligibility on outcomes estimated for alternative income measures is presented in Tables A.14 and A.15 for the samples ages 12-13, 16-17 and 20-21. There is some sensitivity across different measures of income, but overall the main results remain the same.

5.5 Estimates from the Structural Equations

The reduced form analysis of table 6 tells us that there are strong effects of eligibility to Head Start on behavior problems, on being overweight, and on the need to use special health equipment at ages 12 and 13. Table 7 shows strong effects on depression and obesity, and table 8 shows large effects on criminal activity. These three tables summarize our main results, but the estimates in these tables do not correspond to the quantitative impact of the program on individuals because compliance with the program is imperfect, and eligibility does not equal participation. These estimates need to be scaled up by the estimated effect of eligibility on participation, and the best way of doing this is to estimate equation (4) jointly with (5) (Lee and Lemieux, 2010). In doing so, the estimated effects became quite imprecise, reflecting some instability in the procedure. In spite of this, in all cases but one the essential patterns of tables (6)-(8) remain unchanged.[41]

Table 12 shows estimates of β coming from the system consisting of (4) and (5) and for a related set of outcomes, which accounts for selection into the program. The table reports estimates of β, as well as average marginal effects of Head Start on outcomes (labeled *Marginal Effect*). For discrete outcomes, the latter is:

[41]Behind the instability problem may be the fact that either one or both equations in this system are non-linear and our specifications include a large number of location and time indicator variables. This is particularly true when we estimate bivariate probit models, which involve maximizing non-concave likelihood functions with more than one local maximum. For each outcome we started the optimization routine using the estimates where Head Start participation is considered exogenous (that is, those estimates included in odd columns of each table), and the results we report correspond to the maximum values of the likelihood that we found. The optimization algorithm used for each outcome is presented in the note of table 12.

$$\frac{1}{N}\sum_{i=1}^{N}\{\Pr[Y_i=1|HS_i=1,Z_i,X_i]-\Pr[Y_i=1|HS_i=0,Z_i,X_i]\}=\frac{1}{N}\sum_{i=1}^{N}\{\Phi[\alpha+\beta+g(Z_i,X_i)]-\Phi[\alpha+g(Z_i,X_i)]\}.$$

At ages 12-13, we estimate that participation in Head Start leads to a 29% reduction in the probability of being overweight, a 24% reduction in the probability of needing special health equipment, and a 0.6 standard deviation decrease in the behavior problems index for the whole sample. At ages 16-17, surprisingly we cannot find any impact of the program on being overweight (perhaps because of numerical difficulties in our procedure), but we estimate that the program leads to a 0.55 standard deviations decrease in the depression score. Finally, at ages 20 or 21 we find a 30% reduction in the probability of ever being arrested or convicted of a crime.[42]

In summary, table 12 and tables 6-8 (and the subsequent sensitivity analysis) present a picture of strong effects of Head Start on behavioral and health outcomes of children, which are sustained at least until early adulthood. It is interesting that in the case of behavioral outcomes we were able to find a consistent set of large and statistically significant results, while that is not true for cognitive outcomes (as in the HSIS). As stressed by Cameron, Heckman, Knudsen and Schonkoff (2007), this may be due to the fact that non-cognitive skills are more plastic than cognitive skills, and early childhood interventions are more likely to have sustained effects on the former than on the latter.[43]

5.6 Heterogeneity

Race As we mentioned above, there are differences across race groups in the composition of the control group (other preschool vs. informal care). For Black children in the younger sample (outcomes at ages 12-13), the control group is mainly ”other preschool”, while those in older groups (ages 16-17 and 20-21) are more likely to leave ”informal care” when become eligible. For non-Black children, the alternative to Head Start is more likely to be some type of informal care, regardless of the age group (see Table A.5). Therefore, it is possible that children from different race groups benefit differently form exposure to Head Start.

In table A.18 in the Appendix we present reduced form estimates for the main outcomes analyzed by

[42]In table A.16 in the Appendix we also include estimates using a linear probability model for discrete outcomes. These results show that our findings hold under a linear regression model.

[43]For comparison, in the appendix we also present estimates of the impacts of Head Start participation versus pre-school and other arrangements using a siblings comparison strategy, as in Currie and Thomas (1995) and Deming (2009). These results are included in table A.17 in Appendix. Because we focus on a different cohort of participants than these papers, we present three columns for each outcome: (1) for children that could have mainly attended the program in the 1980s, born in 1972-1986, (2) for children that could have enrolled in the 1990s, and (3) for the entire sample (we focus on children that could have been eligible to attend the program in the 1980s and 1990s, whereas the youngest child in Currie and Thomas (1995) and Deming (2009) was born in 1986). We present estimates for three outcomes: (1) an index created following Deming (2009) which is the average of PIAT-Math, PIAT-Reading Recognition and PPVT, (2) for PIAT-RR and (3) for BPI. Our findings are not the same as in those papers but this is mainly because we have greatly extended the sample to include younger cohorts of children.

child's race. We find that for children ages 12-13 the effects on reduction of chronic health conditions are present on both races, but HS impacts on BPI are driven by the non-Black, whereas the effects on obesity come from the Black sample. One additional effect (and robust to multiple hypothesis testing) found for non-Black children is a decrease in the probability of enrolment in special education, which is consistent with improvement in white children's school performance also find in Currie and Thomas (1995). Among adolescents 16-17, the effects on being overweight are driven by Black adolescents (which is consistent with the findings of Frisvold, 2011), whereas those on mental health come from the non-Black. Finally, the effects on crime related activity among young adults are due to reduced engagement in criminal activities by those who are non-Black.

We choose not to include the results by child's race in the main text given the substantial decrease in precision when the sample is divided, so that most effects are no longer robust to multiple hypothesis testing (except at ages 12-13).

Cohort differences It is interesting to check whether there are differences in program impacts across cohorts of children, because they may tell us something about changes in the program over time. In doing this exercise, one needs to be careful about changes occurring in the control group across cohorts. In particular, we have mentioned before (table 4) that, relatively to the younger cohorts, there is a stronger substitution of Head Start for informal care in the older cohorts. This is consistent with the recent expansion in state provided care. The estimates we discuss in this section are shown in table A.19 in Appendix. We focus the discussion on results for the youngest age group (12-13), for which we have cohort variation.

We separated children 12-13 years old into two groups: those that could have been eligible to attended the program in the 1980s (born between 1977 and 1984), and those that could have attended it in the 1990s (born in 1985-2000). The reduced form estimates for five outcomes are presented in Table A.19: indicators for a health condition that requires the use of special equipment, being overweight, grade retention, enrolment in special education classes, and an index of behavioral problems (BPI). The estimates show that most of the effects we find in the overall sample are driven by the set of children that attended the program in the 1980s. This is interesting and perhaps surprising, given the large increase in expenditure in Head Start in subsequent years, but which can be rationalized by the change in the control over time.[44]

[44]Since eligibility cutoffs vary with income and family size, in principle we could examine how the impacts of Head Start varied with these two variables, within the support of the cutoffs. This would be an extra dimension of heterogeneity. Our attempts to do so resulted in imprecise estimates, as we show in the Appendix D. Estimates for equation (7) are presented in Figures (D.2)-(D.4). The results in appendix show larger effects for children living in larger families (see probability of being overweight and use of special equipment), but when we separate the analysis by race we obtain the opposite effect (larger effect of children living in smaller families and with higher income; not presented).

6 A simple cost-benefit calculation

We now present a simple cost-benefit calculation of a marginal expansion in the eligibility to Head Start, achieved through a small increase in eligibility thresholds. As we showed in Table 2, eligibility to Head Start is a associated with an increase in the probability of participation at the margin of 21% among males. The calculations are much simpler to present if we focus on a single cohort of participants, and we choose the cohort of children entering Head Start in 1992, as it is the most contemporaneous year that allows an analysis to all three age groups we studied. As a comparison, we also consider the cost-effectiveness of the version of the program available in 2003, computing the costs in that year but assuming the same expected benefits associated to the 1992-cohort. All monetary values are measured in 2009 dollars.

6.1 Costs

We computed the cost of attendance per child in 1992 using the federal appropriation and the number of enrolled children obtained from the Head Start Fact Sheets. These figures were $1,439,903,924 for the former and 621,078 for the later. Note that this analysis is valid for a specific cohort of Head Start participants: children born in 1988. We then discount the annual cost of attendance at age four to age 0 (using an interest rate of 4%). When comparing costs and benefits we take into account that eligibility does not imply participation and vice-versa. Our calculations assume that a marginal increase in eligibility thresholds leads to a 21% increase in the probability of participating in Head Start.

6.2 Benefits

Health We showed that Head Start is associated with improvements in mental health among adolescents, a reduction in the incidence of a chronic condition that requires the use of special equipment (a brace, crutches, a wheelchair, special shoes, a helmet, a special bed, a breathing mask, an air filter, or a catheter), and a reduction in the probability of being overweight. We provide an estimate for the lower bound of the benefits associated with an improvement in health conditions, by considering only the savings associated with reduction in obesity (for which we can get more reliable figures).

Childhood and adolescence obesity increases the risk of developing diseases such as high cholesterol, hypertension, respiratory ailments, orthopedic problems, depression, and type 2 diabetes[45]. Childhood obesity is estimated to cost $14 billion annually in direct health expenses and one third of all children and teenagers in the US are obese, which adds to more than 23 million children and teenagers (see NCCOR, 2009). Assuming an equal cost across all obese children, the annual per capita cost of obesity is about $609.

However, it is likely that an individual is obese for more than one year. Furthermore, obese adolescents have a 70 percent chance of becoming overweight or obese adults, which increases to 80 percent if

[45]See http://aspe.hhs.gov/health/reports/child_obesity/. Access on January 3, 2012.

at least one parent is overweight or obese[46]. To account for this persistence we also estimate the present value of costs for a child who is obese throughout adolescence and young adulthood, in particular between ages 13 and 30. Therefore, even if we exclude the days lost of work and college due to conditions related to obesity, the discounted value of direct health expenses with diseases associated to obesity between the ages of 13 and 30 adds up to $4,815 (see Panel ”Alternative 1” of Table 13). Using the more conservative measure presented in the previous paragraph, the present value of savings associated with a reduced probability of being obese at just age 13 adds up to $366 (see Panel ”Alternative 2” of Table 13).

Crime To compute the savings associated to reduction in criminality we use the direct expenditures in criminal justice for 2007, which are part of *Justice Expenditure and Employment Statistical Extracts for 2007*.[47] The series we use are national level data from which it is possible to obtain disaggregated expenditures according to different activities. We start by using the direct expenditures per year for criminal corrections, which include correctional supervision (adults supervised in the community on probation or parole and those incarcerated in state or federal prisons and local jails). The total expenditure in correctional supervision adds up to 72 billion dollars (in 2009 USD). This figure ignores direct judicial and police expenditures, which together with correctional supervision add up to $219 billion dollars. The latter number ignores victims costs, which are accounted in the cost-benefit analysis of Perry Preschool Program performed in Heckman et al. (2009).[48]

We obtain the savings associated with reduction in criminal activity in just one year, in which the individual is taken to be at 20 years old. Additionally, we assume uniform costs across all types of criminal activities and locations. In this respect, our analysis is similar to Belfield et al. (2006) for Perry Preschool.

To obtain the cost per criminal offence we use the total number of offenders supervised by correctional authorities in 2007. This figure added up to 7,267,500 offenders. The average cost by offender including only the costs associated with supervision is about $9,867 per year. Discounted to age 0 these represent $4,683 (see Panel ”Alternative 2”). If we use broader definition of costs including also court and policy costs, then the average annual cost per offender reaches $30,269 ($14,364 after discounted for age 0). The marginal individual due to expanding eligibility by $1 is 22% less likely to have been sentenced of any criminal charge by age 20, therefore the expected saving in justice functions due to reduced criminal activity ranges between $1,030 and $3,160, depending on whether we use a narrower or wider definition of expenditures.

[46]See `http://aspe.hhs.gov/health/reports/child_obesity/`. Access on January 3, 2012.

[47]We choose to measure expenses associated with criminal activity in 2007 because this is the most recent available data for expenditures by criminal function and we are interested in the analysis for the cohort of children born in 1988, who are 20 years old in 2008.

[48]Victimization costs are incurred when the crime committed produces victims. Heckman et al. (2009) estimated that this is about three times the judicial costs, although their numbers might be inflated by the existence of murders in their sample.

Summary The net benefits of the policy are present in table 13. Even conservative cost-benefit calculation shows that expanding Head Start is cost effective. If one considers only the static returns on health (the reduction on the health costs associated with childhood obesity at age 13), relaxing the eligibility thresholds implies that the marginal child by age 13 will be 10% less likely to be obese, which represents an expected saving of $36. By age 20 she will be 22% less likely to be under correctional supervision, and the expected discounted saving adds up to $1,030. Then, the program reaches the break-even point, even without accounting for its benefits on mental health, reduction in probability of presenting a chronic condition, and future days lost of work due to worse health (the present value of net savings add up to $650). These calculations are presented in Panel ”Alternative 2”.

When we account for the savings associated with (1) improvement in health through reducing the likelihood of obesity during adolescence and early adulthood and (2) for a broader measure of criminal expenses, which includes court and police protection costs, the present value of net savings associated to expanding eligibility to Head Start reaches $3225 (see Panel ”Alternative 1” in table 13). These calculations suggest that the internal rate of return of the program is at least 4%.

The current version of Head Start costs about three times more per child than the version available to the 1992 (see column for FY2003). Assuming that the pattern of benefits does not vary across the 1992 and the 2003 cohorts, the net benefits associated with Head Start through expanded eligibility add up to $2513 under the broader definition of benefits associated with reducing the prevalence of obesity and justice expenditures in Panel ”Alternative 1”. However, under the more strict definition included in Panel ”Alternative 2”, the combined benefits through improvement in mental health, reduction in the probability of presenting a chronic condition, and future days lost of work due to worse health associated with obesity, must reach $62 for the program to break-even.

7 Summary and Conclusions

In this paper we study the impact of Head Start (a preschool program for poor children) on the risky behaviors and health of children, adolescents and young adults. A recent experimental evaluation of this program, the HSIS, reports little or no effects on children outcomes. However, its focus is on short terms impacts, while our paper focuses on mid to long term impacts of the program.

Identification of the effects of the program is based on the fact that the probability of program participation is a discontinuous function of household income (and family size) because of the program's eligibility rules, enabling us to use a “fuzzy” regression discontinuity design. There is a large range of discontinuity cutoffs, which vary with family size, family structure, year and state. Therefore we are able to identify the effect of the program for a wide range of individuals.

We find that Head Start decreases behavioral problems, prevalence of chronic conditions and obesity at ages 12 to 13, depression and obesity at ages 16 and 17 and crime at ages 20-21. These effects are large, sustained and remarkably robust to a battery of tests. They show the potential for preschool programs to improve outcomes of poor children, even when they are universal programs such as Head Start.

We also present a simple cost-benefit calculation. Even our most conservative analysis shows that expanding Head Start is cost-effective.

References

[1] Administration of Children and Families, Department of Health and Human Services, 2007, `http://eclkc.ohs.acf.hhs.gov/`.

[2] Anderson, Kathryn, James Foster and David Frisvold, 2010. "Investing In Health: The Long-Term Impact Of Head Start On Smoking," Economic Inquiry, Western Economic Association International, vol. 48(3), pages 587-602, 07.

[3] Angrist, Joshua and Victor Lavy, 1999, "Using Maimonides' Rule To Estimate the Effect of Class Size of Scholastic Achievement", *The Quarterly Journal of Economics*, vol. 114, May 1999, pp. 533-576.

[4] Barnett, Steven, 2004. "Cost-Benefit Analysis of Preschool Education", PowerPoint presentation, `http://nieer.org/resources/files/BarnettBenefits.ppt`.

[5] Battistin, E. and Enrico Rettore, 2003, Another Look at the Regression Discontinuity Design, CEMMAP working paper CWP01/03.

[6] Belfield, C. R., M. Nores, W. S. Barnett, and L. Schweinhart, 2006. The High/Scope Perry Preschool program: Cost-benefit analysis using data from the age-40 followup. Journal of Human Resources 41 (1), 162-190.

[7] Benitez-Silva, Hugo, Moshe Buchinsky and John Rust, 2004, "How Large Are the Classification errors in the Social Security Disability Award Process", NBER Working Paper 10219, Feb. 2004.

[8] Berlinski, Samuel, Sebastian Galiani, and Paul Gertler. 2009. The Effect of Pre-Primary Education on Primary School Performance. Journal of Public Economics, 93(12): 21934.

[9] Berlinski, Samuel, Sebastian Galiani, and Marco Manacorda. 2008. Giving Children a Better Start: Preschool Attendance and School-Age Profiles. Journal of Public Economics, 92(56): 141640.

[10] Bitler, Marianne, Janet Currie and John Karl Scholz, 2003, "WIC Participation and Eligibility", *Journal of Human Resources*, v38, 2003, 1139-1179.

[11] Bitler, Marianne, Gelbach, Jonah and Hoynes, Hilary W., 2006, What mean impacts miss: Distributional effects of welfare reform experiments, *The American Economic Review* 96(4), 2006.

[12] Bitler, Marianne, Gelbach, Jonah and Hoynes, Hilary W.,2007, Can Subgroup-Specific Mean Treatment Effects Explain Heterogeneity in Welfare Reform?, June 2007, manuscript.

[13] Black, Dan A., Jose Galdo, and Jeffrey A.Smith, 2005, "Evaluating the Regression Discontinuity Design Using Experimental Data." Unpublished manuscript.

[14] Blank, Rebecca and Patricia Ruggles, 1996, ”When Do Women Use AFDC & Food Stamps? The Dynamics of Eligibility vs. Participation”, *The Journal of Human Resources,* 31(1), 1996, 57-89.

[15] Bowles, Samuel, Herbert Gintis, and Melissa Osborne, 2000, ”The Determinants of Individual Earnings: Skills, Preferences, and Schooling,” 2000. University of Massachusetts.

[16] Cameron, Heckman, Knudsen, Shonkoff, 2006, ”Economic, Neurobiological and Behavioral Perspectives on Building America’s Future Workforce”, NBER Working Paper, 12298, June 2006, National Bureau of Economic Research.

[17] Card, David, and Lara D. Shore-Sheppard, 2002, ”Using Discontinuous Eligibility Rules to Identify the Effects of the Federal Medicaid Expansions on Low Income Children”, NBER Working Paper, 9058, July 2002, National Bureau of Economic Research.

[18] Cascio, Elizabeth U. 2009. Do Investments in Universal Early Education Pay off? Long-term Effects of Introducing Kindergartens into Public Schools. National Bureau of Economic Research Working Paper 14951.

[19] Chay, Kenneth, Patrick McEwan and Miguel Urquiola, 2005, The Central Role of Noise in Evaluating Interventions That Use Test Scores to Rank Schools, *The American Economic Review* 95, 1237-1258.

[20] Cunha, Flavio, and J. Heckman, 2006, ”Formulating, Identifying and Estimating the Technology of Cognitive and Noncognitive Skill Formation”, Manuscript, University of Chicago.

[21] Currie, Janet, 2001, ”Early Childhood Intervention Programs: What Do We Know?,” *The Journal of Economic Perspectives*,15:2, Spring 2001, 213-238.

[22] Currie, Janet, 2006, ”The Take-up of Social Benefits,” in Alan Auerbach, David Card, and John Quigley (eds.) Poverty, the Distribution of Income, and Public Policy, (New York: Russell Sage) 2006.

[23] Currie, Janet, Eliana Garces and Duncan Thomas, 2002, ”Longer Term Effects of Head Start”, *The American Economic Review*, 92:4.

[24] Currie, Janet and Jonathan Gruber, 1996, ”Health insurance eligibility and Child Health: lessons from recent expansions of the Medicaid program,” *The Quarterly Journal of Economics*, May, 1996, 431-466.

[25] Cutler, David and Jonathan Gruber, 1996, ”Does Public Insurance Crowdout Private Insurance?”, *The Quarterly Journal of Economics*, CXI, May 1996, 391-430.

[26] Currie, Janet and Matthew Neidell, 2007, ”Getting Inside the ’Black Box’ of Head Start Quality: What Matters and What Doesn’t,” *Economics of Education Review*, Vol. 26, Issue 1, February 2007, pg. 83-99.

[27] Currie, Janet and Duncan Thomas, 1995, ”Does Head Start make a difference?”, *The American Economic Review*, 85(3), 341 – 364.

[28] Currie, Janet and Duncan Thomas, 1999, ”Does Head Start help Hispanic children?”, *The Journal of Public Economics*, 74 (1999), 235-262,

[29] Currie, Janet and Duncan Thomas, 2000, ”School Quality and the Longer-Term Effects of Head Start”, *The Journal of Human Resources*, 35(4), 755-774.

[30] Daly, Mary and Richard V. Burkhauser, 2002, The Supplemental Security Income Program, Federal Reserve Bank of San Francisco, Working Paper 2002-20.

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

[32] Department of Health and Human Services, 2006, Early Head Start Benefits Children and Families, available at `http://www.acf.hhs.gov/programs/opre/ehs/ehs_resrch/reports/dissemination/research_briefs/research_brief_overall.pdf`.

[33] Department of Health and Human Services, 2011, Income Eligibility for Enrollment in Head Start and Early Head Start Programs `http://www.acf.hhs.gov/programs/ohs/`.

[34] Department of Health and Human Services, 2010, Head Start Impact Study: Final Report, Executive Summary, January 2010.

[35] Economic Report of the President, 2006, United States Government Printing Office Washington, February 2006.

[36] Frisvold, D. (2011), Head Start Participation and Childhood Obesity, September 2011, mimeo.

[37] Frisvold, D. and J. Lumeng (2007), Expanding Exposure: Can Increasing the Daily Duration of Head Start Reduce Childhood Obesity?, November 2007.

[38] Fuerst, J.S. and D. Fuerst(1993). Chicago Experience with an Early Childhood Program: The Special Case of the Child Parent Center Program. Urban Education, 28(1):69-96.

[39] General Accounting Office, 1981, Head Start: An Effective Program but the Fund Distribution Formula Needs Revision and Management Controls Need Improvement, `http://www.gao.gov/assets/140/134239.pdf`

[40] General Accounting Office, 1989, Head Start: Information on Sponsoring Organizations and Center Facilities, `http://archive.gao.gov/d26t7/139159.pdf`

[41] General Accounting Office, 1998, Head Start Programs: Participant Characteristics, Services, and Funding `http://www.gao.gov/archive/1998/he98065.pdf`

[42] General Accounting Office, 2003, Better Data and Processes Needed to Monitor Underenrollment, `http://www.gao.gov/new.items/d0417.pdf`

[43] General Accounting Office, 2003b, Increased Percentage of Teachers Nationwide Have Required Degrees, but Better Information on Classroom Teachers Qualifications Needed, `http://www.gao.gov/new.items/d045.pdf`

[44] Gruber, Jonathan, 2003, ”Medicaid” in Means Tested Transfer Programs in the United States Robert Moffitt (ed) (Chicago: University of Chicago Press for NBER).

[45] Havnes, Tarjei, and Magne Mogstad. 2011. ”No Child Left Behind: Subsidized Child Care and Children’s Long-Run Outcomes.” American Economic Journal: Economic Policy, 3(2): 97129.

[46] Hahn, J., P. Todd and W. Van der Klaauw, 2001, Identification and Estimation of Treatment Effects with a Regression Discontinuity Design, *Econometrica* 69 (2001), 201-209.

[47] Head Start Office, 2011, Head Start Program Performance Standards and Other Regulations, `http://eclkc.ohs.acf.hhs.gov/hslc/Head%20Start%20Program/Program%20Design%20and%20Management/Head%20Start%20Requirements/Head%20Start%20Requirements`

[48] Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev and Adam Yavitz, 2010. ”The rate of return to the HighScope Perry Preschool Program”, *Journal of Public Economics*, Elsevier, vol. 94(1-2), pages 114-128, February.

[49] Heckman, J. J., J. Stixrud, and S. Urzua, 2006, The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior, *Journal of Labor Economics* 24(3), 411-482, July 2006.

[50] Hoynes, H., 1996, Welfare transfers in two-parents families: labor supply and welfare particcipation under AFDC-UP, *Econometrica*, vol.64(2), March 1996, 295-332.

[51] Imbens, Guido and Thomas Lemieux, 2007, Regression Discontinuity Designs: a guide to practice, NBER Technical Working Paper 337, April 2007.

[52] Imbens, Guido and J. D. Angrist, 1994, Identification and estimation of local average treatment effects. *Econometrica* 62(2), 467-475, March 1994.

[53] Keane, Michael and Robert Moffitt, 1998, A Structural Model of Labor Supply and Multiple Welfare Program Participation. *International Economic Review*, August 1998.

[54] Lee, David S., and Thomas Lemieux. 2010. Regression Discontinuity Designs in Economics. Journal of Economic Literature, 48(2): 281355.

[55] Lee, David and John DiNardo, 2004, “Economic Impacts of New Unionization on U.S. Private Sector Employers: 1984-2001”, *The Quarterly Journal of Economics*, 119(4), 1383-1442.

[56] LoSasso, Anthony. and Thomas C. Buchmueller, 2002. ”The Effect of the State Children’s Health Insurance Program on Health Insurance Coverage”, NBER Working Paper 9405, National Bureau of Economic Research, December 2002.

[57] Ludwig, J., and D. 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.

[58] Ludwig, J. and D. Phillips, 2007, The Benefits and Costs of Head Start, NBER wp 12973.

[59] Magnuson, Katherine A., Christopher Ruhm, and Jane Waldfogel. 2007. Does Prekindergarten Improve School Preparation and Performance?, Economics of Education Review, 26(1): 3351.

[60] McCrary, J., 2007, Testing for Manipulation of the Running Variable in the Regression Discontinuity Design, *Journal of Econometrics*, forthcoming.

[61] Moffitt, Robert A., 1983, ”An Economic Model of Welfare Stigma”, *The American Economic Review*, 73(5), 1023-1035.

[62] Moffitt, Robert A., 2002, The Temporary Assistance for Needy Families Program. In Means-Tested Transfer Programs in the U.S., NBER Working Paper 8749, National Bureau of Economic Research, January 2002.

[63] The National Institute for Early Education Research, 2005, The State of Preschool Yearbook, By W. Steven Barnett, Ph.D., Jason T. Hustedt, Ph.D., Kenneth B. Robin, Psy.D., and Karen L. Schulman, M.P.P., Rutgers, The State University of New Jersey.

[64] Neidell, Matthew and Jane Waldfogel, 2006, “Spillover effects of early education: evidence from Head Start”, January 2006, manuscript.

[65] National Collaborative on Childhood Obesity Research, 2009, `http://www.nccor.org/downloads/ChildhoodObesity_020509.pdf`. Accessed on Jan. 2012.

[66] Peterson, James L., and Nicholas Zill, 1986, “Marital Disruption, Parent-Child Relationships, and Behavioral Problems in Children.” Journal of Marriage and the Family 48,2 (May 1986).

[67] Romano, Joseph P. and Michael Wolf, 2005, ”Stepwise Multiple Testing as Formalized Data Snooping”, *Econometrica*, Vol. 73, No. 4, (Jul., 2005), pp. 1237-1282.

[68] Scholz, John Karl, 1994, ”The Earned Income Tax Credit: Participation, Effectiveness”, *National Tax Journal*, March, 1994, 59-81.

[69] Social Security Administration, 2011, Office of Policy Office of Research, Evaluation, and Statistics, Statistical Supplement to the Social Security Bulletin.

[70] U.S. Department of Health and Human Services, Administration for Children and Families, 2005. ”Head Start Impact Study: First Year Findings.” Washington, DC.

[71] U.S. Committee on Ways and Means. Green Book 1994, Washington DC, U.S. Government Printing Office, July 1994.

[72] Van Der Klaauw, Wilbert, 2002, Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression–discontinuity Approach, *International Economic Review* 43, 1249-1287.

[73] Westfall, P. H. and S. S. Young, 1993, Resampling-Based Multiple Testing: Examples and Methods for P-Value Adjustment, New York, Wiley.

[74] Zigler, E., 2010, Putting the National Head Start Impact Study into a Proper Perspective, National Head Start Association (U.S.).

[75] Zigler, E., and Valentine, J. (Eds.). Project Head Start: A Legacy of the War on Poverty. New York: The Free Press/Macmillan, 1979.

Tables

Table 1: Descriptive Statistics

	(1)	(2)	(3)
Variables	Obs	Mean	SD
Age 12-13			
BPI	1226	0.668	1.032
Overweight	1280	0.188	0.390
Special Education	1270	0.270	0.444
Grade Repetition	1310	0.336	0.472
Health: Use special equip.	1308	0.053	0.224
Any health limitation	1300	0.090	0.286
Age 16-17			
Overweight	1215	0.178	0.382
Drunk	1242	0.130	0.337
Ever convicted/arrested	1237	0.140	0.347
CESD	1066	-0.109	0.872
Age 20-21			
Overweight	960	0.140	0.347
Drunk	966	0.140	0.347
Ever convicted/arrested	964	0.257	0.437
High School Diploma	945	0.593	0.492
Treatment between ages 3-5			
Head Start	1676	0.305	0.461
Other preschool	1676	0.498	0.500
Other child care	1676	0.196	0.397
Mother's Characteristics			
AFQT	1615	-5.724	20.553
High School Dropout	1676	0.403	0.491
High School graduate	1676	0.496	0.500
College	1676	0.098	0.297
Age at child's birth	1671	22.998	4.516
Characteristics at entry (age 4)			
Total Family Income	1676	18348.040	9707.420
Father Figure present	1676	0.554	0.497
Family Size	1676	4.479	1.808
Poor	1676	0.578	0.494
Eligible	1676	0.594	0.491
Child's characteristics			
Birth weight (ounces)	1676	116.174	23.062
Low birth weight	1676	0.097	0.296
Breastfed	1660	0.349	0.477

Note: This table reports means and standard deviations for outcomes and control variables in our sample. Statistics are reported for males whose controls are all not missing and whose family income at age four is between 15% and 185% of the maximum level of income that would allow participation in Head Start. We report means and standard deviation using only one observation per individual.

Table 2: First Stage Estimates.

		(1)	(2)	(3)
Sample		All	Males	Females
1[HS Eligible at 4]	No controls	0.496***	0.668***	0.330***
		[0.0576]	[0.0819]	[0.0788]
Marginal Effect		0.169	0.224	0.114
1[HS Eligible at 4]	RD	0.277**	0.685***	-0.0484
		[0.118]	[0.169]	[0.170]
Marginal Effect		0.0884	0.209	-0.0149
Observations		2,560	1,295	1,255
Control Mean		0.244	0.216	0.272
SD		0.430	0.413	0.446

Note: The table reports results of probit regressions of Head Start participation on income eligibility. The marginal effect is the average marginal change in the probability of Head Start participation across individuals as the eligibility status changes and all other controls are kept constant. The first row of estimates does not include any controls, and the second row (RD) controls for: cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. The sample used in estimation includes only children ages 12-13. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Participation in Head Start and eligibility.

	(1)	(2)	(3)
Ages	3	4	5
		Panel A: Ages 12-13	
1[HS Eligible at 4]	0.130	0.685***	0.460***
	[0.177]	[0.169]	[0.171]
Marginal Effect	0.0394	0.209	0.148
		Panel B: Ages 16-17	
1[HS Eligible at 4]	0.0308	0.641***	0.535***
	[0.185]	[0.176]	[0.170]
Marginal Effect	0.00952	0.198	0.170
		Panel C: Ages 20-21	
1[HS Eligible at 4]	0.156	0.746***	0.471**
	[0.202]	[0.197]	[0.184]
Marginal Effect	0.0484	0.225	0.149

Note: Table of probit estimates of Head Start participation between ages 3-5 on income eligibility between ages 3-5. The marginal effect is the average marginal change in the probability of Head Start participation across individuals as the eligibility status changes and all other controls are kept constant. Controls excluded from table include: cubic in log family income and family size at age when eligibility is assessed, an interaction between these two variables, a dummy indicating the presence of a father figure in the household, race and age dummies, and dummies for year and state of residence. Robust standard errors are reported in brackets clustered at state-year at eligibility. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 4: Control Group - Alternative Child Care.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample		Ages 12-13			Ages 16-17			Ages 20-21	
Program	HS	Preschool	Informal	HS	Preschool	Informal	HS	Preschool	Informal
1[HS Eligible at 4]	0.685***	-0.386***	-0.335	0.641***	-0.210	-0.635***	0.746***	-0.323*	-0.678***
	[0.169]	[0.146]	[0.210]	[0.176]	[0.148]	[0.228]	[0.197]	[0.166]	[0.260]
Marginal Effect	0.209	-0.134	-0.0704	0.198	-0.0734	-0.134	0.225	-0.113	-0.128
Observations	1,295	1,295	1,295	1,229	1,229	1,229	954	954	954
Control Mean	0.219	0.585	0.203	0.225	0.572	0.207	0.191	0.649	0.178
SD	0.414	0.494	0.404	0.419	0.496	0.406	0.395	0.479	0.384

Note: The table reports results of probit regressions of different child care arrangements at ages 3-5 on income eligibility at age four. The marginal effect is the average marginal change in the probability of participation in an arrangement across individuals as the eligibility status changes and all other controls are kept constant. Controls excluded from table include: cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Falsification results: Pre-Head Start age outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Motor 0-2	Birth weight	Mom's Educ. 0-2	Grandm.'s Education	Mom married before age 3	Mom's AFQT	Family Income 0-2	Family Size 0-2	Mom lived in south at 14	Lived w/ parents at 14	Mom's siblings at 14	Mom lived in rural area at 14
							Panel A: All Males					
1[HS Eligible at 4]	-0.0970	-6.604**	0.105	-0.386	-0.0541	-3.469	-0.140*	-0.180	-0.0387	-0.0552	0.203	-0.0730
	[0.169]	[2.820]	[0.220]	[0.369]	[0.0344]	[2.486]	[0.0777]	[0.162]	[0.0356]	[0.0578]	[0.333]	[0.0483]
RW algorithm H0 rejected at 10%	No	No	No	No	No	No	No	No	No	No	No	No
Observations	584	1,310	1,310	1,202	1,310	1,269	1,310	1,310	1,255	1,299	1,298	1,296
Control Mean	0.0258	119.3	11.92	10.27	0.908	34.21	9.979	4.297	0.403	0.652	4.352	0.187
SD	0.872	23.72	1.997	3.188	0.290	23.98	0.635	1.614	0.492	0.478	2.874	0.391

Note: The table reports OLS estimates of family and child's outcomes measured before age three on income eligibility. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. The sample used includes children ages 12-13. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6: Reduced Form Estimates: Ages 12-13.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
								Health condition requires				
	Drug Use	Over-weight	Grade Retention	Alcohol Use	School Damage	Ever smoke	Special Education	sp. equip.	doctor	medicines	Health limitat.	BPI
1[HS Eligible at 4]	-0.130	-0.379**	-0.252	-0.251	0.0274	-0.158	-0.260	-0.736**	-0.322*	-0.218	-0.158	-0.279**
	[0.173]	[0.164]	[0.161]	[0.160]	[0.161]	[0.161]	[0.169]	[0.289]	[0.184]	[0.180]	[0.216]	[0.126]
RW algorithm												
H0 rejected at 10%	No	Yes	No	No	No	No	No	Yes	No	No	No	Yes
Observations	1,284	1,243	1,286	1,290	1,210	1,282	1,255	1,057	1,272	1,250	1,116	1,212
Control Mean	0.632	0.198	0.286	0.467	0.143	0.359	0.239	0.0921	0.190	0.207	0.0683	0.654
SD	0.484	0.399	0.453	0.500	0.351	0.481	0.428	0.290	0.393	0.406	0.253	0.908
Marginal Effect	-0.0405	-0.0953	-0.0820	-0.0859	0.00673	-0.0523	-0.0786	-0.0974	-0.0828	-0.0569	-0.0262	

Note: Probit (and OLS for BPI) estimates for several outcomes. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in the outcome as the eligibility status changes and all other controls are kept constant. "Control Mean" is the mean outcome among observations just above the cutoff, that is, at most 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7: Reduced Form Estimates: Ages 16-17.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
	In	Health	Birth	Overweight	Ever	Ever	Arrested or	Expectations by age 35				Mental Health		
	High School	Status	Control		drunk	Sex	Convicted	be married	with family	working	Eating	Effort	Sad	CESD
1[HS Eligible at 4]	0.0858	0.196	0.0965	-0.468**	-0.236	-0.0809	-0.109	0.0916	-0.491**	-0.0334	-0.341*	-0.342**	-0.364**	-0.336***
	[0.226]	[0.169]	[0.172]	[0.191]	[0.193]	[0.164]	[0.194]	[0.202]	[0.232]	[0.196]	[0.188]	[0.160]	[0.177]	[0.109]
RW algorithm														
H0 rejected at 10%	No	No	No	Yes	No	No	No	No	No	No	No	No	No	Yes
Observations	1,081	1,214	905	1,168	1,168	1,215	1,166	823	760	1,117	1,058	1,058	1,049	1,054
Control Mean	0.914	0.669	0.612	0.230	0.108	0.688	0.136	0.824	0.872	0.849	0.225	0.748	0.286	-0.0988
SD	0.281	0.472	0.489	0.422	0.311	0.465	0.344	0.383	0.336	0.359	0.419	0.435	0.453	0.837
Marginal Effect	0.0117	0.0662	0.0343	-0.117	-0.0467	-0.0246	-0.0217	0.0246	-0.115	-0.00726	-0.0974	-0.105	-0.114	

Note: Probit estimates for several outcomes (OLS estimates for CESD). Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in the outcome as the eligibility status changes and all other controls are kept constant. "Control Mean" is the mean outcome among observations at most 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 8: Reduced Form Estimates: Ages 20-21

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Health Status	Birth Control	Ever Sex	Overweight	Ever drunk	Arrested or Convicted	Ever in College	HS Diploma
1[HS Eligible at 4]	0.264	0.132	-0.143	-0.0644	-0.146	-0.421**	-0.134	0.0543
	[0.179]	[0.197]	[0.263]	[0.247]	[0.215]	[0.199]	[0.177]	[0.184]
RW algorithm								
H0 rejected at 10%	No	No	No	No	No	Yes	No	No
Observations	940	766	905	880	895	944	924	931
Control Mean	0.594	0.567	0.866	0.113	0.120	0.284	0.446	0.614
SD	0.493	0.498	0.342	0.318	0.326	0.452	0.499	0.489
Marginal Effect	0.0930	0.0473	-0.0258	-0.0136	-0.0293	-0.120	-0.0470	0.0189

Note: Probit estimates for several outcomes. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in the outcome as the eligibility status changes and all other controls are kept constant. ”Control Mean” is the mean outcome among observations just above the cutoff, that is, at most 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 9: Reduced Form Estimates: Other specifications.

	(1)	(2)	(3)	(4)	(5)
	Basic	No-Pre HS age controls	All controls	Quadratic	Quartic
			Panel A: Ages 12-13		
			A.1: BPI		
1[HS Eligible at 4]	-0.279**	-0.250**	-0.256*	-0.260**	-0.317**
	[0.126]	[0.125]	[0.145]	[0.124]	[0.127]
			A.2: Overweight		
1[HS Eligible at 4]	-0.379**	-0.380**	-0.491**	-0.360**	-0.395**
	[0.164]	[0.158]	[0.207]	[0.159]	[0.169]
Marginal Effect	-0.0953	-0.0970	-0.117	-0.0912	-0.0991
			A.3: Health condition that requires use of special equipment		
1[HS Eligible at 4]	-0.736**	-0.647**	-0.820**	-0.664**	-0.732**
	[0.289]	[0.281]	[0.332]	[0.261]	[0.288]
Marginal Effect	-0.0974	-0.0870	-0.111	-0.0886	-0.0969
			Panel B: Ages 16-17		
			B.1: Overweight		
1[HS Eligible at 4]	-0.468**	-0.493***	-0.318	-0.482**	-0.471**
	[0.191]	[0.187]	[0.227]	[0.190]	[0.194]
Marginal Effect	-0.117	-0.126	-0.0775	-0.121	-0.118
			B.2: CESD		
1[HS Eligible at 4]	-0.336***	-0.329***	-0.304**	-0.329***	-0.328***
	[0.109]	[0.108]	[0.126]	[0.107]	[0.112]
			Panel C: Ages 20-21		
			Arrested or Convicted		
1[HS Eligible at 4]	-0.421**	-0.315	-0.479**	-0.419**	-0.459**
	[0.199]	[0.199]	[0.237]	[0.198]	[0.207]
Marginal Effect	-0.120	-0.0940	-0.132	-0.120	-0.131

Note: Table includes reduced form estimates for selected outcomes using several specifications. Controls excluded from the table include: ”Basic” is specification used throughout the paper (cubic in log family income and family size at age 4, an interaction between these two variables, a dummy for the presence of a father figure in the child’s household at age 4, cubic in average log family income and average family size between ages 0 and 2, an interaction between the two, and cubic in birth weight, race and age dummies and dummies for year and state of residence at age 4). ”No pre-Head Start age controls” includes the same controls than in column (1), except those measured before age 3. ”All Controls” includes the same controls than in column (1) and dummies for highest grade completed by mother before child turned 3, maternal AFQT score, maternal grandmother’s highest grade completed and indicators of maternal situation at 14 years old (whether the mother lived in a Southern, whether she lived with parents and whether she lived in a rural area). ”Quadratic” and ”Quartic” are the same specification as ”Basic” but using polynomials up to the second and fourth order, respectively, in (log) income, family size and birth weight variables. Robust standard errors in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 10: Reduced Form Estimates: Trimming around cutoff.

	(1)	(2)	(3)	(4)	(5)	(6)
	[75%-125%]	[50%-150%]	[25%-175%]	[15%-185%]	300%	Full Sample
			Panel A: Ages 12-13			
			A.1: BPI			
1[HS Eligible at 4]	0.135	-0.170	-0.327**	-0.279**	-0.165*	-0.0663
	[0.203]	[0.153]	[0.134]	[0.126]	[0.0939]	[0.0821]
Observations	359	775	1,111	1,212	1,793	2,327
			A.2: Overweight			
1[HS Eligible at 4]	-1.008***	-0.472**	-0.399**	-0.379**	-0.00571	-0.0852
	[0.366]	[0.199]	[0.165]	[0.164]	[0.134]	[0.111]
Observations	300	773	1,129	1,243	1,852	2,392
			A.3: Health condition that requires use of special equipment			
1[HS Eligible at 4]	-0.353	-0.560	-0.738**	-0.736**	-0.270	-0.265
	[0.491]	[0.349]	[0.297]	[0.289]	[0.197]	[0.169]
Observations	236	557	941	1,057	1,776	2,299
			Panel B: Ages 16-17			
			B.1: Overweight			
1[HS Eligible at 4]	-0.751**	-0.573**	-0.443**	-0.468**	-0.376***	-0.361***
	[0.339]	[0.239]	[0.195]	[0.191]	[0.145]	[0.131]
Observations	303	725	1,077	1,168	1,755	2,248
			B.2: CESD			
1[HS Eligible at 4]	-0.453**	-0.255**	-0.308***	-0.336***	-0.178**	-0.0932
	[0.207]	[0.124]	[0.112]	[0.109]	[0.0826]	[0.0684]
Observations	315	672	974	1,055	1,534	1,960
			Panel C: Ages 20-21			
			Arrested/Convicted			
1[HS Eligible at 4]	-0.0583	-0.599**	-0.476**	-0.421**	-0.0638	-0.0617
	[0.395]	[0.242]	[0.212]	[0.199]	[0.145]	[0.127]
Observations	256	591	868	944	1,377	1,693

Note: Estimates using the same specification as table (6)-(8), but different trimming of data around the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 11: Income eligibility to Head Start at ages 0-7.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Eligible at age	0	1	2	3	4	5	6	7
				Panel A: Ages 12-13				
				A.1: BPI				
1[HS Eligible at 4]	0.00948	-0.0750	0.129	0.168	-0.279**	-0.232*	0.0564	-0.0971
	[0.141]	[0.145]	[0.138]	[0.136]	[0.126]	[0.137]	[0.123]	[0.141]
Observations	984	1,089	1,173	1,110	1,212	1,013	1,093	984
				A.2: Overweight				
1[HS Eligible at 4]	0.110	-0.157	-0.0151	0.206	-0.379**	-0.495**	-0.199	0.117
	[0.244]	[0.232]	[0.211]	[0.225]	[0.164]	[0.228]	[0.218]	[0.253]
Observations	713	814	928	917	1,243	870	876	761
				A.3: Health condition that requires use of special equipment				
1[HS Eligible at 4]	0.532	0.820**	-0.355	0.0377	-0.736**	-0.0713	0.424	-0.575*
	[0.417]	[0.384]	[0.259]	[0.353]	[0.289]	[0.316]	[0.269]	[0.311]
Observations	526	647	841	768	1,057	707	793	630
				Panel B: Ages 16-17				
				B.1: Overweight				
1[HS Eligible at 4]	-0.259	-0.200	-0.127	0.00716	-0.468**	-0.531***	0.0368	-0.181
	[0.276]	[0.235]	[0.223]	[0.229]	[0.191]	[0.199]	[0.238]	[0.256]
Observations	678	791	862	895	1,168	856	829	721
				B.2: CESD				
1[HS Eligible at 4]	0.0163	0.100	-0.00387	-0.0406	-0.336***	0.0301	-0.101	-0.0493
	[0.120]	[0.111]	[0.133]	[0.120]	[0.109]	[0.108]	[0.124]	[0.130]
Observations	831	975	1,010	990	1,054	924	939	841
				Panel C: Ages 20-21				
				Arrested/Convicted				
1[HS Eligible at 4]	0.116	0.407	0.516**	-0.0358	-0.421**	-0.207	0.210	-0.0378
	[0.271]	[0.249]	[0.238]	[0.232]	[0.199]	[0.226]	[0.212]	[0.263]
Observations	592	655	711	760	944	775	741	667

Note: Table includes reduced form results of table (6)-(8) with income eligibility measured at different ages between 0 and 7. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 12: Estimates of structural equations.

	(1)	(2)	(3)	(4)	(5)	(6)
	Ages 12-13			Ages 16-17		Ages 20-21
	Overweight	Special Equipment	BPI	Overweight	CESD	Arrested or Convicted
Head Start	-1.281***	-1.753***	-0.598	0.0230	-0.552	-1.161***
	[0.327]	[0.102]	[0.524]	[1.523]	[0.499]	[0.304]
Marginal Effect	-0.292	-0.243		0.00548		-0.306

Note: Participation in program is instrumented with eligibility status at age four. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in outcome across individuals as the participation in Head Start between ages 3 and 5 changes and all other controls are kept constant. Estimates obtained by bivariate probit allow for a tolerance of 0.0001 in the likelihood using the Newton-Raphson algorithm. Standard errors are obtained using the observed information matrix. For the continuous outcomes, BPI and CESD, the standard errors for the 2SLS estimates they are obtained by block-bootstrap (500 replications; the block unit is state-year of residence when child was 4). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 13: Cost-benefit analysis.

	FY 1992	FY 2003
COSTS		
Total appropriation (HS Program Fact Sheet)	1,439,903,924	5,718,484,287
Nb children	621,078	909,608
Cost per child	2,318	6,287
(Ct) Annual cost of attendance at age 4 (discounted for age 0)	1981.77	5373.95
BENEFITS - Alternative 1		
1) CRIME		
Total cost of justice system in 2007	219,930,610,736	
Adults under correctional supervision 2007	7,267,500	
Cost/correction	30,262	
(C1) Cost at age 20 (discounted for age 0)	14,364	
2) OBESITY		
Per Capita Cost	609	
(O1) Direct costs of obesity between ages 13-30 (discounted for age 0)	4,815	
NET SAVINGS		
(A1) 22% less likely to be arrested/sentenced at age 20-21 ($0.22 \times C1$)	3160.02	
(A2) 10% less likely to be obese at 13-30 ($0.10 \times O1$)	481.53	
(A3) 21% more likely to participate in HS ($0.21 \times Ct$)	416.17	1128.53
Total savings (A1 + A2 - A3)	3225.38	2513.03
BENEFITS - Alternative 2		
1) CRIME		
Total cost of corrections in 2007	71,709,452,250	
Adults under correctional supervision 2007	7,267,500	
Cost/correction	9867.14	
(C2) Cost at age 20 (discounted for age 0)	4683.36	
2) OBESITY		
Per Capita Cost	609	
(O2) Cost at age 13 (discounted for age 0)	366	
NET SAVINGS		
(B1) 22% less likely to be arrested/sentenced at age 20-21 ($0.22 \times C2$)	1030.34	
(B2) 10% less likely to be obese at 12-13 ($0.10 \times O2$)	36.57	
(B3) 21% more likely to participate in HS ($0.21 \times Ct$)	416.17	1128.53
Total savings (B1 + B2 - B3)	650.74	-61.61

Note: All monetary values are in 2009 dollars. Values discounted to age 0 are calculated using an interest rate of 4%. Costs related to criminal activity are obtained from the Justice Expenditure and Employment Extracts for 2007. The figures for "total justice system" include the costs with police protection, judicial and legal expenses and corrections. The number of adults under correctional supervision is obtained from U.S. Department of Justice, 2010.

Figures

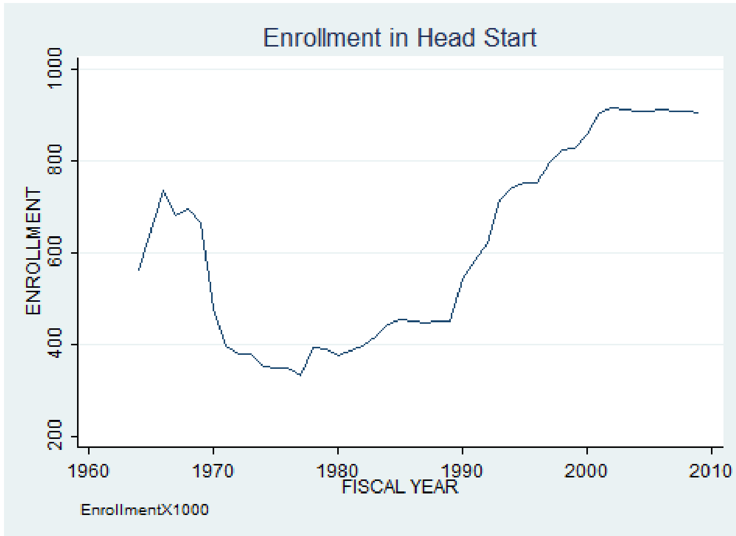

Figure 1: Historical enrolment in Head Start.

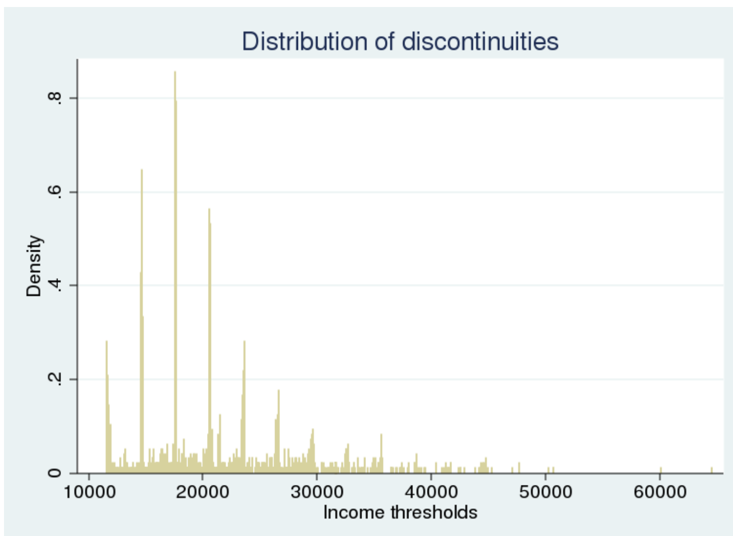

Figure 2: Distribution of Income thresholds at age 4.
Note: Support of income at age 4 for sample used in estimation. Includes all children used in the regressions whose family income at age 4 was 15-185% of the discontinuity level of income.

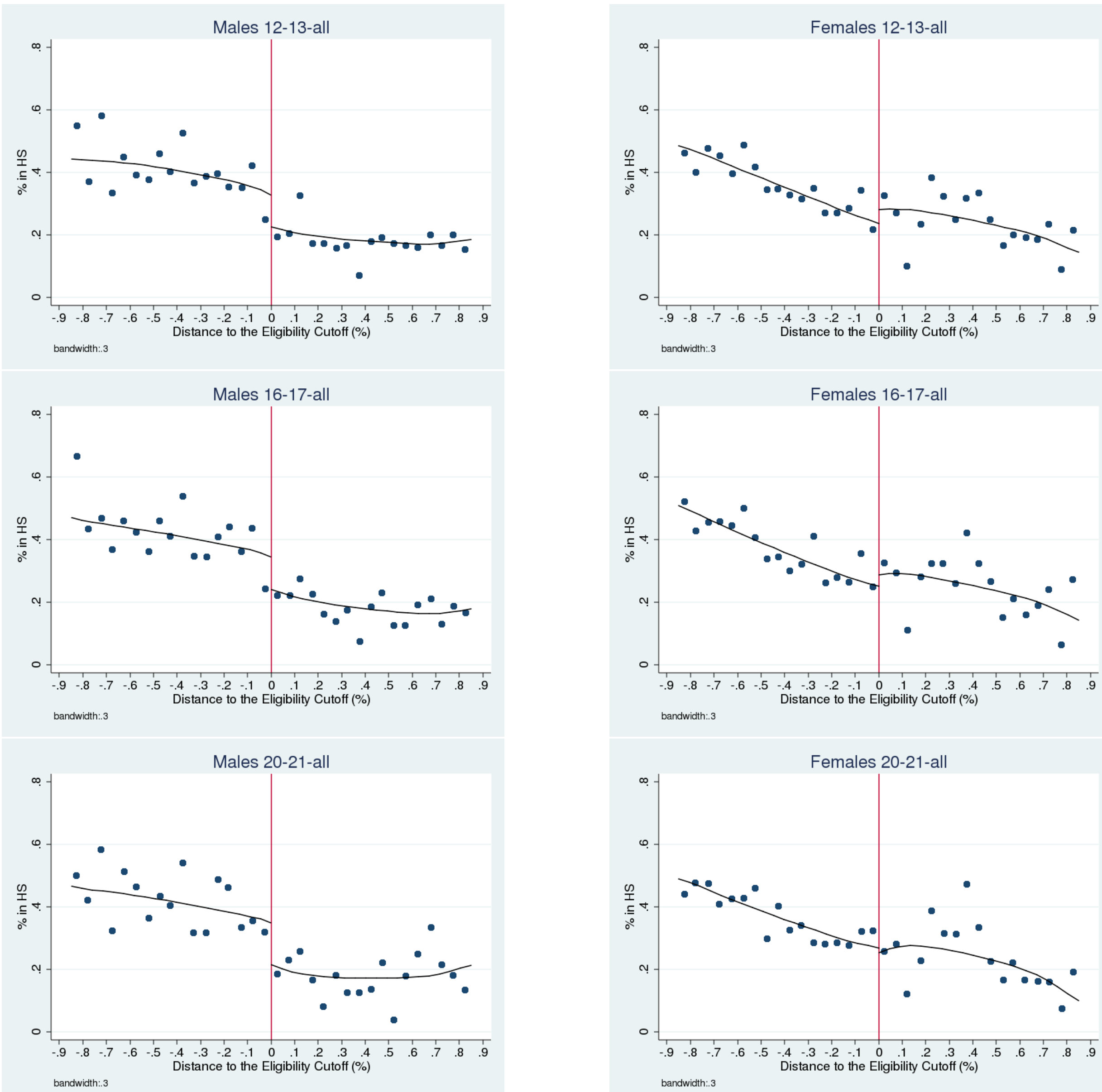

Figure 3: Proportion of children in Head Start, by eligibility status.

Note: The continuous lines in Figure are local linear regression estimates of Head Start participation on percentage distance to cutoff; regressions were run separately on both sides of the cutoff and the bandwidth was set to 0.3. Circles in figures represent mean Head Start participation by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

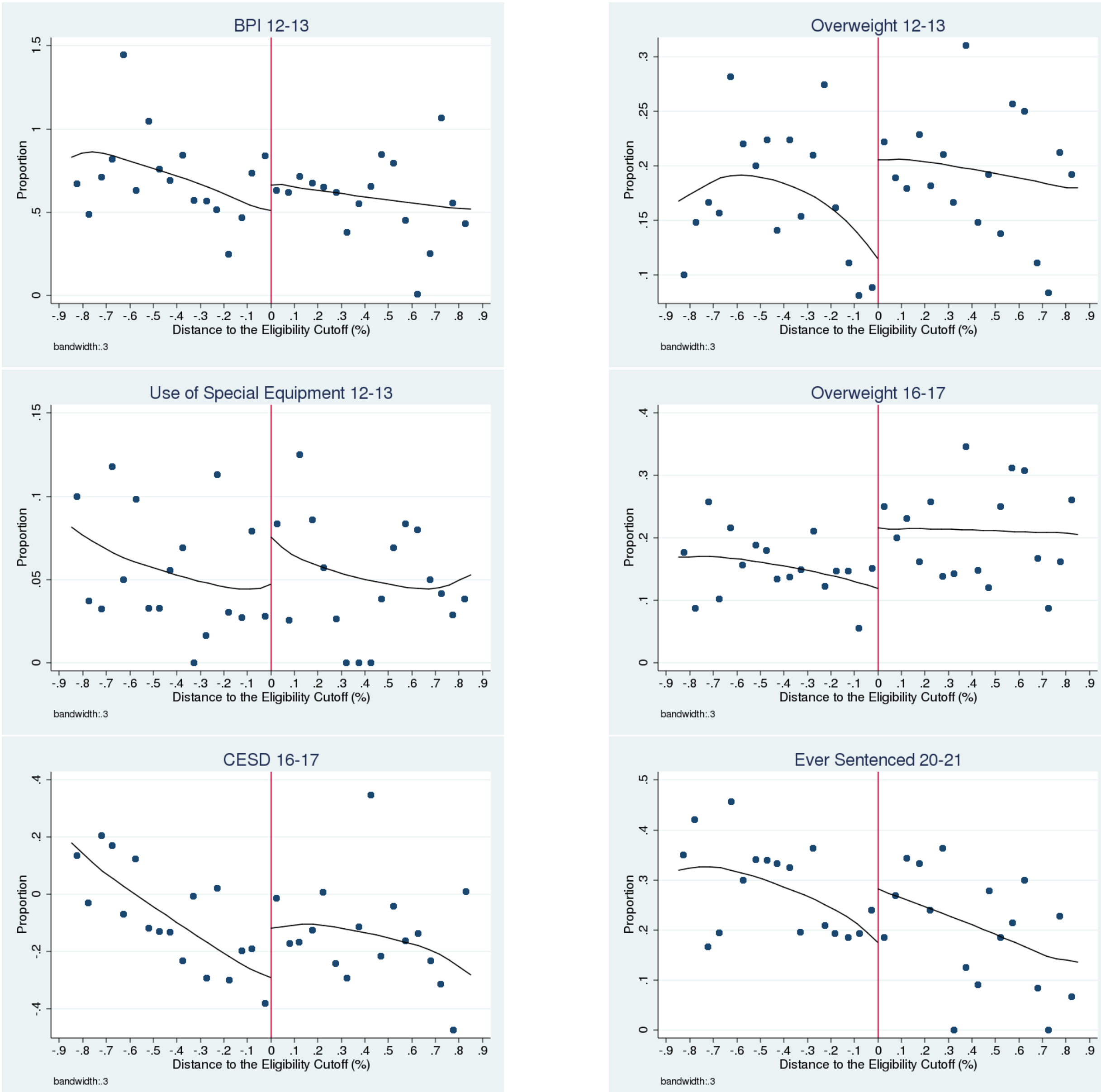

Figure 4: Average outcomes by eligibility status, Bandwidth = 0.3.

Note: The continuous lines in Figure present local linear regression estimates of several outcomes on percentage distance to cutoff. The bandwidth was set to 0.3. Circles in figures represent the mean outcome by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

Appendix for: Long Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start

A Tables

Table A.1: Summary of legislation

Date	Law Number	Title	Description
1964	88-452	Economic Opportunity Act	Anti-poverty bill to ”strengthen, supplement, and coordinate efforts in furtherance” of a policy of the U.S. ”to eliminate the paradox of poverty in the midst of plenty”. Head Start was not mention in the original act, but it was considered part of the Community Action Program.
1966	P.L. 89-794	Economic Opportunity Act Amendments of 1966	A section was added to Title II making Head Start a part of the Economic Opportunity Act.
1967	P.L. 90-222	Economic Opportunity Act Amendments of 1967	”Follow Through” was added in Title II, to continue services for Head Start children when they enter kindergarten and elementary school. This program was administered by the Office of Education.
1969	P.L. 91-177	Economic Opportunity Act Amendment of 1969	A provision was added allowing children from families above the poverty level to receive Head Start services for a fee.
1972	P.L. 92-424	Economic Opportunity Act Amendment of 1972	A fee schedule for non-poor participants in Head Start was required; fees were prohibited for families below the poverty line. Added a requirement that at least 10 percent of Head Start's enrollment include children with disabilities.
1973	93-202	Postponement of a Head Start Fee Schedule	Prior approval by Congress was required before any Head Start fee schedule could be established.
1974	P.L. 93-644	Head Start, Economic Opportunity, and Community Partnership Act of 1974	Reauthorized Head Start through the fiscal year of 1978. Head Start funds should be allocated to states proportionately based upon each state's relative numbre of children living in families with income below the poverty line and the relative number of public assistance recipients in each state.
1978	P.L. 95-568	Economic Opportunity Act Amendment of 1978	Reauthorized Head Start for three more years. Minor changes to the law.
1981	P.L. 97-35 (42 USC 9831 et. Seq.)	Economic Opportunity Amendment of 1981	The Head Start Act was attached to the OBRA of 1981. To ”promote school readiness by enhancing the social and cognitive development of low-income children.”
1984	P.L. 98-558	Human Services Reauthorization Act of 1984	Head Start Reauthorization for 2 years. In 1984, the Indian and Migrant branches of Head Start became separate regions; prohibited changes in methods for determining eligibility for low income if they would reduced participation fo persons in the program. HS may provide services to children age 3 to the age of compulsory school attendance.
1989	P.L. 101-120	Head Start Supplemental Authorization Act of 1989	Reauthorized Head Start for FY of 1990.
1990	P.L. 101-597	National Health Service Corps Revitalization Amendments of 1990	Minor amend to Head Start Act (library of congress 101th congress pub laws)
1990	P.L. 101-501	Head Start Reauthorization Act of 1990.	Reinforced importance of parental involvement, improved information on Head Start programs.
1992	P.L. 102-763	Head Start Improvement Act	Facilities purchase; Extended waivers for non-federal regulations; Establishment of transportation regulations; Health services to younger siblings; Protection of the quality set-aside; Literacy and child development training to parents; Elimination of priority status to a grantee once funded.
1994	P.L. 103-218	Head Start Act Amendments of 1994	Reauthorized Head Start for the years of 1995 through 1998. Required the development of quality standards (including revising the Program Performance Standards), the development of performance measures, and improved monitoring of local programs. It authorized family-centered programs for infants and toddlers. It established new standards for classroom teachers and family service workers.
1998	P.L. 105-285	Coats Human Services Reauthorization Act of 1998	Reauthorized Head Start for 5 years.
2007	110-134	Improving Head Start for School Readiness Act of 2007	Allows grantees to serve additional children from families with income up to 130% of poverty to be served; formula allocation remains at 100% of poverty; expansion of both Head Start and Early Head Start programs with additional funds going to states serving fewer than 60 percent of eligible children; establishes standards for the curriculum of teachers.

Note: Source of regulations relevant to Head Start: 45 CFR (Code of Federal Regulations), Parts 1301 to 1311. Additional Program Instructions and Information Memorandums can be found at the Early Childhood Learning and Knowledge Center web site: http://eclkc.ohs.acf.hhs.gov/hslc.

Table A.2: Number of children and average cutoff per year/family size.

	Family Size at age 4													
Year	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1981	8	17	17	11	5	2	3	2	1					
	12369.04	15561.19	18606.65	21264.72	23426.72	26310.65	29194.59	33631.24	34962.45					
1982	13	21	33	18	1	9	3	3	1	1		1		
	13232.16	15094.13	17878.9	20693.89	23452.54	26645.22	29287.26	33120.24	35121.98	38039.34		43874.06		
1983	15	30	44	24	11	9	3	2	5		1	1		1
	12397.69	14986.1	17790.46	21112.79	23661.89	26716.88	29657.66	33383.6	35653.43		41649.2	44647.09		50642.86
1984	19	43	39	35	10	10	4	2	2	3	2		1	
	12550.84	15299.46	18026.74	21177.34	23651.57	26659.88	29712.28	32676.51	35684.82	38693.13	41701.45		47718.07	
1985	24	40	56	35	13	6	3	5	3	3	1	2		
	15002.53	17192.84	18632.14	22652.98	25053.74	28349.69	33580.82	35235.4	39347.12	41727.43	41516.94	44500.2		
1986	13	36	53	35	11	4	3	2		1	1	1		
	13229.18	16769.68	18934.03	22625.68	23621.49	26630.19	32442.84	32647.58		38664.98	42537.88	44682.38		
1987	10	33	31	30	10	3	5	1	2		1			
	15920.71	17670.95	18704.52	22533.42	24368.81	26552.74	29537.96	32523.17	35508.39		41478.83			
1988	6	28	40	22	16	6	4	2						
	12808.94	17625.27	19682.56	21725.57	25513.44	28083.51	29543.82	35238.85						
1989	12	31	37	42	9	5	4							
	15646.89	17027.51	19778.27	22106.93	25003.76	26521.42	31449.06							
1990	9	21	27	16	12	4	3	1				2		
	12531.59	17587.47	18685.67	21704.04	26430.83	26552.13	29523.97	32495.81				44383.16		
1991	1	18	38	21	13	6		1	1		1			
	12774.68	16827.94	19072.06	22441.04	23609.98	26587.58		32542.77	35520.37		41475.56			
1992	9	18	33	19	9	3		4	1	1	1			
	15201.42	18220.12	19792.78	21698.49	26467.22	26664.45		33907.66	35691.68	38700.75	50719.35			
1993	5	16	38	31	19	4	6	2						
	13886.37	16404.64	19549.22	21567.42	24049.47	27095.42	29986.9	32709.41						
1994	3	15	23	21	11	7	1	1	1					
	15732.29	18308.64	19071	21929.75	27597.85	26534.29	29458.71	39368.85	35369.52					
1996	2	12	12	13	8	5	3	1	2					
	12767.88	17676.26	19227.76	21895.09	25054.25	26507.95	29468.35	32428.74	39917.86					
1998	1	9	7	6	9	7								
	11640.93	15956.91	20919.16	20653.27	23668.11	26661.49								
2000	3	3	8	6	2	2								
	12869.11	16714.29	17932.83	22722.86	23618.07	26615.55								
2002	2	4	4	4	5			1						
	11609.64	18820.51	17599.21	20593.99	23588.78			32573.12						
2004	1	2	2		3									
	14150.35	16255.96	17641.14		24801.17									

Note: Average number of children (top number in each cell) and value of cutoff (bottom value) per year-family size.

Table A.3: Description of variables.

Variables		Ages		
PIAT-Math	Test that measures a child's attainment in mathematics as taught in mainstream education. Standardized score with population mean 0 and standard deviation 1 (normed within age).	6-12		
PIAT-RR	Test that measures word recognition and pronunciation ability. Standard score with population mean 0 and standard deviation 1 (normed within age).	6-13		
PIAT-RC	Measures child's ability to derive meaning from sentences. Standard score with population mean 0 and standard deviation 1 (normed within age).	6-14		
PPVT	Measures an individuals receptive (hearing) vocabulary for Standard American English. Standard score with population mean 0 and standard deviation 1 (normed within age).	10-11		
Overweight	Indicator that takes value 1 if the individual's Body Mass Index (BMI) is above the 95th percentile for her/his age and gender.	12-13	16-17	20-21
CESD	Center for Epidemiological Studies Depression Scale: percentile score that measures symptoms of depression (higher scores are negative)		16-17	
Drug Use	Ever tried cocaine or marijuana	12-13	16-17	
Alcohol Use	Ever tried alcohol	12-13		
Ever Drunk	Ever got drunk		16-17	20-21
Ever smoked	Ever smoked	12-13		
Ever sex	Ever had sexual intercourse		16-17	20-21
Grade Repetition	Ever repeated a grade up to a given age	12-13		
Special Education	Attending classes for remedial work	12-13		
School damage	Ever damaged school property	12-13		
High School	Attending high school		16-17	
Ever in college	ever in college			20-21
High School Diploma	1 if has HS diploma or was ever in college as opposed to be GED/not HS diploma			20-21
Health				
Any limitation	Child has health problem that limits school attention, work or play activity	12-13		
Health Status	Self-reported health status (1 if has good/excellent health status)		16-17	20-21
Health: equipment	Child had between ages 6 and 12-13 health condition requires use of special equipment, such as a brace, crutches, a wheelchair, special shoes, a helmet, a special bed, a breathing mask, an air filter, or a catheter and so on	12-13		
Health: doctor	Child had between ages 6 and 12-13 health condition that requires medical attention	12-13		
Health: medicines	Child had between ages 6 and 12-13 health condition that requires regular use of medicines	12-13		
Sad	Respond rarely feels sad (at most 2 days/week)		16-17	
Effort	Respond rarely feels (at most 2 days/week) that everything is an effort		16-17	
Eating	Respond feels poor appetite at most 2 days/week		16-17	
Illicit activities				
Ever sentenced	Indicator for whether the individual has ever been convicted of any charge other than minor traffic violations or sentenced to a corrections institution/jail/reform school.		16-17	20-21

Table A.4: First Stage Estimates.

		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Sample		All	Males	Females	Black Males	Black Females	Non-Black Males	Non-Black Females
					Panel A: ages 12-13			
1[HS Eligible at 4]	no controls	0.496***	0.668***	0.330***	0.359***	0.210*	0.782***	0.271**
		[0.0576]	[0.0819]	[0.0788]	[0.120]	[0.119]	[0.114]	[0.107]
Marginal Effect		0.169	0.224	0.114	0.137	0.0809	0.228	0.0802
1[HS Eligible at 4]	RD	0.277**	0.685***	-0.0484	0.874***	0.172	0.474**	-0.293
		[0.118]	[0.169]	[0.170]	[0.259]	[0.260]	[0.234]	[0.238]
Marginal Effect		0.0884	0.209	-0.0149	0.273	0.0582	0.126	-0.0770
Observations		2,560	1,295	1,255	564	566	708	654
Control Mean		0.244	0.216	0.272	0.359	0.348	0.140	0.225
SD		0.430	0.413	0.446	0.484	0.480	0.349	0.420
					Panel B: ages 16-17			
1[HS Eligible at 4]	no controls	0.510***	0.687***	0.337***	0.384***	0.180	0.817***	0.351***
		[0.0591]	[0.0836]	[0.0810]	[0.126]	[0.123]	[0.115]	[0.111]
Marginal Effect		0.175	0.231	0.118	0.146	0.0693	0.242	0.108
1[HS Eligible at 4]	RD	0.307**	0.641***	0.0142	0.811***	0.115	0.450*	-0.127
		[0.122]	[0.176]	[0.176]	[0.274]	[0.268]	[0.243]	[0.254]
Marginal Effect		0.0993	0.198	0.00449	0.254	0.0398	0.124	-0.0349
Observations		2,423	1,229	1,187	536	550	664	598
Control Mean		0.249	0.224	0.275	0.361	0.348	0.150	0.223
SD		0.433	0.418	0.448	0.484	0.480	0.359	0.419
					Panel C: ages 20-21			
1[HS Eligible at 4]	no controls	0.538***	0.722***	0.376***	0.475***	0.247*	0.796***	0.337***
		[0.0656]	[0.0933]	[0.0905]	[0.137]	[0.135]	[0.129]	[0.127]
Marginal Effect		0.183	0.241	0.130	0.177	0.0945	0.240	0.101
1[HS Eligible at 4]	RD	0.304**	0.746***	-0.00240	1.063***	0.229	0.496*	-0.183
		[0.136]	[0.197]	[0.186]	[0.328]	[0.296]	[0.261]	[0.281]
Marginal Effect		0.0978	0.225	-0.000755	0.314	0.0763	0.134	-0.0487
Observations		1,986	954	1,023	414	484	518	509
Control Mean		0.225	0.190	0.261	0.300	0.346	0.144	0.209
SD		0.419	0.394	0.441	0.464	0.480	0.353	0.409

Note: The table reports results of probit regressions of Head Start participation on income eligibility. The marginal effect is the average marginal change in the probability of Head Start participation across individuals as the eligibility status changes and all other controls are kept constant. The first row of estimates in each panel does not include any controls, and the second row (RD) controls for: cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.5: Control Group: Alternative Child Care by child's race.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample		Black			Non-Black	
Program	HS	Preschool	Informal	HS	Preschool	Informal
			Panel A: ages 12-13			
1[HS Eligible at 4]	0.874***	-0.661***	-0.399	0.474**	-0.160	-0.306
	[0.259]	[0.244]	[0.356]	[0.234]	[0.194]	[0.266]
Marginal Effect	0.273	-0.220	-0.0706	0.126	-0.0538	-0.0711
Observations	564	564	564	708	708	708
Control Mean	0.339	0.500	0.211	0.147	0.647	0.212
SD	0.477	0.504	0.411	0.355	0.480	0.411
			Panel B: ages 16-17			
1[HS Eligible at 4]	0.811***	-0.369	-1.179***	0.450*	-0.0784	-0.493*
	[0.274]	[0.249]	[0.387]	[0.243]	[0.195]	[0.299]
Marginal Effect	0.254	-0.125	-0.232	0.124	-0.0265	-0.112
Observations	536	536	536	664	664	664
Control Mean	0.339	0.491	0.240	0.159	0.636	0.212
SD	0.477	0.505	0.431	0.367	0.484	0.410
			Panel C: ages 20-21			
1[HS Eligible at 4]	1.063***	-0.714**	-1.077**	0.496*	-0.0248	-0.605*
	[0.328]	[0.293]	[0.534]	[0.261]	[0.215]	[0.342]
Marginal Effect	0.314	-0.243	-0.217	0.134	-0.00801	-0.122
Observations	414	414	414	518	518	518
Control Mean	0.263	0.622	0.208	0.151	0.681	0.195
SD	0.446	0.492	0.415	0.360	0.469	0.399

Note: The table reports results of probit regressions of different child care arrangements at ages 3-5 on income eligibility at age four. The marginal effect is the average marginal change in the probability of participation in an arrangement across individuals as the eligibility status changes and all other controls are kept constant. Controls excluded from table include: cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.6: Falsification results: Pre-Head Start age outcomes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Motor 0-2	Birth weight	Mom's Educ. 0-2	Grandm.'s Education	Mom married before age 3	Mom's AFQT	Family Income 0-2	Family Size 0-2	Mom lived in south at 14	Lived w/ parents at 14	Mom's siblings at 14	Mom lived in rural area at 14
							Panel A: Ages 16-17					
1[HS Eligible at 4]	0.0103	-5.396*	0.157	-0.385	-0.0296	-3.031	-0.121	-0.0884	-0.0468	-0.0207	0.481	-0.0278
	[0.174]	[2.769]	[0.216]	[0.402]	[0.0360]	[2.610]	[0.0778]	[0.168]	[0.0358]	[0.0598]	[0.377]	[0.0498]
Observations	562	1,242	1,242	1,123	1,242	1,178	1,242	1,242	1,156	1,206	1,206	1,205
Control Mean	-0.0199	119.4	11.87	10.38	0.891	34.31	9.955	4.262	0.401	0.626	4.278	0.165
SD	0.916	22.66	1.966	3.072	0.313	24.20	0.631	1.630	0.492	0.485	2.928	0.372
							Panel B: Ages 20-21					
1[HS Eligible at 4]	-0.283	-4.687	0.392	0.0686	-0.0248	-3.092	-0.128	-0.219	-0.0262	-0.0699	-0.0351	-0.0164
	[0.215]	[3.106]	[0.262]	[0.444]	[0.0373]	[2.978]	[0.0875]	[0.197]	[0.0430]	[0.0690]	[0.452]	[0.0636]
Observations	367	966	966	852	966	894	966	966	880	918	918	919
Control Mean	-0.00395	117.4	11.64	10.21	0.905	34.80	9.961	4.249	0.403	0.669	4.453	0.185
SD	0.957	24.40	2.045	3.434	0.294	23.97	0.642	1.644	0.493	0.472	3.224	0.389

Note: The table reports OLS estimates of family and child's outcomes measured before age three on income eligibility. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.7: Reduced Form Estimates: Test Scores.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Males			Black Males			non Black Males		
	PIAT-M	PIAT-RC	PIAT-R	PIAT-M	PIAT-RC	PIAT-R	PIAT-M	PIAT-RC	PIAT-R
Panel A: Ages 6-7									
1[HS Eligible at 4]	-0.0185	0.107	0.0428	0.0127	-0.123	0.0396	-0.153	0.201	0.00275
	[0.0895]	[0.105]	[0.0830]	[0.168]	[0.194]	[0.127]	[0.124]	[0.149]	[0.119]
RW algorithm									
H0 rejected at 10%	No	No	No	No	No	No	No	No	No
Observations	1,371	706	1,343	538	269	523	772	354	761
Control Mean	-0.0901	0.265	0.0599	-0.184	0.453	0.250	-0.0314	0.105	-0.0701
SD	0.794	0.668	0.789	0.716	0.729	0.861	0.820	0.596	0.709
Panel B: Ages 9-10									
1[HS Eligible at 4]	-0.0330	-0.00748	-0.0326	0.0571	0.0501	0.139	-0.158	-0.0138	-0.0962
	[0.0983]	[0.120]	[0.117]	[0.173]	[0.188]	[0.186]	[0.132]	[0.172]	[0.172]
RW algorithm									
H0 rejected at 10%	No	No	No	No	No	No	No	No	No
Observations	1,364	1,324	1,358	563	549	564	775	749	768
Control Mean	0.0291	-0.0294	0.114	-0.0979	-0.00860	0.0913	0.0708	-0.0513	0.0979
SD	0.911	0.972	1.043	0.982	0.940	0.971	0.864	0.999	1.072
Panel C: Ages 12-13									
1[HS Eligible at 4]	0.0340	-0.135	-0.227*	0.0216	-0.165	-0.153	0.0311	-0.0873	-0.197
	[0.101]	[0.113]	[0.133]	[0.181]	[0.170]	[0.195]	[0.132]	[0.166]	[0.210]
RW algorithm									
H0 rejected at 10%	No	No	No	No	No	No	No	No	No
Observations	1,198	1,182	1,197	528	521	528	650	636	644
Control Mean	0.0468	-0.161	0.156	-0.0384	-0.335	0.00452	0.0891	-0.0843	0.227
SD	0.916	0.915	1.085	0.907	0.784	0.976	0.915	0.944	1.116

Note: OLS estimates for test scores. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birthweight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. The sample used includes children ages 12-13. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.8: Participation in Food Stamps, Medicaid, AFDC/TANF and SSI at ages 3-5 and eligibility to Head Start.

	(1)	(2)	(3)	(4)
	Medicaid	Food Stamps	AFDC/TANF	SSI
		Eligibility measured at age 0		
		Ages 12-13		
1[Eligible at 4]	0.520**	0.244	0.398**	0.592**
	[0.217]	[0.170]	[0.197]	[0.244]
Marginal Effect	0.155	0.0806	0.119	0.117
control mean	0.214	0.407	0.250	0.105
		Ages 16-17		
1[Eligible at 4]	0.391**	0.297*	0.388*	0.489**
	[0.199]	[0.174]	[0.200]	[0.232]
Marginal Effect	0.118	0.100	0.116	0.0982
control mean	0.244	0.423	0.245	0.113
		Eligibility measured at age 2		
		Ages 12-13		
1[Eligible at 4]	0.221	0.535***	0.387**	0.194
	[0.231]	[0.161]	[0.177]	[0.190]
Marginal Effect	0.0623	0.176	0.108	0.0412
control mean	0.271	0.437	0.250	0.128
		Ages 16-17		
1[Eligible at 4]	0.0520	0.447***	0.271	0.203
	[0.236]	[0.172]	[0.184]	[0.201]
Marginal Effect	0.0148	0.146	0.0757	0.0437
control mean	0.301	0.463	0.259	0.132
		Eligibility measured at age 4		
		Ages 12-13		
1[Eligible at 4]	0.644***	0.490***	0.750***	0.481**
	[0.214]	[0.159]	[0.168]	[0.244]
Marginal Effect	0.183	0.157	0.209	0.0998
control mean	0.229	0.427	0.236	0.101
		Ages 16-17		
1[Eligible at 4]	0.505**	0.466***	0.707***	0.440*
	[0.222]	[0.158]	[0.168]	[0.245]
Marginal Effect	0.148	0.152	0.198	0.0917
control mean	0.256	0.466	0.250	0.110
		Eligibility measured at age 5		
		Ages 12-13		
Food Stamps	0.398	0.393**	0.389**	0.172
	[0.250]	[0.178]	[0.183]	[0.216]
Marginal Effect	0.0984	0.118	0.104	0.0368
control mean	0.254	0.497	0.261	0.115
		Ages 16-17		
Medicaid	0.465**	0.402**	0.505***	0.197
	[0.235]	[0.191]	[0.190]	[0.204]
Marginal Effect	0.125	0.123	0.138	0.0422
control mean	0.250	0.484	0.250	0.115

Note: Specification than in table (2)

Table A.9: Reduced Form Estimates: Functions of distance to cutoff as running variable.

	(1)	(2)
	Distance to cutoff	
	Quadratic	Cubic
	Panel A: Ages 12-13	
	A.1: BPI	
1[HS Eligible at 4]	-0.117	-0.150
	[0.106]	[0.126]
	A.2: Overweight	
1[HS Eligible at 4]	-0.156	-0.170
	[0.145]	[0.172]
Marginal Effect	-0.0398	-0.0434
	A.3: Health - use of sp. equip.	
1[HS Eligible at 4]	-0.318	-0.354
	[0.215]	[0.249]
Marginal Effect	-0.0403	-0.0452
	Panel B: Ages 16-17	
	B.1: Overweight	
1[HS Eligible at 4]	-0.328**	-0.388**
	[0.166]	[0.188]
Marginal Effect	-0.0835	-0.0991
	B.2: CESD	
1[HS Eligible at 4]	-0.270***	-0.197*
	[0.0921]	[0.114]
	Panel C: Ages 20-21	
	Arrested or Convicted	
1[HS Eligible at 4]	-0.273	-0.476**
	[0.182]	[0.200]
Marginal Effect	-0.0828	-0.142

Note: Table includes reduced form estimates for selected outcomes using several specifications. Controls excluded from the table include polynomials of the difference between family income and cutoff instead of polynomials in log income and family size at age 4, a dummy for the presence of a father figure in the child's household at age 4, race and age dummies and dummies for year and state of residence at age 4. Robust standard errors in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.10: Reduced Form estimates using fake income cutoff.

	(1)	(2)	(3)	(4)
Cutoff at	0.5	-0.5	0,0.5	-0.5,0
		Panel A: Ages 12-13		
		A.1: BPI		
Fake cutoff	-0.00933	0.329***	-0.0755	0.281**
	[0.131]	[0.122]	[0.139]	[0.127]
1[HS Eligible at 4]			-0.298**	-0.217*
			[0.133]	[0.129]
		A.2: Overweight		
Fake cutoff	0.0343	0.00477	-0.0579	-0.0984
	[0.175]	[0.177]	[0.183]	[0.179]
1[HS Eligible at 4]			-0.419**	-0.431**
			[0.170]	[0.167]
		A.3: Special Equipment		
Fake cutoff	0.271	-0.274	0.112	-0.429
	[0.231]	[0.266]	[0.234]	[0.280]
1[HS Eligible at 4]			-0.706**	-0.831**
			[0.291]	[0.326]
		Panel B: Ages 16-17		
		B.1: Overweight		
Fake cutoff	0.154	-0.247	0.0441	-0.362*
	[0.179]	[0.187]	[0.191]	[0.196]
1[HS Eligible at 4]			-0.491**	-0.583***
			[0.201]	[0.205]
		B.2: CESD		
Fake cutoff	0.134	0.115	0.0721	0.0343
	[0.0957]	[0.118]	[0.0936]	[0.120]
1[HS Eligible at 4]			-0.320***	-0.326***
			[0.108]	[0.111]
		Panel C: Ages 20-21		
		Arrested/Convicted		
Fake cutoff	0.101	-0.0559	0.0216	-0.151
	[0.191]	[0.219]	[0.196]	[0.222]
1[HS Eligible at 4]			-0.393*	-0.431**
			[0.203]	[0.203]

Note: Estimates using the same specification as table (7), but different cutoffs are set at the relative distance indicated in each column. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.11: Reduced Form clustering SE by state of residence at age 4: Children surveyed at ages 12-13, 16-17 and 20-21.

	(1)	(2)	(3)	(4)	(5)	(6)
Age groups		12-13			16-17	20-21
Variables	Overweight	Use of Special Equipment	BPI	Overweight	CESD	Arrested or Convicted
1[HS Eligible at 4]	-0.379***	-0.736**	-0.227*	-0.468**	-0.272***	-0.421**
	[0.126]	[0.301]	[0.125]	[0.212]	[0.0973]	[0.207]
Marginal Effect	-0.0953	-0.0974		-0.117		-0.120

Note: Estimates using the same specification as tables (6)-(8). Robust standard errors are reported in brackets clustered by state at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.12: Reduced Form accounting for attrition: Children surveyed at ages 12-13, 16-17 and 20-21.

	(1)	(2)	(3)	(4)	(5)	(6)
Age groups		12-13			16-17	20-21
Variables	Overweight	Use of Special Equipment	BPI	Overweight	CESD	Arrested or Convicted
1[HS Eligible at 4]	-0.0383	-0.751	-0.276	-0.231	-0.380**	-0.598**
	[0.303]	[0.484]	[0.176]	[0.311]	[0.155]	[0.264]
Observations	533	428	601	500	601	565
Control Mean	0.195	0.0968	0.735	0.164	-0.103	0.329
SD	0.399	0.298	0.902	0.373	0.920	0.473
Marginal Effect	-0.00846	-0.0852		-0.0505		-0.167

Note: Estimates using the same specification as tables (6)-(8), with sample restricted to children used in to estimate effects of Head Start at ages 12-13, 16-17 and 20-21.

Table A.13: First Stage: Alternative measures of income.

	(1)	(2)	(3)	(4)
Sample	Males	Females	Black Males	Non-Black Males
		Panel A: Alternative 1		
		A.1: Ages 12-13		
1[HS Eligible at 4]	0.496***	0.0307	0.658**	0.411
	[0.186]	[0.208]	[0.307]	[0.263]
Marginal Effect	0.149	0.00922	0.200	0.111
Observations	1,071	1,052	457	614
Control Mean	0.226	0.254	0.400	0.130
SD	0.419	0.437	0.494	0.338
		A.2: Ages 16-17		
1[HS Eligible at 4]	0.517***	0.102	0.525*	0.530**
	[0.190]	[0.212]	[0.313]	[0.265]
Marginal Effect	0.156	0.0317	0.165	0.146
Observations	1,016	996	440	576
Control Mean	0.220	0.258	0.388	0.130
SD	0.416	0.440	0.492	0.339
		Panel B: Alternative 2		
		B.1: Ages 12-13		
1[HS Eligible at 4]	0.603***	0.0606	0.824***	0.427*
	[0.182]	[0.179]	[0.278]	[0.249]
Marginal Effect	0.184	0.0188	0.260	0.112
Observations	1,280	1,275	559	721
Control Mean	0.243	0.287	0.464	0.133
SD	0.430	0.453	0.503	0.341
		B.2: Ages 16-17		
1[HS Eligible at 4]	0.507***	0.146	0.648**	0.361
	[0.188]	[0.185]	[0.303]	[0.252]
Marginal Effect	0.156	0.0467	0.209	0.0984
Observations	1,205	1,209	527	678
Control Mean	0.255	0.299	0.481	0.143
SD	0.437	0.459	0.505	0.352

Note: Probit estimates of Head Start participation on income eligibility. The marginal effect is the average change in the probability of Head Start participation across individuals as the eligibility status changes and all other controls are kept constant. Controls excluded: cubic in log family income and family size at age 4, an interaction between these two variables, a dummy indicating the presence of a father figure in the household at age 4, dummies for race, child's age and for year and state of residence at age 4. *Alternative measure 1* is equal to the measure of income used in paper, excluding Food Stamps. *Alternative 2* is equal to the sum of different income components (see subsection 5.4). Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.14: Reduced Form Estimates with alternative income measure.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample				Panel A: Ages 12-13				
Variables	Overweight	Grade Retention	Alcohol Use	School Damage	Special Education	Health condition requires sp. equip.	doctor	BPI
1[HS Eligible at 4]	-0.450**	-0.279	-0.186	0.0409	-0.301	-0.795***	-0.374*	-0.336**
	[0.204]	[0.191]	[0.184]	[0.210]	[0.190]	[0.307]	[0.224]	[0.145]
Observations	959	1,004	1,006	921	964	817	981	937
Control Mean	0.188	0.301	0.445	0.132	0.239	0.0959	0.205	0.686
SD	0.392	0.460	0.499	0.340	0.428	0.295	0.405	0.960
Marginal Effect	-0.109	-0.0881	-0.0636	0.0101	-0.0903	-0.108	-0.0941	
Sample				Panel B: Ages 16-17				
Variables	In High School	Health Status	Overweight	Drug Use	Arrested or Convicted	Expectations by age 35 be married	with family	CESD
1[HS Eligible at 4]	0.416	0.184	-0.292	-0.385**	-0.387	-0.159	-0.536**	-0.217*
	[0.274]	[0.200]	[0.222]	[0.196]	[0.268]	[0.225]	[0.258]	[0.131]
Observations	837	955	918	964	877	630	563	806
Control Mean	0.909	0.682	0.192	0.705	0.144	0.833	0.889	-0.0906
SD	0.289	0.468	0.396	0.458	0.352	0.375	0.316	0.870
Marginal Effect	0.0667	0.0608	-0.0692	-0.116	-0.0803	-0.0417	-0.125	
Sample				Panel C: Ages 20-21				
Variables	HS Diploma	Health Status	Overweight	Ever drunk	Arrested or Convicted	Ever in College	Birth Control	
1[HS Eligible at 4]	0.171	0.137	-0.297	0.0597	-0.248	-0.125	0.0523	
	[0.236]	[0.213]	[0.282]	[0.318]	[0.250]	[0.217]	[0.241]	
Observations	715	717	679	669	721	711	576	
Control Mean	0.593	0.609	0.100	0.100	0.282	0.407	0.607	
SD	0.494	0.490	0.301	0.301	0.452	0.494	0.491	
Marginal Effect	0.0592	0.0474	-0.0582	0.0112	-0.0709	-0.0420	0.0177	

Note: Probit estimates for several outcomes (OLS estimates for BPI and CESD). Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in outcome across individuals as the eligibility status changes and all other controls are kept constant. "Control Mean" is the mean outcome among observations just above the cutoff, that is, at most 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level.
The income measured used is *Alternative measure 1*, which is equal to the measure of income used in paper, excluding Food Stamps.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.15: Reduced Form Estimates with alternative income measure.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Sample				Panel A: Ages 12-13				
Variables	Overweight	Grade Retention	Alcohol Use	School Damage	Special Education	Health condition requires sp. equip.	doctor	BPI
1[HS Eligible at 4]	-0.105	-0.414***	0.0153	0.0910	-0.212	-0.496*	-0.343*	-0.367***
	[0.163]	[0.160]	[0.162]	[0.178]	[0.183]	[0.282]	[0.187]	[0.136]
Observations	1,211	1,255	1,259	1,199	1,225	1,035	1,241	1,181
Control Mean	0.160	0.289	0.416	0.119	0.241	0.0719	0.190	0.689
SD	0.367	0.455	0.494	0.325	0.429	0.259	0.394	0.968
Marginal Effect	-0.0268	-0.132	0.00530	0.0221	-0.0627	-0.0634	-0.0862	
Sample				Panel B: Ages 16-17				
Variables	In High School	Health Status	Overweight	Drug Use	Arrested or Convicted	Expectations by age 35 be married	with family	CESD
1[HS Eligible at 4]	-0.0565	0.0711	-0.233	-0.00238	-0.0726	-0.110	-0.713***	-0.250**
	[0.249]	[0.173]	[0.201]	[0.189]	[0.221]	[0.203]	[0.258]	[0.120]
Observations	1,047	1,188	1,124	1,191	1,139	800	742	1,025
Control Mean	0.926	0.688	0.199	0.667	0.145	0.862	0.889	-0.0723
SD	0.262	0.465	0.400	0.473	0.353	0.346	0.316	0.877
Marginal Effect	-0.00763	0.0239	-0.0573	-0.000756	-0.0146	-0.0288	-0.163	
Sample				Panel C: Ages 20-21				
Variables	HS	Health	Overweight	Ever	Arrested or	Ever in	Birth	
	Diploma	Status		drunk	Convicted	College	Control	
1[HS Eligible at 4]	-0.198	-0.0188	-0.00778	0.146	-0.362*	-0.327*	-0.116	
	[0.183]	[0.185]	[0.247]	[0.259]	[0.207]	[0.196]	[0.212]	
Observations	914	924	844	878	922	906	751	
Control Mean	0.629	0.608	0.123	0.128	0.288	0.467	0.604	
SD	0.485	0.490	0.330	0.336	0.455	0.501	0.492	
Marginal Effect	-0.0692	-0.00657	-0.00167	0.0284	-0.104	-0.114	-0.0416	

Note: Probit estimates for several outcomes (OLS estimates for BPI and CESD). Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in outcome across individuals as the eligibility status changes and all other controls are kept constant. "Control Mean" is the mean outcome among observations just above the cutoff, that is, at most 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level.
The income measured used is *Alternative measure 2*.
* significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.16: Linear Probability Model: Two Stage Least Squares Estimates.

	(1)	(2)	(3)	(4)
	Ages 12-13		Ages 16-17	Ages 20-21
	Special Equipment	Overweight	Overweight	Arrested or Convicted
Head Start	-0.202*	-0.362**	-0.442**	-0.140
	[0.108]	[0.162]	[0.192]	[0.190]

Note: Participation in program is instrumented with eligibility status at age four. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. The standard errors are obtained by block-bootstrap (500 replications; block is the cell year-state of residence at age 4). * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.17: Mother Fixed Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
VARIABLES		Test Score			PIAT-RR			BPI	
Year of Birth	1972-1986	1987-2000	1972-2000	1972-1986	1987-2000	1972-2000	1972-1986	1987-2000	1972-2000
Data	1986-2000	1986-2008	1986-2008	1986-2000	1986-2008	1986-2008	1986-2000	1986-2008	1986-2008
					Panel A: All Sample				
Head Start	0.0348	-0.116	-0.0258	0.191***	0.162*	0.151***	0.0131	0.0408	0.0499
	[0.0557]	[0.0823]	[0.0401]	[0.0676]	[0.0982]	[0.0480]	[0.0524]	[0.0637]	[0.0347]
Head StartXAge 7-10	0.0754*	0.151***	0.0994***	-0.150***	-0.129**	-0.130***	0.00642	0.0307	0.0141
	[0.0419]	[0.0484]	[0.0316]	[0.0472]	[0.0562]	[0.0360]	[0.0407]	[0.0410]	[0.0284]
Head StartXAge 11-14	0.0338	0.113**	0.0593*	-0.257***	-0.195***	-0.218***	-0.0126	-0.0275	-0.0224
	[0.0452]	[0.0539]	[0.0344]	[0.0535]	[0.0662]	[0.0416]	[0.0480]	[0.0500]	[0.0341]
Preschool	0.0409	0.111**	0.0969***	0.000273	0.103	0.0857**	0.0161	0.0407	0.0757**
	[0.0582]	[0.0568]	[0.0341]	[0.0697]	[0.0712]	[0.0426]	[0.0580]	[0.0523]	[0.0315]
PreschoolXAge 7-10	-0.0706*	-0.0854**	-0.0811***	-0.0714*	-0.0762*	-0.0905***	0.0272	0.0228	0.0269
	[0.0372]	[0.0374]	[0.0264]	[0.0431]	[0.0448]	[0.0308]	[0.0350]	[0.0303]	[0.0221]
PreschoolXAge 11-14	-0.0731*	-0.0640	-0.0755***	-0.0467	-0.0639	-0.0755**	-0.0470	-0.0521	-0.0462*
	[0.0409]	[0.0401]	[0.0285]	[0.0491]	[0.0503]	[0.0344]	[0.0418]	[0.0370]	[0.0270]
Observations	19,860	16,837	36,697	17,283	14,396	31,679	19,000	16,832	35,832
Mean Dependent Variable	-0.208	0.0988	0.592	0.136	0.430	0.269	0.522	0.147	0.346
SD	0.972	1.068	-0.0674	0.966	1.007	0.996	0.978	0.988	1.000
P-Value: Joint Test HS	0.0213	0.0161	0.00236	5.11e-06	0.0278	1.07e-06	0.948	0.441	0.181

Note: Controls excluded from table include child's gender, first born status, and age at test and year fixed effects and mother fixed effects

Table A.18: Reduced Form: Estimates by Race.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Overweight	Grade Retention	Alcohol Use	Ages 12-13 Special Education	Health condition requires sp. equip.	BPI	Overweight	Ages 16-17 Arrested or Convicted	CESD	Ages 20-21 Arrested or Convicted
Panel A: Black										
1[HS Eligible at 4]	-0.816***	-0.355	-0.507*	0.161	-0.911*	-0.105	-0.713**	-0.668*	-0.218	-0.127
	[0.269]	[0.283]	[0.266]	[0.284]	[0.474]	[0.213]	[0.304]	[0.367]	[0.151]	[0.359]
RW algorithm										
H0 rejected at 10%	Yes	No	No	No	No	No	No	No	No	No
Observations	536	540	553	529	387	524	500	495	457	410
Control Mean	0.305	0.200	0.426	0.196	0.186	0.541	0.321	0.0417	-0.233	0.0811
SD	0.464	0.404	0.499	0.401	0.394	0.838	0.471	0.202	0.821	0.277
Marginal Effect	-0.211	-0.116	-0.163	0.0464	-0.152		-0.182	-0.123		-0.0330
Panel B: Non Black										
1[HS Eligible at 4]	-0.259	-0.280	-0.0937	-0.669***	-1.288***	-0.393**	-0.261	0.0781	-0.318**	-0.499*
	[0.228]	[0.228]	[0.212]	[0.244]	[0.432]	[0.176]	[0.257]	[0.241]	[0.161]	[0.261]
RW algorithm										
H0 rejected at 10%	No	No	No	Yes	Yes	No	No	No	No	No
Observations	668	699	673	679	469	648	596	569	559	503
Control Mean	0.153	0.348	0.495	0.265	0.0822	0.722	0.181	0.227	-0.0653	0.367
SD	0.362	0.478	0.502	0.444	0.277	0.950	0.387	0.421	0.809	0.485
Marginal Effect	-0.0617	-0.0861	-0.0331	-0.174	-0.152		-0.0635	0.0168		-0.137

Note: Probit estimates for several outcomes (OLS estimates for BPI and CESD). Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birthweight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in the probability of Head Start participation across individuals as the eligibility status changes and all other controls are kept constant. "Control Mean" is the mean outcome among observations just above the cutoff, that is, at 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table A.19: Reduced Form Estimates at Ages 12-13, by year of birth.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Health condition requires special equipment		Overweight		Grade Retention		Special Education		BPI	
Year of birth	1977-1984	1985-2000	1977-1984	1985-2000	1977-1984	1985-2000	1977-1984	1985-2000	1977-1984	1985-2000
1[HS Eligible at 4]	-1.324***	-0.184	-0.374*	-0.362	-0.333	-0.282	-0.474*	0.0233	-0.348**	-0.187
	[0.442]	[0.447]	[0.219]	[0.261]	[0.213]	[0.265]	[0.246]	[0.247]	[0.156]	[0.235]
Observations	578	325	656	496	737	519	695	523	689	503
Control Mean	0.0952	0.109	0.161	0.264	0.270	0.316	0.216	0.266	0.078	0.140
SD	0.295	0.315	0.370	0.444	0.446	0.468	0.414	0.445	0.787	0.458
Marginal Effect	-0.194	-0.0285	-0.0902	-0.104	-0.111	-0.0894	-0.143	0.00713	0.890	0.918

Note: Probit (and OLS for BPI) estimates for several outcomes. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birthweight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Marginal effect is average marginal change in the probability of Head Start participation across individuals as the eligibility status changes and all other controls are kept constant. "Control Mean" is the mean outcome among observations just above the cutoff, that is, at 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

B Figures

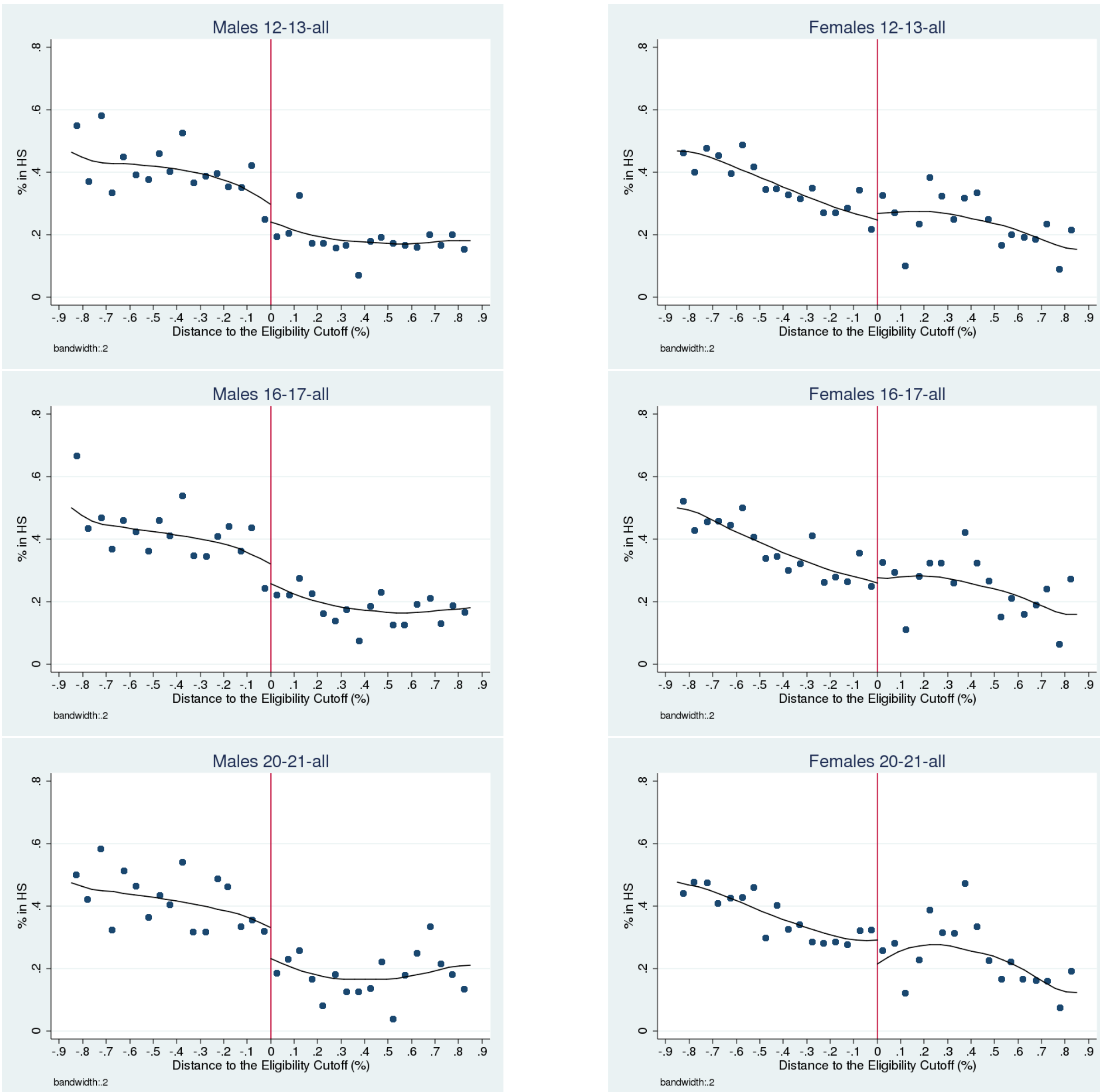

Figure B.1: Proportion of children in Head Start, by eligibility status.

Note: The continuous lines in figure are local linear regression estimates of Head Start participation on percentage distance to cutoff; regressions were run separately on both sides of the cutoff and the bandwidth was set to 0.2. Circles in figures represent mean Head Start participation by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

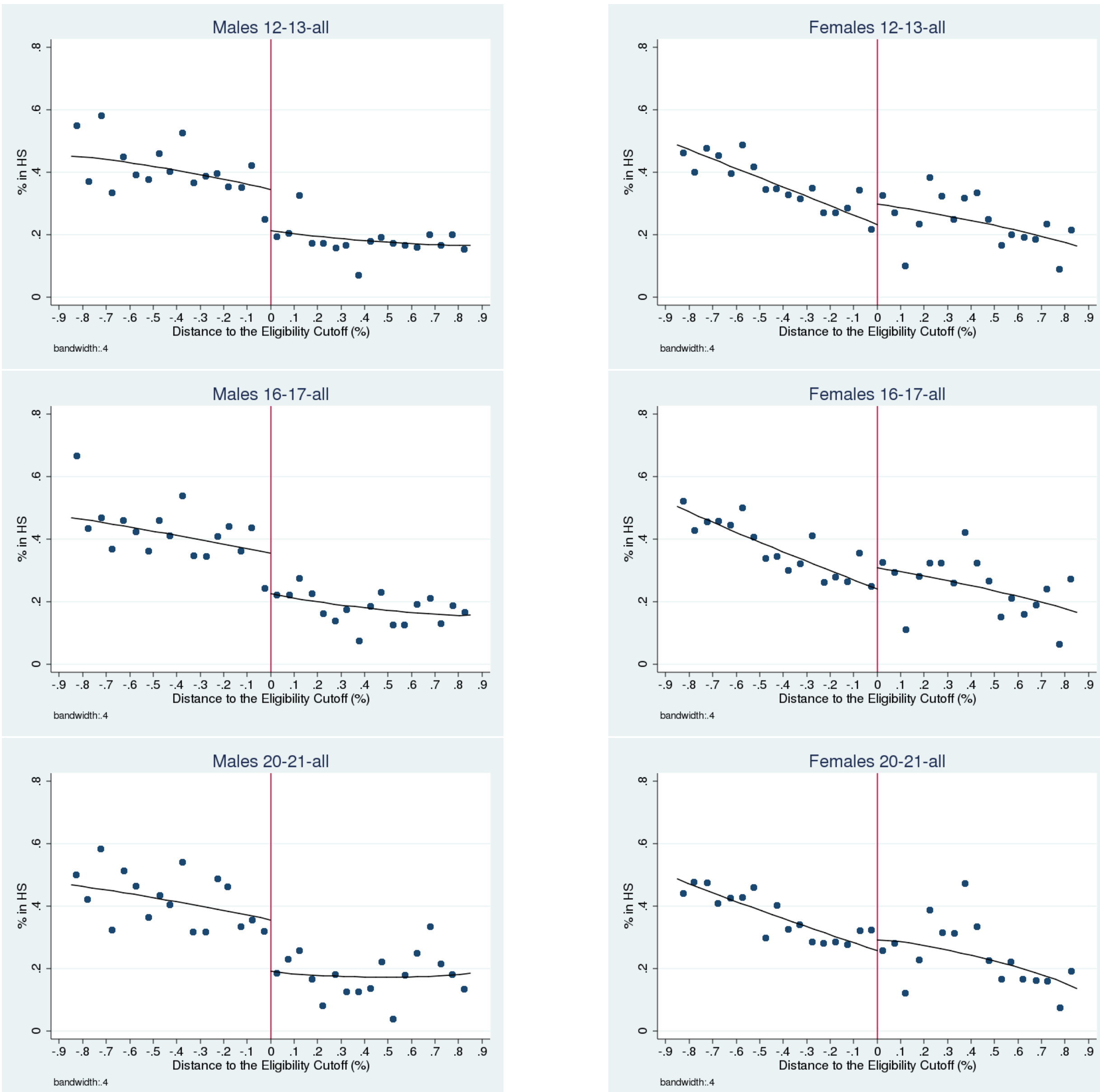

Figure B.2: Proportion of children in Head Start, by eligibility status.

Note: The continuous lines in figure are local linear regression estimates of Head Start participation on percentage distance to cutoff; regressions were run separately on both sides of the cutoff and the bandwidth was set to 0.4. Circles in figures represent mean Head Start participation by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

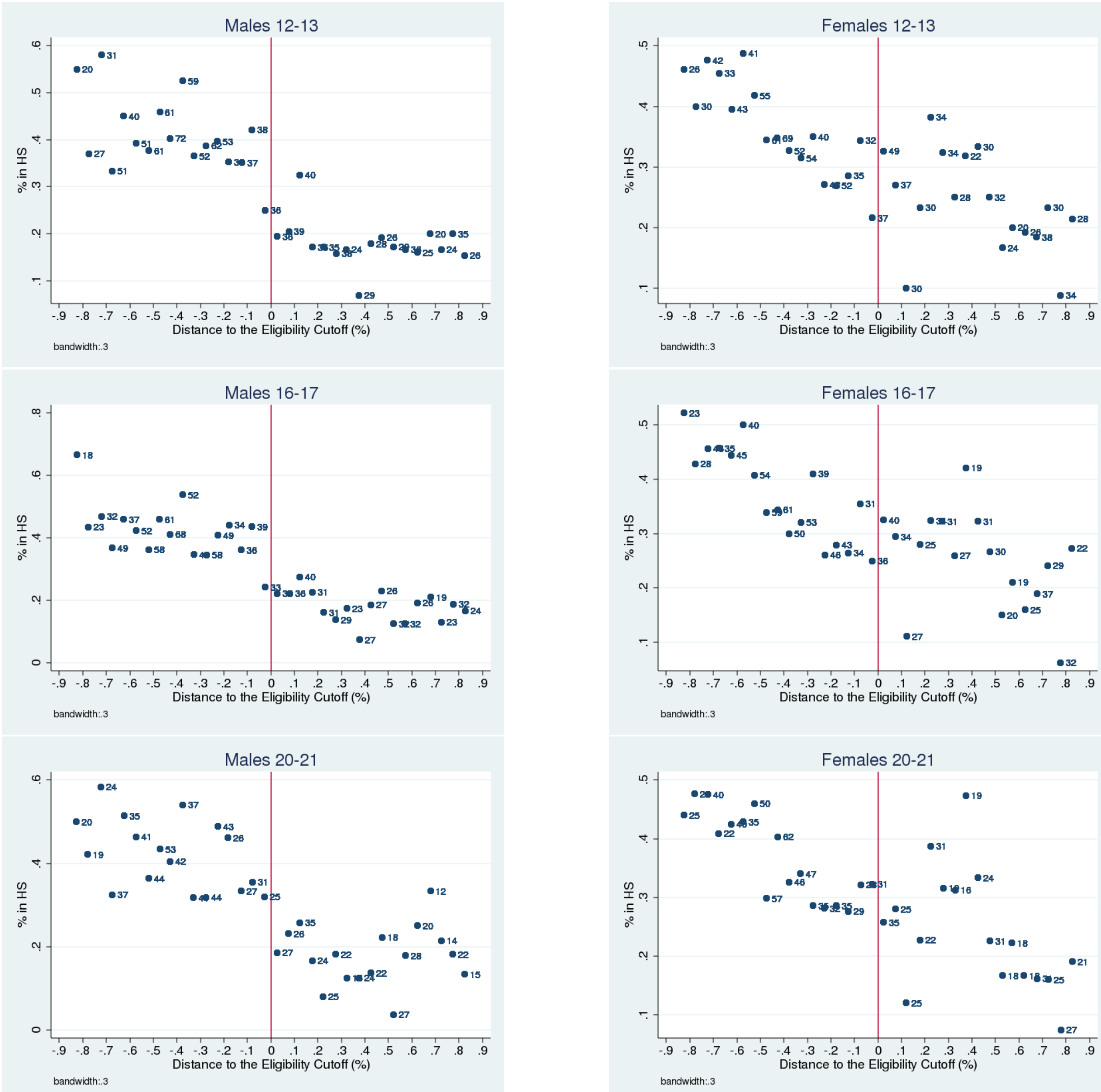

Figure B.3: Proportion of children in Head Start, by eligibility status and gender.

Note: The circles include the average proportion of children in Head Start (these are the same as in figure 3) and the number next to each circle is the sample size used to compute the average. The kernel used was epanechnikov.

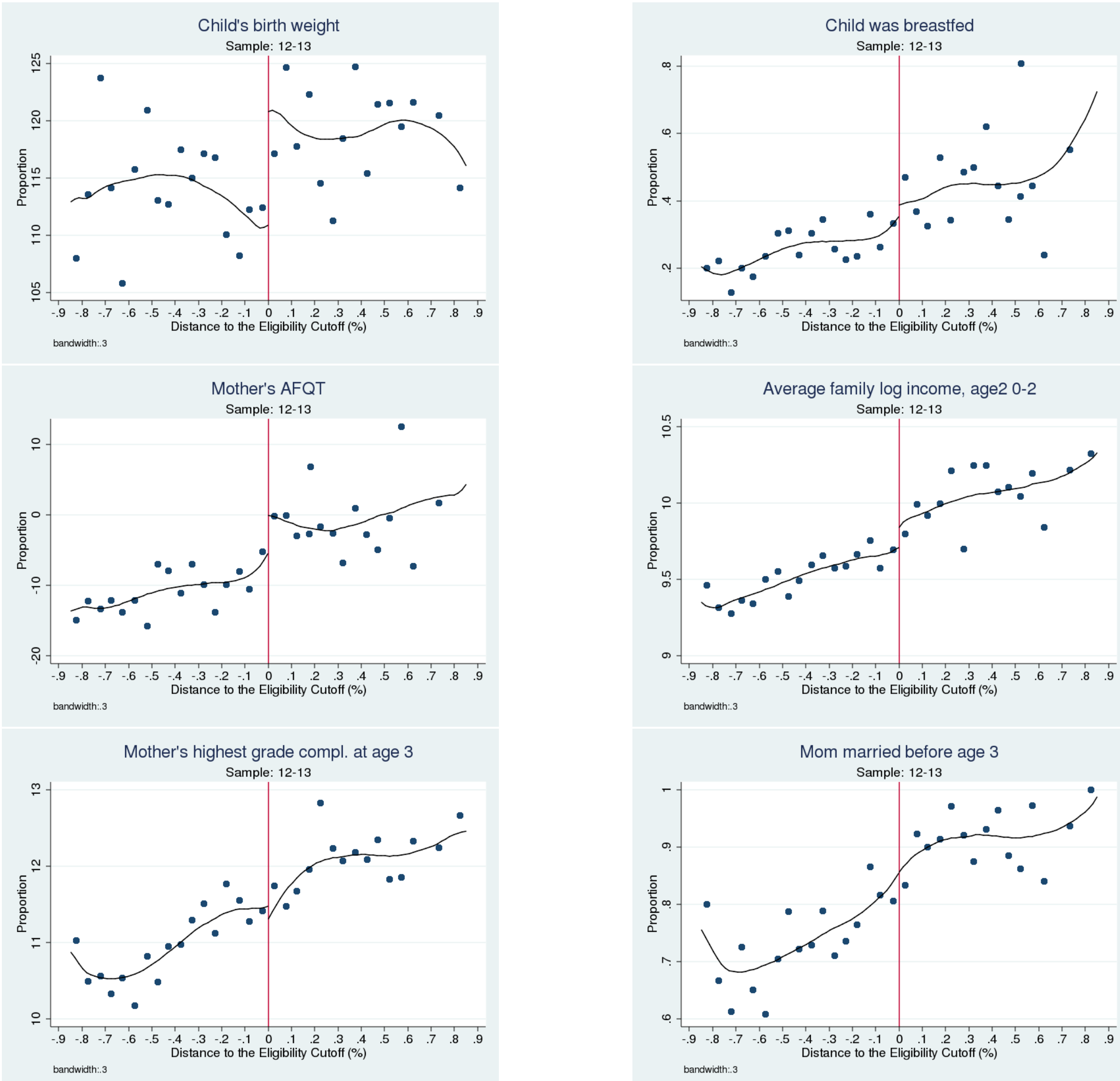

Figure B.4: Average outcomes by eligibility status: Falsification, Bandwidth = 0.3.

Note: The continuous lines in Figure present local linear regression estimates of several outcomes on percentage distance to cutoff. The kernel used was epanechnikov.

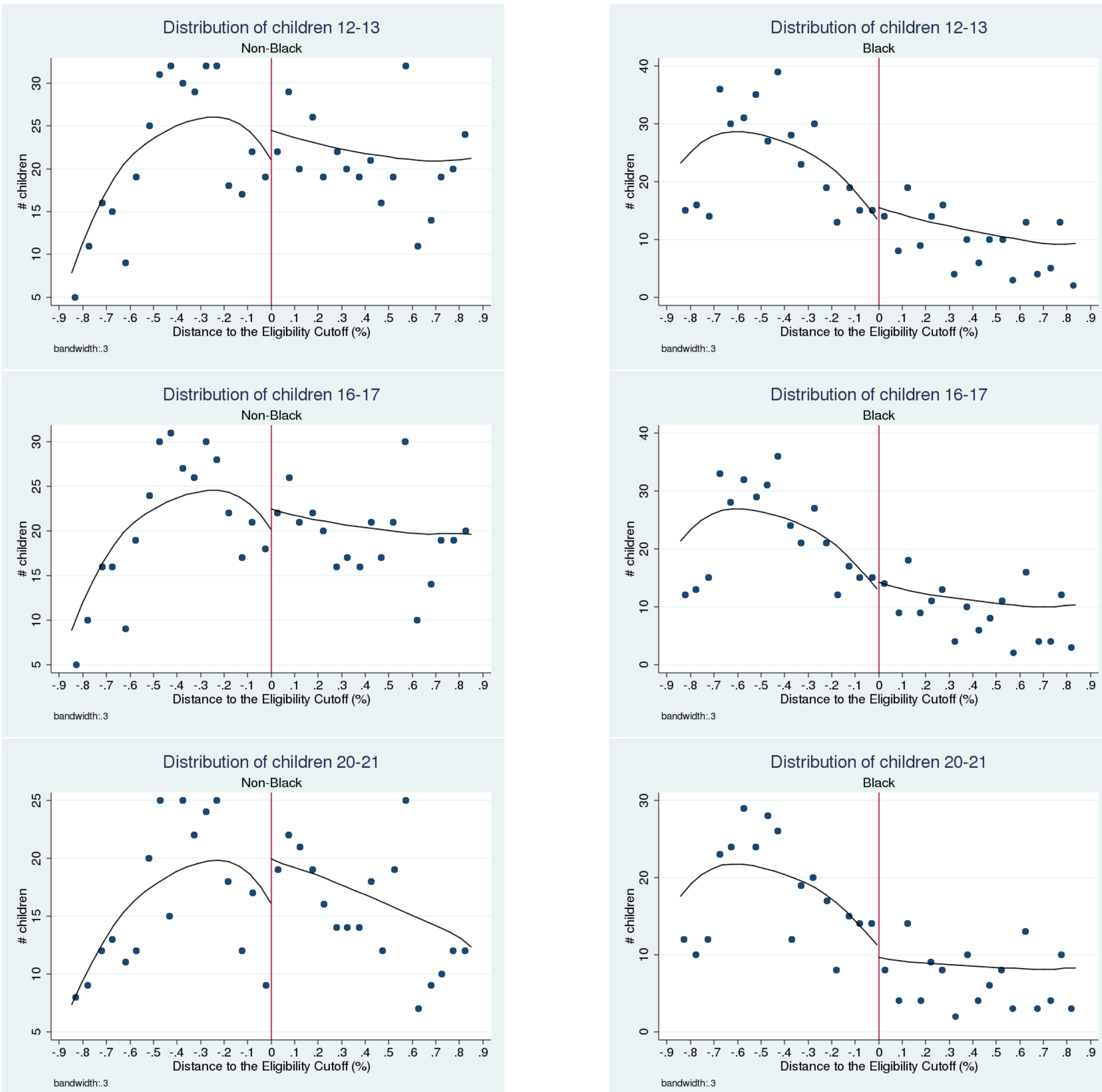

Figure B.5: Distribution of children around the cutoff.

Note: The continuous lines in figure are local linear regression estimates of the number of children per interval of width 0.05 of relative income-distance to cutoff; regressions were run separately on both sides of the cutoff and the bandwidth was set to 0.3. Circles in figures represent mean number of children by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

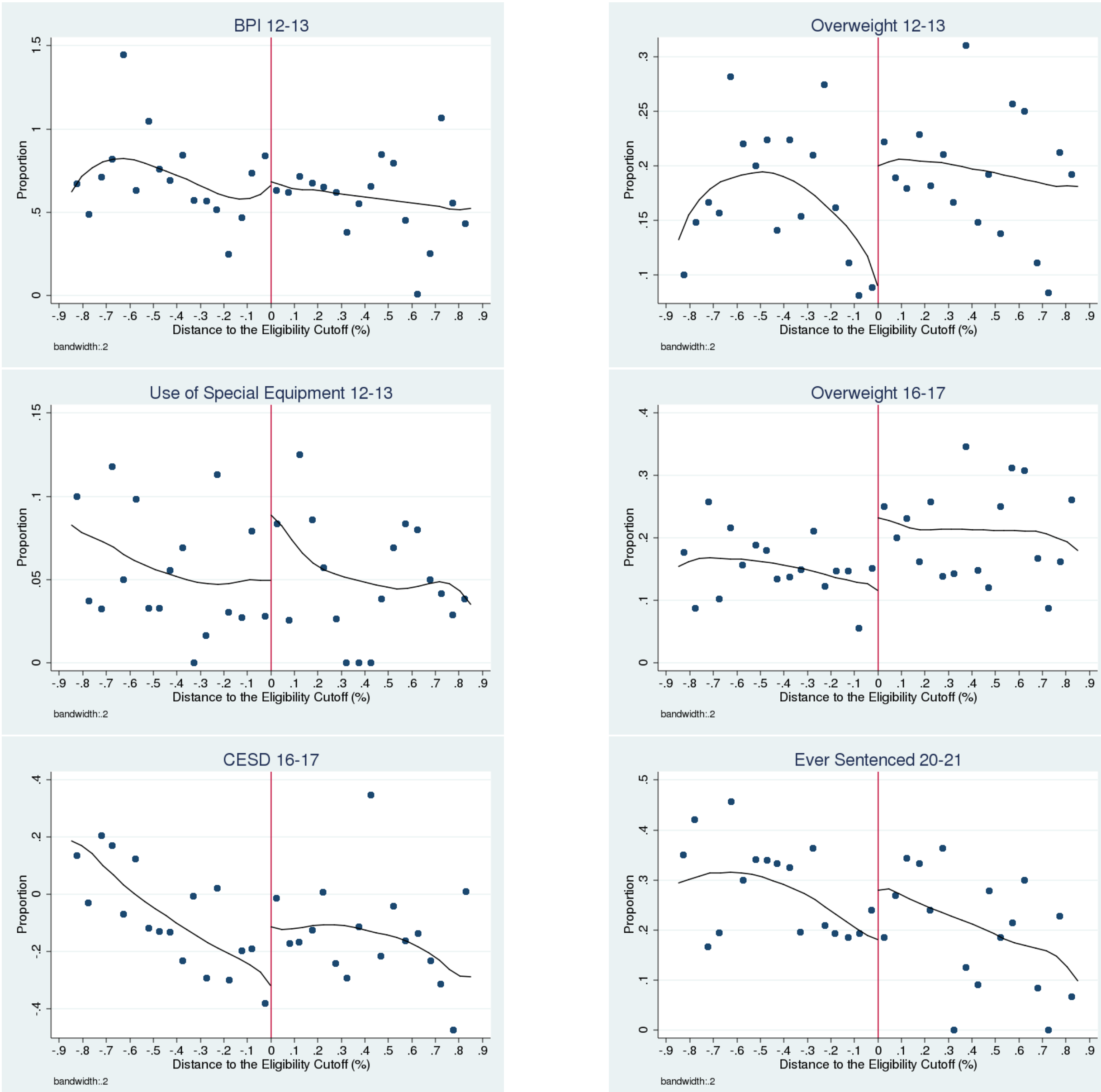

Figure B.6: Average outcomes by eligibility status, Bandwidth = 0.2.

Note: The continuous lines in Figure present local linear regression estimates of several outcomes on percentage distance to cutoff. The bandwidth was set to 0.2. Circles in figures represent the mean outcome by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

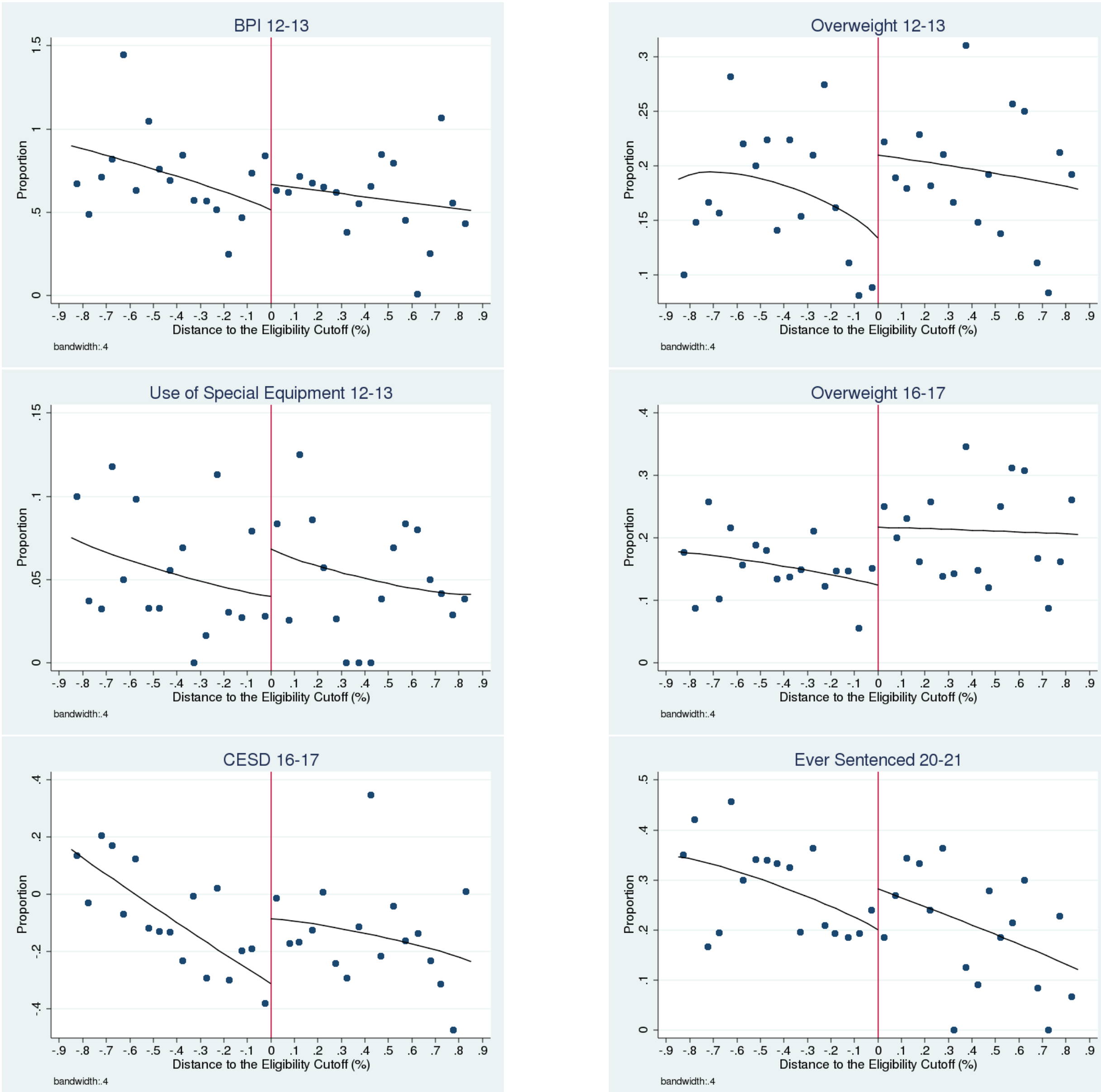

Figure B.7: Average outcomes by eligibility status, Bandwidth = 0.4.

Note: The continuous lines in Figure present local linear regression estimates of several outcomes on percentage distance to cutoff. The bandwidth was set to 0.4. Circles in figures represent the mean outcome by cell within intervals of 0.05 of distance to cutoff. The kernel used was epanechnikov.

C Eligibility to Head Start

According to the Head Start Act, Sec. 645(a)(2)(A) "children from low-income families shall be eligible for participation in programs assisted under this subchapter (*Head Start*) if their families' incomes are below the poverty line, or if their families are eligible or, in the absence of child care, would potentially be eligible for public assistance"[49]. Alternatively, grantees may enroll up to 10% of children from "over-income" families, as allowed by 45 CFR Part 1305 and, if applicable, by Section 645(a)(2) and (d) of the Head Start Act (*the latter refers to Head Start centers that operate in an Indian tribe*)[50]. See table A.1 in Appendix A for a summary of Head Start's legislation since the program was launched in 1965. The eligibility criteria have been unchanged throughout the period of analysis.[51]

A low-income family is a family whose income before taxes is below the poverty line or a family that is receiving public assistance, even if the family's income exceeds the poverty line. The U.S. Department of Health and Human Services considers public assistance as AFDC/TANF and SSI (See 45 CFR Part 1305.2). In section 3 we explain why we did not impute SSI eligibility.

The income should include:

> total cash receipts before taxes from all sources, with the exceptions noted below. Income includes money wages or salary before deductions; net income from non-farm self-employment; net income from farm self-employment; regular payments from Social Security or railroad retirement; payments from unemployment compensation, strike benefits from union funds, workers' compensation, veterans benefits (with the exception noted below), public assistance (including Temporary Assistance for Needy Families, Supplemental Security Income, Emergency Assistance money payments, and non-Federally funded General Assistance or General Relief money payments); training stipends; alimony, child support, and military family allotments or other regular support from an absent family member or someone not living in the household; private pensions, government employee pensions (including military retirement pay), and regular insurance or annuity payments; college or university scholarships, grants, fellowships, and assistantships; and dividends, interest, net rental income, net royalties, and periodic receipts from estates or trusts; and net gambling or lottery winnings.

Income does not include capital gains; any assets drawn down as withdrawals from a bank, the sale of property, a house or a car; or tax refunds, gifts, loans, lump-sum inheritances, one-time insurance payments, or compensation for injury. Also excluded are noncash benefits, such as the employer-paid or union-paid portion of health insurance or other employee fringe benefits; food or housing received in lieu of wages; the value of food and fuel produced and consumed on farms; the imputed value of rent from owner-occupied non-farm or farm housing; and such Federal non-cash benefit programs as Medicare, Medicaid, food stamps, school lunches, and housing assistance, and certain disability payments made to disabled children of Vietnam veterans as prescribed by the Secretary of Veterans Affairs.[52]

The Department of Health and Human Services considers that the income period of time to be considered for eligibility is the 12 months immediately preceding the month in which application or reapplication for enrollment of a child in a Head Start program is made, or the calendar year immediately

[49]See Title VI, Subtitle A, Chapter 8, Subchapter B, of the Omnibus Budget Reconciliation Act of 1981, Public Law 97-35 (42 USC 9840) and its amends (http://www.law.cornell.edu/uscode/html/uscode42/usc_sec_42_00009840—-000-.html).

[50]Indian Tribes meeting the conditions specified in 45 CFR 1305.4(b)(3) are excepted from this limitation (see 45 CRF Part 1305 - source 57 FR 46725, Oct. 9, 1992, as amended at 61 FR 57226, Nov. 5, 1996).

[51]See www.eric.ed.gov and Zigler and Valentine, 1979.

[52]See http://eclkc.ohs.acf.hhs.gov/hslc for the Head Start Program Definition of income and Federal Poverty Guidelines.

preceding the calendar year in which the application or reapplication is made, whichever more accurately reflects the family's current needs. We use income from last calendar year because it is the income measure available in NLSY79. As of our knowledge D.H.H.S. does not issue any specific definition of "family unity" and therefore we use NLSY79's definition.

To check the veracity of declared income, centers are required to verify the following proofs: Individual Income Tax Form 1040, W-2 forms, pay stubs, pay envelopes, written statements from employers, or documentation showing current status as recipients of public assistance, and should keep a signed statement by an employee identifying which of these documents was examined and stating that the child is eligible to participate in the program. Some centers do not keep an accurate register.

Given that there are two routes of eligibility to Head Start for each child we perform two separate comparisons:

1. Impute child's poverty status: the child is in a poor family if the annual family gross income is below or equal to the Federal Poverty Guideline for each year of data available.

2. Impute child's family AFDC/TANF eligibility. See "AFDC Eligibility Requirements" below for a detailed description.

Finally, the child is eligible to participate in Head Start if she is in a poor family or if she is in an AFDC/TANF eligible family.

In order to restrict our analysis to a comparable group of individuals we restrict our sample to children whose family income at age four was between 15 percent and 185 percent of the maximum level of income that would enable them to be Head Start eligible. To obtain this level of income we perform several comparisons:

1. If the child's family does not verify the categorical requirements to be entitled to AFDC/TANF, the maximum gross income that would have allowed Head Start eligibility is the Federal Poverty Guideline.

2. If the child's family is categorically eligible to AFDC/TANF, several scenarios may emerge[53]:

   (a) if the family is not receiving income from AFDC, or if this information is missing, then two income tests must be verified in order to become AFDC income eligible:

      i. if the *gross income test* is not valid, the maximum level of income that would allow Head Start eligibility is $MAX(m \times 12 \times Need\ Standard, Federal\ Poverty\ Guideline)$ where m is 1.5 for the years of 1982, 1983 and 1984, and 1.85 from then onwards. We use the Need Standard in the state of residence and year at age four and the Federal Poverty Guideline of the year in which the child turned four years old.

      ii. if *the gross income test* is verified, then the relevant cutoff point will be given by $MAX(MIN(m \times 12 \times NS, 12 \times NS + Annual\ Deductions), Federal\ Poverty\ Guideline)$ where NS is the Need Standard in the state of residence at age four.

   (b) since 1982, if the family is currently receiving income from AFDC/TANF only *the gross income test* is performed and the maximum level of income above which the family no longer is income eligible is given by $MAX(m \times 12 \times Need\ Standard, Federal\ Poverty\ Guideline)$

   (c) the gross income test had not been implemented as of 1979, 1980 and 1981, and the cutoff is given by $MAX(12 \times NS + Annual\ Deductions, Federal\ Poverty\ Guideline)$

[53]See "AFDC Eligibility Requirements" for detailed description of the gross and countable income tests.

We then define a variable we call *"percentage distance to cutoff"* which results from the percentage difference between the family income and the threshold income level that results from the previous set of comparisons, and use it to restrict our sample to the set of individuals located near their relevant discontinuity cutoffs.

AFDC Eligibility Requirements[54]

Eligibility for AFDC requires that household contains at least one child less than eighteen years old, and has sufficiently low income and assets levels. Additionally, children in two-parents families may be eligible to AFDC-UP (Unemployed Parent), which requires that parents satisfy a work history requirement and work less than 100 hours per month while on welfare.[55] There are two income tests that an applicant family must pass in order to become AFDC income eligible (U.S. Congress, 1994):

- the gross income test: a gross income limit for AFDC eligibility of 150 percent of the state's standard was imposed by The Omnibus Reconciliation Act (OBRA) of 1981, and raised to 185 percent by The Deficit Reduction Act of 1984;
- the countable income test: it requires that family income after some disregards must be less than the state's need standard. The countable income is the gross income subtracted of work related expenses, child care expenses, child support disregards up to a maximum.[56]

Eligibility is re-assessed annually, and for those who are already recipients of AFDC/TANF only the first income test is required. To impute the threshold for AFDC/TANF income eligibility for each child we merge the need standard, child support disregards, child care expenses and work related expenses information with the child-level data from the CNLSY by state of residence and family size for each year.

Federal AFDC law requires that all income received by an AFDC recipient or applicant must be counted against the AFDC grant except that income explicitly excluded by definition or deduction. The disregards can be computed as follows. Prior to 1981 there was no allowance for work related expenses and child care expenses were capped at 160 dollars per month. The OBRA of 1981 continued to cap the deduction for child care costs at 160 dollars per month and set the work incentive disregard for work expenses at 75 dollars per month. These allowances were increased by the Family Support Act of 1988 that raised work expenses disregards to 90 dollars per month and the child care expenses to 200 dollars per each child under two years old and 175 dollars for month per each child two years or older. This was effective from October 1, 1989, but as our income values are annual we used it from 1990 onwards. In 1996, work related expenses were subsequently raised to 100 dollars, 200 dollars in 1997 and 250 dollars per month since 1998. Between 1997 and 1999, child care expenses were set at 200 dollars per each child either she was under or older than two years old. Additionally, the Deficit Reduction Act of 1984 established a monthly disregard of 50 dollars of child support received by family, that is valid from 1985 (inclusive) onwards. As the last age in which we impute program eligibility is 7 years old and the youngest child in our sample was born in 2000, 2006 is the last year in which eligibility status should be imputed.[57]

---

[54] See Hoynes, 1996, for a description of AFDC eligibility rules.

[55] The Family Support Act of 1988 mandates that states set up AFDC-UP programs, but it allows states to limit benefits to six months per year.

[56] Details on all state-specific values can be found in the Welfare Rules Database of the Urban Institute.

[57] NLSY79 and CNLSY surveys were not conducted in odd years after 1994, and income, family size, child's mother cohabiting status and state of residence were not imputed for these years.

Since NLSY79 does not contain systematic collection of child care and work related expenses we assume that families fully deduct the full disregard of child care expenses for all children under 6 years old and no disregard for older children (as is imposed by AFDC requirements), and deduct the full amount of work related expenses if the mother or her spouse is working.

Need standard, work related expenses, child care expenses and child support disregards are defined in monthly levels but were converted into annual values to be comparable with the annual gross income measure available in the NLSY79.

Treatment of Earned Income Tax Credit (EITC) has changed over time. Prior to 1981, EITC was counted only when received, however the OBRA of 1981 requires to assume that working AFDC recipients received a monthly EITC if they appeared eligible for it and regardless of when or if the credit was actually available. The 1984 legislation returned to prior law policy with respect to the EITC: it was to be counted only when actually received. The Family Support Act of 1988 required to disregard the EITC in determining eligibility for and benefits under the AFDC program. As EITC information in NLSY79 started to be recovered with the 2000 wave we ignore the EITC in our analysis.[58]

States are also required by Federal law to disregard certain earned income when determining the amount of benefits to which a recipient family is entitled, which we did not take into account as we only impute income eligibility to the program.

Our treatment of the data regarding stepparent's or mother's partner income was as follows. The OBRA of 1981 required that a portion of the stepparent income to count as part of the income, however, as NLSY79 total income does not include mother's partner income, we do not include it in the definition of income, but as long as child's mother is married, her husband's income is included in the definition of family income (regardless of whether she is married or not with the child's natural father). Also, if mother is cohabiting, her partner will not be included in the family size variable.

When determining AFDC/TANF eligibility we took into account the program categorical requirements with respect to the family structure. Eligibility under the traditional AFDC program requires that a child resides in a female-headed household, which we considered as a family where a father-figure is missing. However, children in two-parents households may still be eligible under the AFDC-Unemployed Parent program in those states in which the program is available[59]. Eligibility for AFDC-UP is limited to those families in which the principal wage earner is unemployed but has a history of work. We consider that the principal wage earner has a "history of work" if the father was employed for less than forty weeks in the previous calendar year (as Currie and Gruber, 1996).[60] We do not perform the assets test required by AFDC, as information on assets is only available after 1985 in NLSY79.

[58]Given the extensive set of robustness checks performed, we are convinced that our results are not sensitive to this.

[59]In 1988, the Family Support Act required all States, effective October 1, 1990, to provide AFDC-UP (except American Samoa, Puerto Rico, Guam, and the Virgin Islands until funding ceilings for AFDC benefits in these areas are removed). The two-parent program reverted to optional status for all States after September 30, 1998.

[60]Since 1971, Federal regulations have specified that an AFDC parent must work fewer than 100 hours in a month to be classified as unemployed, unless hours are of a temporary nature for intermittent work and the individual met the 100-hour rule in the two preceding months and is expected to meet it the following month. See U.S. Congress, 1994, for the specific requirements.

D Heterogenous Effects

Figure D.1 displays the range of income cutoffs and family size in our data over which we are able to identify the effects of Head Start. This figure shows that relatively to the standard regression discontinuity, we can exploit additional variation. This "continuum of discontinuities" allows us to go beyond the traditional regression discontinuity design and identify treatment effects for individuals over a wide range of values for income and family size (the two main running variables). Black, Galdo and Smith (2005) also recognize the potential of multiple discontinuities to identify heterogeneous effects of the program. Figure D.1 also includes the number of children per family size and intervals of income cutoffs of \$571. This figure shows that β in equation (4) is a weighted average of effects for children around different cutoffs, where the largest weights correspond to children living in families with 2 to 6 members and with cutoffs varying between \$US10,000-20,000.

Therefore, we also consider models where β varies explicitly across individuals:

$$Y_i = \alpha + \beta_i HS_i + g(Z_i, X_i) + \varepsilon_i. \tag{6}$$

If there is perfect compliance, in the sense that $HS_i = E_i$, then Hahn, Todd and van der Klauww (2001) show that using regression discontinuity we can estimate $E(\beta_i | Z_i = \bar{Z}_i, X_i)$ (the average effect of the program conditional on Z) over the support of $\bar{Z}$ in the data. Under the weaker condition that $HS_i = E_i$ only when $E_i = 0$, we estimate $E(\beta_i | Z_i = \bar{Z}_i, X_i, HS = 1)$ (Battistin and Rettore, 2007). More generally, one can have non-compliance on both sides of the discontinuity, in which case we obtain an estimate of a Local Average Treatment Effect (LATE; Imbens and Angrist, 1994) at $Z_i = \bar{Z}_i$, over the support of $\bar{Z}$ (the set of income eligibility cutoffs), or $E(\beta_i | HS(\bar{Z}_i - \delta_i) - HS(\bar{Z}_i + \delta_i) = 1, X_i, Z_i = \bar{Z}_i)$, for $\delta_i > 0$. The latter is the case in our data. There are two reasons why this parameter may vary with $\bar{Z}$: i) β_i is a function of Z_i (income at the time eligibility is measured); or ii) even if there is independence between β_i and Z_i in the sample, independence may not hold conditional on program participation if HS_i depends on β_i.[61] In our setting, it is impossible to distinguish the two.

We implement this estimator as follows. In our case, the income cutoffs vary mainly with family size (besides state, year and mother's marital status), so there is a threshold for $\bar{Z}_f$ for each family size f. Therefore, it is possible to estimate effect of the program around each cutoff $\bar{Z}_f, f = \{2, ..., 15\}$, that is, $E\left(\beta_f | HS\left(\bar{Z}_f - \delta\right) - HS\left(\bar{Z}_f + \delta\right) = 1, Z_i = \bar{Z}_f\right)$. The effect estimated under homogeneous effects is a weighted average of the effect around each threshold

$$\beta = \sum_{f=2}^{15} \frac{\beta_f 1\left[\left|Z_{if} - \bar{Z}_f\right| \leq h\right]}{\sum_{f=2}^{15} 1\left[\left|Z_{if} - \bar{Z}_f\right| \leq h\right]}$$

where h is the bandwidth to trim the sample around each cutoff. Therefore, we recover β_f by estimating the following model:

$$Y_i = \alpha + h^*(Z_i, X_i) \times E_i + g(Z_i, X_i) + \varepsilon_i \tag{7}$$

[61] Say program participation is determined by equation (5), and β_i is correlated with v_i. Even if β_i is independent of Z_i,

$$E[\beta_i | HS(\bar{Z}_i - \delta_i) - HS(\bar{Z}_i + \delta_i) = 1, X_i] = E[\beta_i | -\eta - \tau - h(\bar{Z}_i, X_i) < v_i \leq -\eta - h(\bar{Z}_i, X_i), X_i],$$

which is a function of $\bar{Z}_i$. Intuitively, the set of v_i for individuals at the margin varies with the level of $\bar{Z}_i$.

For simplicity, we model $h^*(Z_i, X_i)$ as:

$$h^*(Z_i, X_i) = \beta_0 + \beta_1 \ln Y_{4i} + \beta_2 (\ln Y_{4i})^2 + \beta_3 (\ln Y_{4i})^3 + \beta_4 FS_{4i} + \beta_5 FS_{4i}{}^2 + \beta_6 FS_{4i}{}^3 + \beta_7 \ln Y_{4i} FS_{4i}$$

where Y_{4i} is child's family income at age 4 and FS_{4i} is the family size. Potentially we would like to estimate $h^*(Z_i, X_i)$ using a more flexible specification, but our sample size forces us to be parsimonious. Even with such parsimonious specification our estimates of this function in the Section 5 are quite imprecise, and results should be seen as suggestive and illustrative of the potential of this approach. We report estimates of β_j, $j = 0, ..., 7$, as well as estimates of the average partial effect of Head Start. We also display graphical representations of $h^*(Z_i, X_i)$.

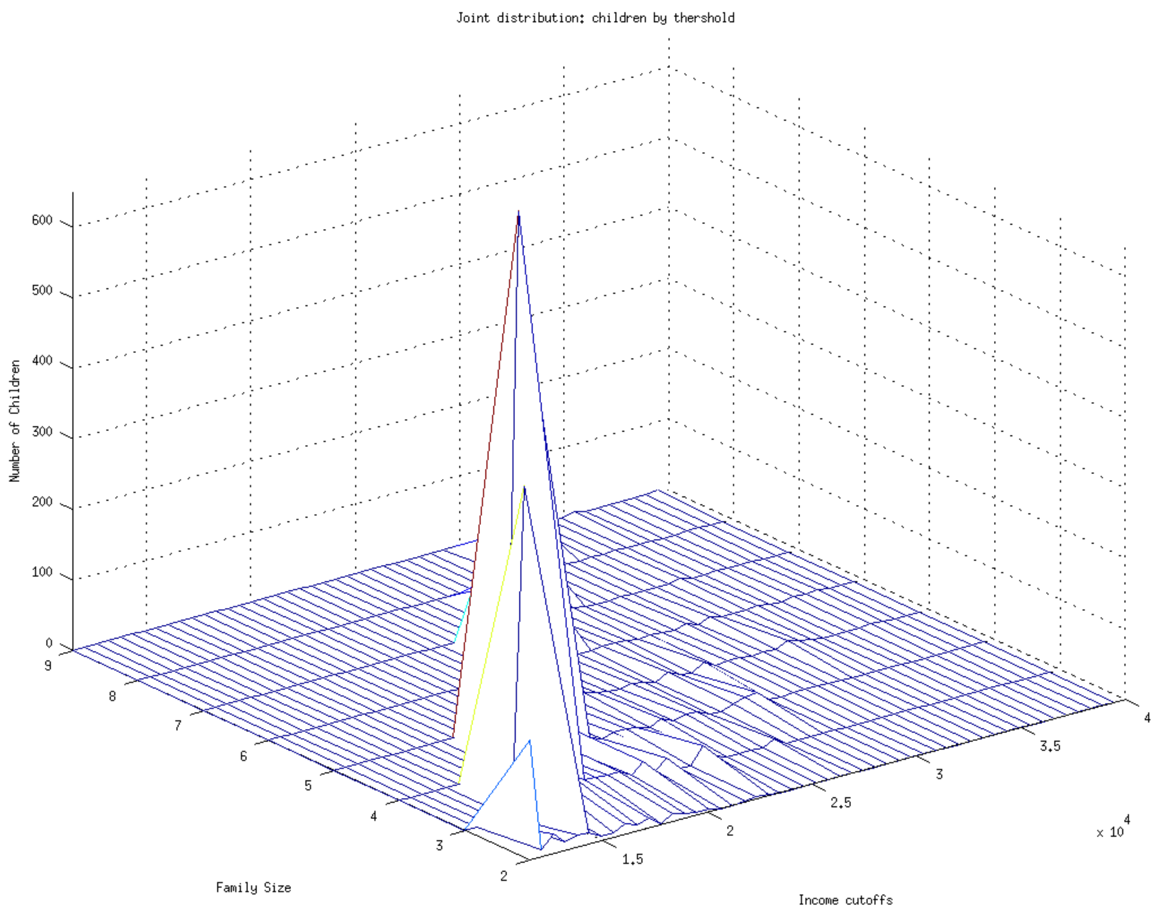

Figure D.1: Density of children around each threshold.
Note: Only children whose family income is at most twice the relevant threshold at age four are included. The sample used is further restricted to children living in families with at most nine members and with a cutoff at most of $40,000.

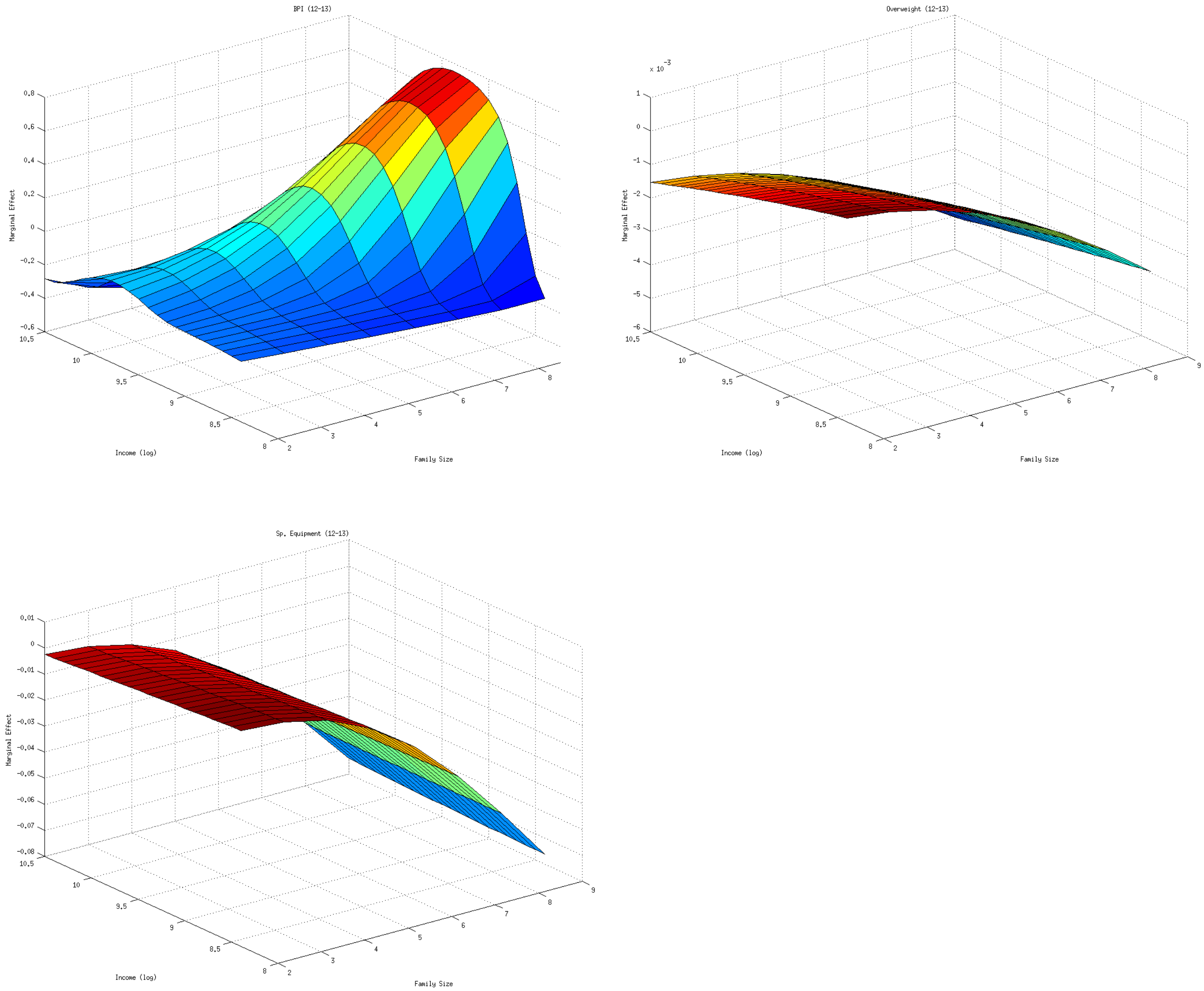

Figure D.2: Ages 12-13.

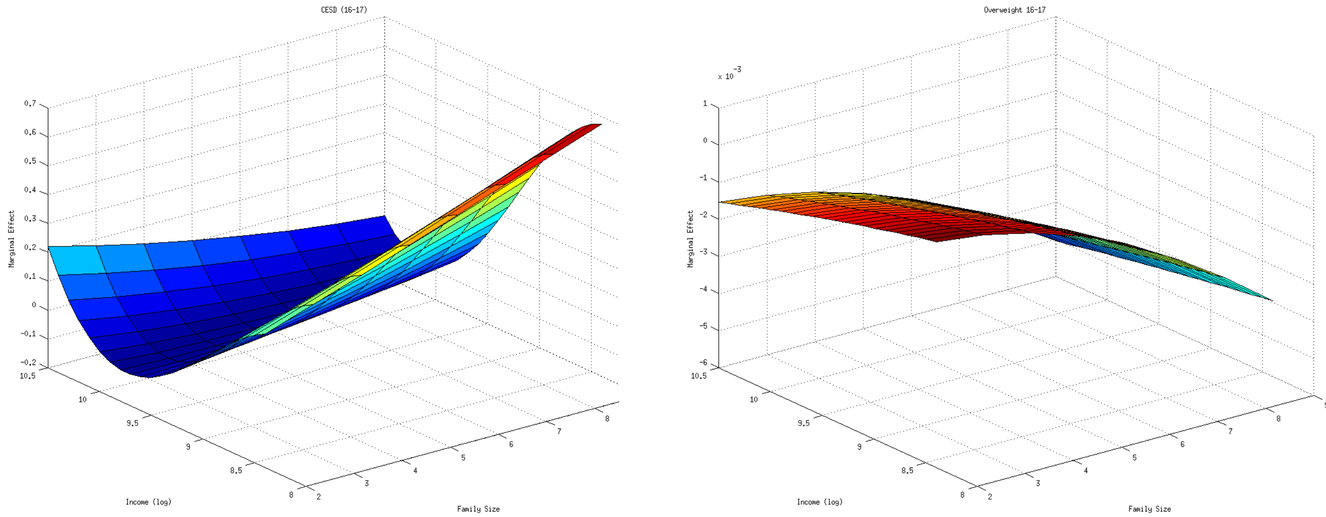

Figure D.3: Ages 16-17

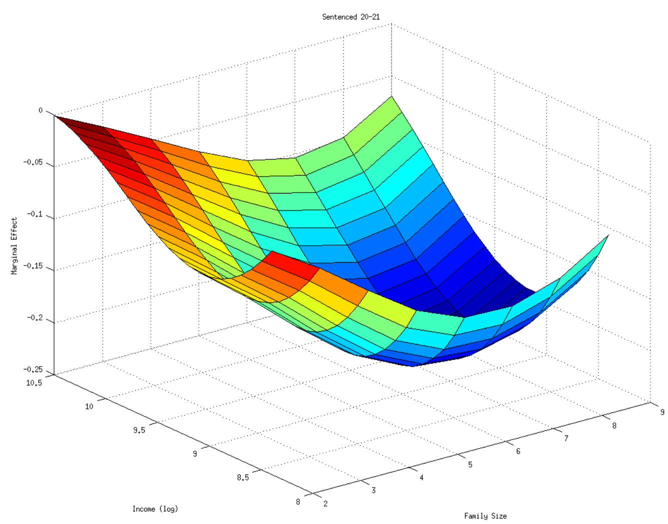

Figure D.4: Ever Sentenced (Ages 20-21)

E Calculation of adjusted p-values

This appendix describes our algorithm for calculating familywise adjusted p-values. It is based on Algorithms 4.1 and 4.2 of Romano and Wolf, 2005.

Let T be the sample size and $s = \{1, ..., S\}$ the number of hypothesis to test. Consider an individual test statistic $z_{T,s} = \hat{\beta_{T,s}} / \hat{\sigma_{T,s}}$, where $\hat{\beta_{T,s}}$ is the estimated coefficient on the Head Start eligibility indicator and $\hat{\sigma_{T,s}}$ is the estimated standard deviation of $\hat{\beta_{T,s}}$. Let X_T^* denote a data matrix generated by bootstrap, and $\hat{\beta^*_{T,s}}$ is the estimated coefficient on Head Start eligibility obtained using X_T^* and $\hat{\sigma^*_{T,s}}$ is the respective estimated standard deviation, such that $z^*_{T,s} = \hat{\beta^*_{T,s}} / \hat{\sigma^*_{T,s}}$. Let $\hat{d}_j$ be a data-dependent critical value obtained as in Algorithms 4.1 and 4.2 of Romano and Wolf, 2005.

We start by re-labelling the hypotheses in ascending order of their p-value. In the first step, we reject the null hypothesis that $\hat{\beta_{T,s}} = 0$ if $|z_{T,s}| > \hat{d}_1$, for all $s = \{1, ..., S\}$. If none of the null hypothesis is rejected, the process stops. If at least one is rejected, we remove it from the data and treat those left as the original data. We construct a new critical value with remaining data, which we denote by $\hat{d}_2$. Again, we reject the null hypothesis of $\hat{\beta_{T,s}} = 0$ if $|z_{T,s}| > \hat{d}_2$, for all $s = \{1, ..., S\}$ (excluding those hypothesis rejected in the first step). The process stops when no hypothesis is rejected.

F Mechanisms

What can be driving the effects we find on behaviors and health? The parameter we identify is the total effect of the program, inclusive of parental reaction to it, who could reinforce or substitute away investments in their children. To understand to which extent parents reaction to Head Start could explain, at least partially, the sustained effects we find we use the panel structure of CNLSY and present reduced form estimates using several measures of parental investments as dependent variables. In table F.1 in this Appendix we use three indices of quality of home environment, the HOME score, and two of its subcomponents (Cognitive Stimulation and Emotional Support), and one individual measure included in these indices (the number of books the child has). Besides the usual set of controls, the estimates control for child unobserved heterogeneity through child fixed effects. The results suggest that there is a complementarity between Head Start and parental investment, through increased cognitive stimulation (see Panel A, column 2).

The estimation of model (2) only requires that we are able to observe the family's income at age four along with the determinants of eligibility described in Section 3. Therefore, we are able to use other data sets in which we measure of family's eligibility status to understand whether other family's behaviors change when a child becomes eligible, for example, in terms of parental use of time in activities with children and expenditures. This is possible utilizing time use and consumer surveys, in particular, the American Time Use Survey (ATUS) and the Consumer Expenditure Survey (CEX).[62] Unfortunately, these data sets are repeated cross-sections and we unable to infer about the lagged eligibility status. Thus, we present suggestive evidence of the contemporaneous effect on families allocation of time and expenditures in tables F.2 and F.3 in this Appendix. Other limitation of this analysis is related to the period of time in which we are able to estimate model (2) in these data sets: (1) the time use data is only available for 2003-2009 and (2) we can use expenditure data only for the years of 1992-2008, as information about the state of residence is only available from 1992 onwards. Given that, as in CNLSY, most of the children are eligible due to poverty status, a family with children in Head Start age (3 to 5) is eligible if family income is below the Federal Poverty Guidelines.[63] Using a time use survey, it seems that Head Start is not associated to changes in parental use of time in educational activities or time spent with children table F.2. From CEX just eligible parents do not shift their expenditures towards children (table F.3). However, using a larger sample around cutoff for these data sets we find that (including all families with income at most twice the cutoff), there is an increase in baby sitting expenditures (no change in log consumption, which shows that families perceive Head Start as a transitory shock) and parents spend more time eating with children (this is consistent with finding of the HSIS on increasing of parental time spent with children).

[62]Table (F.5) compares the measure of income used from the three data sets used (NLSY79, CEX and ATUS).

[63]Unlike CNLSY, ATUS and CEX are household surveys. We refer to the notes of tables F.2 and F.3 for the selection of sample.

Table F.1: Mechanisms: Labor market outcomes and investments in human capital (ages 0-14).

	(1)	(2)	(3)	(4)
	HOME Score	Cognitive Stimulation	Emotional Support	Child's books
		Panel A: Males		
1[HS Eligible at 4]XAge0-2	0.163	0.172	0.136	-0.0974
	[0.102]	[0.122]	[0.116]	[0.113]
1[HS Eligible at 4]XAge6-9	0.00556	0.137	-0.105	0.0290
	[0.0807]	[0.0865]	[0.0922]	[0.0535]
1[HS Eligible at 4]XAge10-14	0.0889	0.248***	-0.0621	0.00722
	[0.0864]	[0.0906]	[0.0961]	[0.0596]
Observations	5,831	5,657	5,831	5,831
control mean	-0.538	-0.522	-0.387	3.210
Joint test: p-value	0.197	0.0370	0.214	0.718
Pre age 3: p-value	0.111	0.159	0.238	0.391
Age 6-9: p-value	0.945	0.114	0.255	0.588
Age 10-14: p-value	0.305	0.00652	0.519	0.904
		Panel B: Black Males		
1[HS Eligible at 4]XAge0-2	0.186	0.199	0.0854	0.0145
	[0.177]	[0.197]	[0.203]	[0.174]
1[HS Eligible at 4]XAge6-9	-0.00402	0.118	-0.0961	-0.0280
	[0.131]	[0.138]	[0.157]	[0.0952]
1[HS Eligible at 4]XAge10-14	0.0789	0.276*	-0.104	0.0408
	[0.144]	[0.154]	[0.161]	[0.103]
Observations	2,512	2,440	2,512	2,512
control mean	-0.661	-0.456	-0.693	3.155
Joint test: p-value	0.582	0.266	0.831	0.888
Pre age 3: p-value	0.296	0.314	0.674	0.933
Age 6-9: p-value	0.976	0.393	0.542	0.769
Age 10-14: p-value	0.585	0.0745	0.518	0.694
		Panel C: Non-Black Males		
1[HS Eligible at 4]XAge0-2	0.158	0.208	0.133	-0.158
	[0.130]	[0.143]	[0.152]	[0.155]
1[HS Eligible at 4]XAge6-9	-0.0265	0.129	-0.147	0.0289
	[0.0995]	[0.102]	[0.120]	[0.0698]
1[HS Eligible at 4]XAge10-14	0.0465	0.183*	-0.0651	-0.0201
	[0.110]	[0.104]	[0.126]	[0.0747]
Observations	3,319	3,217	3,319	3,319
control mean	-0.446	-0.570	-0.160	3.250
Joint test: p-value	0.352	0.285	0.230	0.598
Pre age 3: p-value	0.224	0.146	0.383	0.310
Age 6-9: p-value	0.790	0.206	0.220	0.679
Age 10-14: p-value	0.672	0.0798	0.607	0.788

Note: Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birthweight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Estimates include also child effects. "Control Mean" is the mean outcome among observations just above the cutoff, that is, at most 25% above the cutoff. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table F.2: Time use pattern around eligible threshold.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Sleeping	Eating	Leisure	Work		Child Care		
					Full	Basic	Teach	Play
Poor	5.520**	4.626	14.03	-4.152	-6.034	-4.646	-2.007	0.619
	(2.703)	(3.995)	(13.90)	(4.996)	(8.853)	(7.385)	(3.186)	(3.421)
Black	0.882	-5.117	-8.080	3.685	-13.22*	-6.953	-1.698	-4.566**
	(3.256)	(3.641)	(10.15)	(4.739)	(7.601)	(6.712)	(1.826)	(2.224)
PoorXBlack	-5.309	2.580	20.59	0.202	2.233	2.433	-0.568	0.368
	(4.274)	(6.028)	(17.42)	(7.127)	(11.99)	(10.54)	(2.482)	(3.019)
Observations	493	493	493	493	493	493	493	493
R-squared	0.175	0.252	0.184	0.144	0.259	0.214	0.187	0.168
Mean	61.85	22.38	63.76	10.63	37.83	26.93	4.662	6.241
SD	15.33	25.29	68.26	21.46	50.11	40.89	19.16	16.13
% non zeros	1	0.925	0.935	0.247	0.799	0.748	0.274	0.268
P-Value	0.0426	0.248	0.314	0.407	0.496	0.530	0.530	0.857

Note: Variables excluded from table include year indicators (for 2004-2009), quadratic on mother's age, cubic on family income and family size, an interaction between the 2, state FE, mother's education (high school graduate or college), marital status of mother, indicator for age of youngest child (3-5, 6-10), race dummy (for black mother). Robust standard errors in parenthesis clustered by state-year. Sample used from ATUS 2003-2009, families with youngest child 6 or less with mothers born between 1950 and 1970. Sample restricted to families whose total family income (taken from March Supplement of CPS) is at most twice times the poverty line. Leisure includes socializing, relaxing, and enjoyment of life and excludes child care. Child care basic includes grooming, rocking a child to sleep, feeding, breastfeeding. Dependent variables are measured in weekly hours. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table F.3: Consumption pattern around eligible threshold.

	(1)	(2)	(3)	(4)	(5)
	Exp. Educ.	Child Clothes	Daycare	Log Consumption	Alcohol and tobacco
Poor	50.62	17.27	107.7	-0.0459	26.69
	(104.8)	(32.05)	(76.06)	(0.0664)	(58.27)
Black	212.3	110.9***	146.3	0.0464	-121.1*
	(141.0)	(30.78)	(117.5)	(0.0784)	(62.93)
PoorXBlack	-239.3	20.10	-178.8	-0.0911	106.9
	(155.7)	(54.69)	(134.2)	(0.0868)	(92.38)
Observations	1,320	1,320	1,320	1,320	1,319
R-squared	0.134	0.158	0.142	0.241	0.145
Mean	321.5	194.2	232.9	9.026	300.9
SD	812.2	287.8	640.2	0.626	549.7
P-Value	0.629	0.590	0.158	0.490	0.647

Note: Variables excluded from table include year indicators (for 1993-2008), quadratic on mother's age, cubic on family income and family size, an interaction between the 2, state FE, marital status of mother, indicator for age of youngest child, race dummy (for black/native mother), education of mother (high school graduate, some college). Robust standard errors in parenthesis clustered by state-year. Sample used from CEX includes years of 1992-2008 and families whose youngest child is 5 or less and mothers born between 1950 and 1970. To comparison with CNLSY sample is restricted to families whose total family income (before taxes) between 15-185% of the poverty line. Dependent variables are measured in levels, except nondurable consumption. Means and SD in table refer to annual expenditures. Expenditures in education include school books, school and college tuition, daycare and baby sitting. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table F.4: Families with more than one age-eligible child.

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	All		Black Families		Non-Black Families	
	OLS	Family FE	OLS	Family FE	OLS	Family FE
1[HS Eligible at 4]	0.137*	0.00609	0.243	0.0762	0.0507	-0.144
	[0.0804]	[0.183]	[0.165]	[0.287]	[0.112]	[0.227]
1[HS Eligible at 4]xMale	0.158**	0.0692	0.0953	-0.0685	0.269***	0.156
	[0.0709]	[0.138]	[0.132]	[0.378]	[0.103]	[0.185]
Male	-0.0504	-0.0513	-0.0420	0.0490	-0.0681	-0.0628
	[0.0506]	[0.114]	[0.118]	[0.338]	[0.0595]	[0.106]
Observations	673	673	300	300	373	373

Note: OLS estimates of participation in Head Start on determinants of eligibility at age 4. The sample includes only families with at least two children and with mixed gender composition. Controls excluded from table include cubic in log family income and family size at age 4, an interaction between these two variables, cubic in log of average family income and family size for ages 0-2, the interaction between these two variables, cubic on child's birth weight, a dummy indicating the presence of a father figure in the household at age 4, race and age dummies, and dummies for year and state of residence at age 4. Robust standard errors are reported in brackets clustered at state-year at age four level. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table F.5: Income measure in NLSY, CEX and ATUS.

	NLSY	CEX	ATUS
1	Income received by mother and spouse in the past calendar year from wages, salary, commissions, or tips from all jobs, before deductions for taxes.	Amount of wage and salary income before deductions received by all CU members in past 12 months.	Hhld income from wages, salary, Armed Forces pay, commissions, tips, piece-rate payments, and cash bonuses earned, before tax deductions.
2	Income received by mother and spouse from own farm or nonfarm business, partnership or professional practice (after expenses).	Income or loss from own farm and nonfarm business, partnership or professional practice received by all CU members.	Hhld income from self employment and farm income.
3	Amount received by the mother and spouse from unemployment insurance.	Income from unemployment compensation received by ALL CU members.	Hhld income from unemployment insurance
4	Amount of child support received by the mother and spouse.	Income from child support payments in other than a lump sum amount received by ALL CU members.	Hhld income from child support.
5	Income from alimony or child support received by the mother from someone living outside the household.	Total amount of income from regular contributions from alimony and other sources such as from persons outside the CU received by ALL CU members.	Household payments from alimony.
6	Amount received from AFDC/TANF, SSI/other public assistance income received by mother or spouse.	Amount of Supplemental Security Income from all sources received by all CU members.	Hhld income from Supplemental Security Income.
7		Income from public assistance or welfare including money received from job training grants such as Job Corps received by ALL CU members.	Hhld income from Public Assistance.
8	Amount of other veteran benefits, worker compensation or disability payments received by the respondent (or spouse).	Income from workers compensation or veterans benefits, including education benefits, but excluding military retirement, received by ALL CU members.	Household income from worker's compensation and disability income.
9	Income received by mother/spouse from other sources. Includes any money from interest on savings, payments from social security, net rental income, or any other regular sources of income.	Amount of Social Security and Railroad Retirement income prior to deductions for medical insurance and Medicare received by all CU members.	Household income from Social Security and retirement income.
10		Income from pensions or annuities from private companies, military, Government, IRA, or Keogh, income from interest on savings accounts or bonds received by ALL CU members.	Household income from dividends, interests and rents.
11	Amount received by the mother or husband from persons living outside household (if living in DU), outside her home in city of permanent residence (if living in Dorm, fraternity or sorority) or not living with respondent (if in military).	Amount of other money income including money received from cash scholarships and fellowships, stipends not based on working, or from the care of foster children received by ALL CU members.	Household income from other sources.
12		Amount of regular income from interests on savings accounts or bonds, dividends, royalties, estates, or trusts received by ALL CU members.	
13	Amount of educational benefits received by mother and spouse in last calendar year.		Household income from veteran payments, education benefits and survivor's benefits.
14	Income received by the mother from adults that also live in the household and are related to her (excluding spouse and children).		
		Relevant Time Unit	
	Last calendar year.	Past 12 months.	Last calendar year.

Note: Income component for ATUS are obtained by merging this data with the March extracts of CPS (Current Population Survey). CU stands for consumer unit in CEX.