Gender Differences in the Benefits of an Influential Early Childhood Program*

Jorge Luis García
Department of Economics
The University of Chicago

James J. Heckman
American Bar Foundation
Department of Economics
The University of Chicago

Anna L. Ziff
Center for the Economics of
Human Development
The University of Chicago

First Draft: January 5, 2016
This Draft: May 11, 2017

*This research was supported in part by previous grants from the Buffett Early Childhood Fund, the Pritzker Children's Initiative, NICHD R01HD054702 and NIA R01AG042390, and current grants from the Robert Wood Johnson Foundation's Policies for Action program, NICHD R37HD065072 and the American Bar Foundation. The views expressed in this paper are solely those of the authors and do not necessarily represent those of the funders or the official views of the National Institutes of Health. The authors wish to thank Frances Campbell, Craig and Sharon Ramey, Margaret Burchinal, Carrie Bynum, Elizabeth Gunn, and the staff of the Frank Porter Graham Child Development Institute at the University of North Carolina Chapel Hill for the use of data and source materials from the Carolina Abecedarian Project and the Carolina Approach to Responsive Education. Years of partnership and collaboration have made this work possible. For information on childcare in North Carolina, we thank Richard Clifford and Sue Russell. We thank Ruby Zhang for excellent research assistance. Collaboration with Andrés Hojman, Yu Kyung Koh, Sylvi Kuperman, Stefano Mosso, Rodrigo Pinto, Joshua Shea, and Jake Torcasso on related work has strengthened the analysis in this paper. For helpful comments on various versions of the paper, we thank Stéphane Bonhomme, Flávio Cunha, Steven Durlauf, David Figlio, Dana Goldman, Ganesh Karapakula, Sidharth Moktan, Tanya Rajan, Azeem Shaikh, Jeffrey Smith, Chris Taber, Matthew Tauzer, Ed Vytlacil, Jim Walker, and Matt Wiswall. We benefited from helpful comments received at the Leonard D. Schaeffer Center for Health Policy and Economics in December, 2016, and at the University of Wisconsin, February, 2017. The set of codes to replicate the computations in this paper are posted in a repository. Interested parties can request to download all the files at https://github.com/jorgelgarcia/abccare-cba. To replicate the results in this paper, contact any of the authors, who will put you in contact with the appropriate individuals to obtain access to restricted data.

Abstract

This paper estimates gender differences in life-cycle impacts across multiple domains of an influential enriched early childhood program targeted toward disadvantaged children that was evaluated by the method of random assignment. We assess the impacts of the program on promoting or alleviating population differences in outcomes by gender. For many outcomes, boys benefit relatively more from high-quality center childcare programs compared to low-quality programs. For them, home care, even in disadvantaged environments, is more beneficial than lower-quality center childcare for many outcomes. This phenomenon is not found for girls. We investigate the sources of the gender differentials in impacts.

Keywords: Gender differences, childcare, early childhood education, health, randomized trials, substitution bias
JEL codes: J13, I28, C93

Jorge Luis García
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
Phone: 773-449-0744
Email: jorgelgarcia@uchicago.edu

James J. Heckman
Department of Economics
University of Chicago
1126 East 59th Street
Chicago, IL 60637
Phone: 773-702-0634
Email: jjh@uchicago.edu

Anna L. Ziff
Center for the Economics of
Human Development
University of Chicago
1126 East 59th Street
Chicago, IL 60637
Phone: 734-277-7379
Email: aziff@uchicago.edu

1 Introduction

Differences in gender are central features of economic and social life. This paper investigates how participation in enriched early childhood programs differentially affects the lives of disadvantaged boys and girls, and whether it promotes or reduces gender gaps. We show that, on many outcomes, low-quality childcare programs harm boys and contribute to gender gaps. This analysis sounds a cautionary note for advocates of early childcare in any form. Quality matters and low-quality programs can cause harm.

There is a rich literature in psychology on the greater vulnerability of boys to adverse life conditions and the importance of fathers in the lives of boys.[1] As a group, girls mature earlier, are more resilient to adversity, and perform better in a variety of life tasks.[2] Less is known about effective strategies for reducing the vulnerability of boys to disadvantage.

Many studies have shown the benefits of early-life interventions for improving the skills of children, especially those from disadvantaged families.[3] Although several of these studies report effects by gender, most do not.[4] Pooling males and females ignores potentially large differences in treatment effects.

This paper investigates this issue using data from a randomized controlled trial of a prototypical intensive early childhood program that enriched the early lives of disadvantaged children. The data come from the Carolina Abecedarian Project (ABC) and its closely aligned sister program, the Carolina Approach to Responsive Education (CARE). These programs were conducted in Chapel Hill, North Carolina for a sample of children born between 1972 and 1980. In this paper, we refer to the combined programs by the acronym

[1]See Golding and Fitzgerald (2016).

[2]See Schore (2017) for an informative overview of this literature. Economists have contributed to this literature. See, e.g., Bertrand and Pan (2013), Kottelenberg and Lehrer (2014), and Autor et al. (2015).

[3]Currie (2011); Elango et al. (2016).

[4]We survey the literature in Appendix C.1. See Elango et al. (2016) for a discussion of the main findings from the literature on early childhood education. Not reporting gender differences is common. Some examples include Bernal and Keane (2011); Cascio and Schanzenbach (2013); Bitler et al. (2014); Kline and Walters (2016). There are some exceptions: Heckman et al. (2010); Campbell et al. (2014); García et al. (2017).

ABC/CARE. It is a template for many current and proposed early interventions.[5] It starts at 8 weeks of age and continues through age 5. Treatment and control subjects are followed through their mid 30s, with data collected on multiple dimensions of human development.

There are positive impacts of the program across the life cycle for both genders. However, there are substantial differences in impacts by gender across domains. The program differentially promotes the labor income, employment, and health of males and reduces their participation in crime. It differentially enhances the cognition, achievement, and educational attainment of girls. Boys placed in childcare benefit relatively more than girls from participation in high-quality center-based care compared to low-quality center-based care.

Our analysis sheds light on recent claims about the harm caused by enrolling children in childcare programs. In an influential analysis, Baker et al. (2008) show that participants in childcare manifest adverse behavioral outcomes. Kottelenberg and Lehrer (2014) localize their effects to boys.

However, all childcare programs are not alike. We produce evidence that high-quality childcare greatly benefits boys relative to low-quality childcare. Staying at home is a better option for them, especially if the family environment is relatively advantaged (e.g. the father is present). This effect is not found for girls. The program analyzed by Baker et al. (2015) and Kottelenberg and Lehrer (2014) is of relatively low quality compared to the enriched program we analyze.[6] There is no contradiction between the claim that low-quality programs impair child development, while high-quality programs do not. Boys are more vulnerable

[5]Programs inspired by ABC/CARE have been (and are currently being) launched around the world. Sparling (2010) and Ramey et al. (2014) list numerous programs based on the ABC/CARE approach. The programs are: IHDP—eight different cities around the U.S. (Spiker et al., 1997); Early Head Start and Head Start in the U.S. (Schneider and McDonald, 2007); John's Hopkins Cerebral Palsy Study in the U.S. (Sparling, 2010); Classroom Literacy Interventions and Outcomes (CLIO) study in the U.S. (Sparling, 2010); Massachusetts Family Child Care Study (Collins et al., 2010); Healthy Child Manitoba Evaluation (Healthy Child Manitoba, 2015); Abecedarian Approach within an Innovative Implementation Framework (Jensen and Nielsen, 2016); and Building a Bridge into Preschool in Remote Northern Territory Communities in Australia (Scull et al., 2015). Educare programs are also based on ABC/CARE (Educare, 2014; Yazejian and Bryant, 2012).

[6]The program they analyze was a modest payment (between 5 and 7 Canadian dollars a day starting in 1997) in the form of a voucher. The voucher component itself reduces received quality given an information problem by parents. The full cost of the program was $44 (2014) a day (Baker et al., 2005). This is less than two thirds of the cost of ABC/CARE, which was close to $75 (2014) a day (García et al., 2017).

and respond adversely to low-quality environments on many outcomes.

Unlike previous studies, we estimate treatment effects comparing treatment group outcomes to different control groups: home care or care in low-quality centers.[7,8] We find that home care is beneficial for boys compared to low-quality center childcare.

To preview our analysis, we present gender differences in outcomes in Table 1. We report the proportion of outcomes, by category, for which males outperform females (we have multiple outcome measures in each category which we explain in greater detail below). These proportions are invariant to the scales used for individual measures. Under the null hypothesis of no difference in treatment effects by gender, the proportion of outcomes favoring any gender should be 50%. We test this hypothesis for the control group and the treatment group separately to determine whether there is a baseline (control) gender difference and whether treatment affects this gender difference.

Pooling the two control groups, males have higher IQs, employment, enhanced parental labor income (through subsidized childcare), and are more likely to participate in criminal activity than females. They do better on an aggregate over all categories. Treatment reverses the gap in achievement between males and females found in the pooled control group. All achievement measures favor males in the pooled control group, but favor females in the treatment group. Education is another outcome category for which treatment reverses the gender gap between treatment and a pooled control group. Males have higher educational attainment in the pooled control group with 66.7% of the educational outcomes favoring males, although the result is not statistically significant. In the treatment group, however, only 11.1% of the educational outcomes favor males. Treatment reverses the gap for an aggregate across all outcome categories and narrows the gap for employment.

[7]Historical documentation, records, and evidence from knowledgeable individuals indicate that although the available alternative centers followed state and federal standards, they were of considerably lower quality than the ABC/CARE program. We document this in Appendix A.5.

[8]Previous studies presenting treatment effects of ABC and CARE include Ramey et al. (1985); Clarke and Campbell (1998); Campbell et al. (2001, 2002, 2008, 2014).

Table 1: Summary of Proportion of Outcomes Males > Females

Category	# Outcomes	Proportion		Difference
		Control	Treatment	Treatment − Control
IQ	15	0.733	0.533	-0.200
Achievement	12	**0.833**	**0.000**	**-0.833**
Social-Emotional	22	0.455	0.364	**-0.091**
Parenting	7	0.571	0.286	**-0.286**
Parental Labor Income	15	0.600	0.733	0.133
Education	9	0.667	**0.111**	**-0.556**
Employment	4	**1.000**	**0.750**	**-0.250**
Crime	4	**0.000**	**0.000**	**0.000**
Risky Behavior	5	0.400	**0.200**	**-0.200**
Health	22	0.545	0.545	0.000
Mental Health	11	**0.818**	0.545	**-0.273**
All	126	0.611	0.413	**-0.198**

Note: This table summarizes comparison of gender gaps across outcome categories by different groups. A bold proportion for the treatment or control group indicates that the proportion is statistically different than 50% at the 10% level. A bold difference between treatment and control indicates that the Wilcoxon signed-rank test comparing the control and treatment proportions, over 1,000 bootstraps, yields a p-value less than or equal to 0.10. The variables for each outcome category are listed in Appendix C.2. The inference procedure is described in detail in Appendix B.3.4. In summary, we bootstrap with replacement 1,000 times. For bootstrap $b \in [1, \ldots, 1,000]$ we compute the proportion of variables in each outcome category for which the males do better than the females. We do this separately for the treatment and control group, estimating the pair of proportions $(P_{T,b}, P_{C,b})$. Over the distribution of the pairs, we compute the signed-rank test. It compares the empirical distribution of the proportions in the treatment group to the empirical distribution of the proportions in the control group. We repeat this procedure for each outcome category.

Table 2 summarizes the gender gaps reported in Table 1, as well as those for different childcare environments for control-group children. In the control group, the proportion of outcomes for which males do better than females is higher than 50% for most of the categories, although not all are statistically significant. Exceptions include social-emotional skills, risky behavior, and crime, in which females surpass males.[9] Treatment reverses the gaps for achievement, education, employment, parental income, and overall outcomes. For controls who stay at home, females have slightly better health outcomes than males. The males who attend lower-quality childcare do not outperform females on any of the outcome categories associated with cognition, education, and induced parental labor income (through

[9]Females commit fewer crimes and report less risky behavior in comparison to males.

subsidy of childcare). These outcomes are concentrated relatively early in the life of a child, indicating an early-life disadvantage that enriched early childcare programs partially correct.

Table 2: Summary of Proportion of Outcomes Males > Females by Home Status

Category	(1) Control Group	(2) Control Group Stay at Home	(3) Control Group in Low-Quality Childcare	(4) Treatment Group
IQ	✓	✓*	×	✓
Achievement	✓*	✓*	×	×*
Social-Emotional	×	✓	×*	×*
Parenting	✓	✓*	×	×
Parental income	✓	✓*	×	✓
Education	✓	✓*	×	×*
Employment	✓*	✓*	=	✓*
Crime	×*	=	×*	×*
Risky Behavior	×	×*	✓	×*
Health	✓	×	=	✓
Mental Health	✓*	✓*	✓	✓
All	✓	✓*	×	×

Note: This table summarizes comparison of gender gaps across outcome categories by different groups. A checkmark indicates that the proportion of outcomes in the corresponding category is larger than 50%, meaning that males outperform females. A checkmark with an asterisk indicates that the proportion is significantly larger than 50% at the 10% level. A cross indicates that the proportion of outcomes is smaller than 50%. A cross with an asterisk indicates that the proportion is significantly smaller than 50% at the 10% level. An equals sign indicates that the proportion is exactly 50%. Column (1) is the difference between males and females in the full control group. Column (2) is the difference between males and females in the control group only considering those who stayed at home. Column (3) is the difference between males and females in the control group only considering those who attended alternative childcare. Column (4) is the difference between males and females in the treatment group.

This paper unfolds in the following way. Section 2 describes the experimental data we analyze and its special features. It documents that a considerable proportion of the control-group subjects attend lower-quality childcare compared to treatment subjects. Section 3 defines the treatment effects we estimate and how we summarize them. Section 4 reports the treatment effects overall and by gender and establishes the existence of sharp gender differences for many categories of outcomes. Section 5 discusses the sources of these differences. Section 6 concludes.

2 Data

We analyze a combined sample of the two closely related programs, ABC and CARE. Table 3 summarizes their characteristics. Both interventions were implemented by researchers at the Frank Porter Graham Center at the University of North Carolina Chapel Hill, and targeted children from disadvantaged families in the Chapel Hill area. ABC had four cohorts born between 1972 and 1977, and CARE had two cohorts born between 1978 and 1980. Eligibility was determined on the basis of the High Risk Index (HRI) developed for ABC and adapted for CARE.[10] Components of the HRI include father's presence, parental employment, and participation in welfare.[11]

Both interventions involved intensive center-based care for subjects in the treatment group starting at 8 weeks and continuing until age 5 before the children started kindergarten. Treatment-group subjects also received daily health screenings and diapers and formula until 6 months. Control-group families received diapers and formula as well for the same period of time.[12] After school entry up until age 8, there was an additional component of treatment in which home visitors worked with children and their parents to tutor the children and to encourage families to be involved in their child's schooling.[13] We do not analyze this part of the treatment because previous work has found no statistically significant treatment effects of it.[14]

[10] See Campbell et al. (2008).

[11] See Appendix A.2 for the full list of the determinants of HRI (Ramey and Smith, 1977; Wasik et al., 1990; Ramey and Campbell, 1991).

[12] Wasik et al. (1990).

[13] In ABC, treatment status of this component was randomized. In CARE, all the subjects who received center-based care also received this school-age component.

[14] Campbell et al. (2002, 2014).

Table 3: Overview of the ABC and CARE Programs

Site	Chapel Hill, North Carolina
Cohorts	4 (ABC), 2 (CARE)
N	58 treatment, 56 control (ABC)
	17 treatment, 23 control (CARE)
Eligibility	HRI > 11
	Biologically healthy
Treatment years	1972–1981 (ABC), 1977–1985 (CARE)
Treatment duration	5 years
Home visits	2.5–2.7 per month (CARE)
Center care	50 weeks per year
	30–45 hours per week
Other treatment components	Formula until 6 months
	Diapers until 6 months
	Health check-ups
	Medical care
	Parenting instruction
	Counseling
	Transportation to center
Control-group incentives	Formula until 6 months
	Diapers until 6 months
	Health check-ups until 1 year (ABC, cohort 1)
Adult-child ratio	1:3–1:6
Teacher requirements	High school through masters
	Experience with children
Specialists	Physician, nurse, social worker

Note: Characteristics that do not specify ABC or CARE were present in both. Biologically healthy includes lack of serious illness, including mental retardation. HRI is the High Risk Index.
Sources: Ramey et al. (1976); Ramey and Smith (1977); Ramey et al. (1985); Wasik et al. (1990); Ramey and Campbell (1991).

The pedagogical approach focused on developing language, cognition, and social-emotional skills. During ABC and CARE, the *Learning Games* curriculum was developed and refined.[15] Other curricula emphasized child-led learning of skills important for future learning.[16]

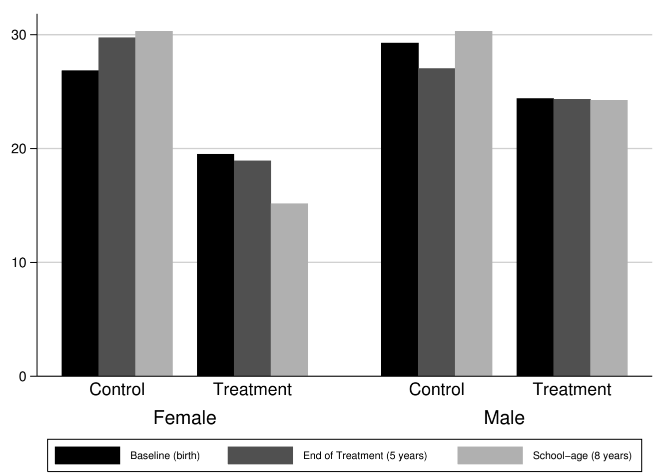
CARE included an additional arm of treatment. Besides the services just described, those in the treatment group also received home visiting from birth to age 5. Home visit-

[15] Sparling and Lewis (1979).
[16] Conti et al. (2016).

ing consisted of biweekly visits focusing on parental problem-solving skills. There was, in addition, an experimental group that received only the home visiting component, but not center-based care.[17] In light of previous analyses, we drop this last group from our analysis. The home visiting component had very weak estimated effects.[18] These analyses justify merging the treatment groups of ABC and CARE, even though that of CARE received the additional home-visiting component.[19] We henceforth analyze the samples so generated as coming from a single ABC/CARE program.

Figure 1 presents one aspect of family disadvantage. The figure shows the presence of the father at home and its evolution over the early life of the child. More fathers are present among the controls. The level in the treatment group is lower for both genders. The decline in the level is less precipitous for males. For both treatments and controls, the proportion of fathers present, conditional on being present at birth, remains roughly stable.[20]

Figure 1: Father Present Over Time



Note: The bars represent the proportion of subjects with fathers present. The bars after baseline are conditional on the father being at home at baseline.

[17]Wasik et al. (1990).

[18]Campbell et al. (2014) test and do not reject the hypothesis of no treatment effects for this additional component of CARE.

[19]García et al. (2016).

[20]For a more detailed description of the sample's disadvantage, see Appendix A.2.

2.1 The Randomization Protocol and its Compromises

Randomization for ABC/CARE was conducted on child pairs matched on family background. Siblings and twins were jointly randomized into either treatment or control groups.[21] Pairing was based on the High Risk Index, as well as maternal education, maternal age, and gender of the subject.[22] ABC collected an initial sample of 121 subjects. We characterize each missing observation in Appendix A.3.[23] We conduct the same analysis for the CARE sample. 22 subjects in ABC did not stay in the program through age 5. Dropouts are evenly balanced across treatments and controls. They are primarily related to the health of the child and the mobility of families rather than any dissatisfaction with the program.[24]

2.2 Control Group Substitution

In ABC/CARE, many control-group subjects (but no treatment-group subjects) attended alternative center childcare.[25] The figure is 75% for ABC and 74% for CARE. This creates both a problem (substitution bias[26]) and an opportunity (we can compare the effect of an enriched treatment against home and low-quality center environments separately).

Figure 2a shows the cumulative distribution of the proportion of time in the first five years that control subjects were enrolled in alternatives. Figure 2b shows the dynamics of enrollment. Those who enroll generally stay enrolled. As control-group children age, they are more likely to enter childcare (see Appendix A.5).

Children in the control group who are enrolled in alternative early childcare programs

[21]For siblings, this occurred when two siblings were close enough in age such that both of them were eligible for the program.

[22]We do not know the original pairs.

[23]In Appendix D.3, we document that our estimates are robust when we adjust for missing data using standard weighting (matching) methods described in Appendix B.2.

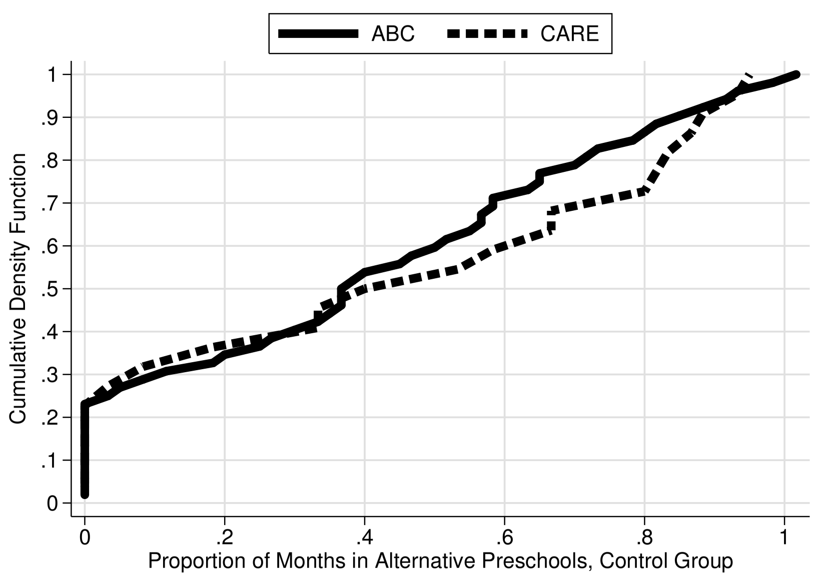
[24]The 22 dropouts include four children who died, four children who left the study because their parents moved, and two children who were diagnosed as developmentally delayed. Details are in Table A.3. Everyone offered the program was randomized to either treatment or control. All eligible families agreed to participate. Dropping out occurs *after* randomization.

[25]See Heckman et al. (2000) on the issue of substitution bias in social experiments.

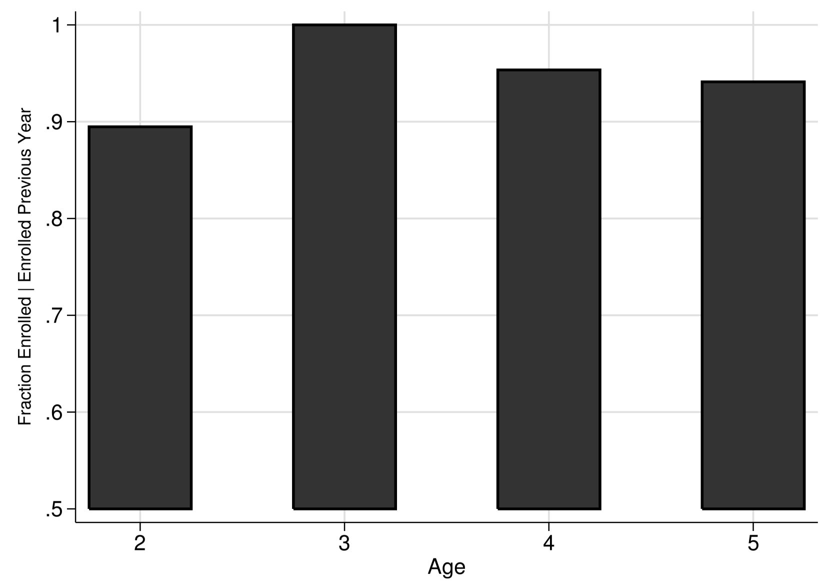
[26]See Heckman (1992), Heckman et al. (2000), and Kline and Walters (2016).

Figure 2: Control Substitution Characteristics, ABC/CARE Control Group

(a) Cumulative Enrollment



(b) Enrollment Dynamics



Note: Panel (a) displays the cumulative distribution function of enrollment in alternatives. Panel (b) displays the fraction of ABC/CARE control-group children enrolled in alternatives, conditional on being enrolled in the previous age (at least one month).

are less economically disadvantaged at baseline compared to children who stay at home. Disadvantage is measured by maternal education, maternal IQ, Apgar scores, and the High Risk Index defining ABC/CARE eligibility. Children who attend low-quality alternative care have fewer siblings. On average, they are children of mothers who are more likely to be working at baseline.[27] Parents of girls are much more likely to use alternative center childcare if assigned to the control group.[28] If the subject is a boy, the father is more likely to be present.

While most of the alternative childcare centers received federal subsidies and were subject to the federal regulations of the era, they were relatively low-quality compared to ABC/CARE.[29,30] The access of control-group children to alternative programs affects the interpretation of estimated treatment effects, as we discuss next.

3 Parameters Estimated

Random assignment to treatment does not guarantee that conventional treatment effects answer policy-relevant questions. In this paper, we define and estimate three parameters that address different policy questions.

Life cycles consist of A discrete periods. Treatment occurs in the first $\bar{a}$ periods of life $[1, \ldots, \bar{a}]$. We have data through age $a^* > \bar{a}$. We lack follow-up data on the remainder of life $(a^*, \ldots, A]$. We define three indicator variables. $W = 1$ indicates that the parents referred to the program participate in the randomization protocol, $W = 0$ indicates otherwise. R indicates randomization into the treatment group ($R = 1$) or to the control group ($R = 0$). D indicates compliance in the initial randomization protocol, i.e., $D = R$ implies full compliance

[27] Statistically significant at 10%.

[28] See Table A.5 for tests of differences across these variables between children in the control group who attended and who did not attend alternative childcare.

[29] Appendix A.5.1 discusses the federal standards of that day. See Department of Health, Education, and Welfare (1968); North Carolina General Assembly (1971); Ramey et al. (1977); Ramey and Campbell (1979); Ramey et al. (1982); Burchinal et al. (1997).

[30] When we compare ABC/CARE treatment to these alternatives, ABC/CARE has substantial treatment effects. Further, as we argue below, parents perceived that ABC/CARE was superior to the alternatives.

in the initial randomization protocol.

Individuals are eligible to participate in the program if baseline background variables $\boldsymbol{B} \in \mathcal{B}_0$. $\mathcal{B}_0$ is the set of scores on the High Risk Index that determines program eligibility. As it turns out, in the ABC/CARE study, all of the eligible persons given the option to participate choose to do so ($W = 1$, and $D = R$). There are very few dropouts. *Ex ante*, parents perceived that ABC/CARE was superior to other childcare alternatives. Thus, we can safely interpret the treatment effects generated by the experiment as average treatment effects for the population for which $\boldsymbol{B} \in \mathcal{B}_0$ and not just treatment effects for the treated (**TOT**).[31]

Let $\boldsymbol{Y}_a^1$ be the outcome vector at age a for the treated. $\boldsymbol{Y}_a^0$ is the age-a outcome vector for the controls. In principle, life-cycle outcomes for the treated and controls can depend on the exposures to various alternative preschools at each age. It would be desirable to estimate treatment effects for each possible exposure but our samples are too small to make credible estimates for very detailed levels of exposure.

All treatment-group children have the same exposure. We simplify the analysis of the controls by creating two categories. "H" indicates that the control-group child is in home care throughout the entire length of the program. "C" indicates that a control-group child is in alternative center childcare for any amount of time.[32] We test the sensitivity of our estimates to the choice of different categorizations in our empirical analysis in Appendix D.12.1.

We thus compress a complex reality into two counterfactual outcome states at age a for

[31] All providers of health care and social services (referral agencies) in the area of the ABC/CARE study were informed of the programs. They referred mothers whom they considered disadvantaged. Eligibility was corroborated before randomization. Our conversations with the program staff indicate that the encouragement from the referral agencies was such that most referred mothers attended and agreed to participate in the initial randomization (Ramey et al., 2012).

[32] This assumption is consistent with Figure 2b. Once parents decide to enroll their children in alternative childcare arrangements, the children stay enrolled up to age 5.

control-group subjects:

$$\boldsymbol{Y}_{a,H}^{0} \quad : \quad \textbf{Subject received home care exclusively}$$

$$\boldsymbol{Y}_{a,C}^{0} \quad : \quad \textbf{Subject received some alternative childcare}.$$

We define V as a dummy variable indicating participation by control-group children in an alternative childcare. $V = 0$ denotes staying at home. The outcome when a child is in control status is

$$\boldsymbol{Y}_a^0 := (1 - V)\,\boldsymbol{Y}_{a,H}^0 + (V)\,\boldsymbol{Y}_{a,C}^0. \tag{1}$$

One parameter of interest addresses the question: what is the effect of the program as implemented? This is the effect of the program compared to the next best alternative as perceived by the parents (or the relevant decision maker) and is defined by

$$\boldsymbol{\Delta}_a := \mathbb{E}\left[\boldsymbol{Y}_a^1 - \boldsymbol{Y}_a^0 | W = 1\right] = \mathbb{E}\left[\boldsymbol{Y}_a^1 - \boldsymbol{Y}_a^0 | \boldsymbol{B} \in \mathcal{B}_0\right], \tag{2}$$

where the second equality follows because everyone eligible wants to participate in the program. For the sample of eligible persons, this parameter addresses the effectiveness of the program relative to the quality of all alternatives available when the program was implemented, including staying at home. It is the parameter intended to be estimated by LATE.

It is fruitful to assess the effectiveness of the program with respect to a counterfactual world in which the child stays at home full time. The associated causal parameter for those who would choose to keep the child at home is:

$$\boldsymbol{\Delta}_a\,(V = 0) := \mathbb{E}\left[\boldsymbol{Y}_a^1 - \boldsymbol{Y}_a^0 | V = 0, W = 1\right] := \mathbb{E}\left[\boldsymbol{Y}_a^1 - \boldsymbol{Y}_{a,H}^0 | V = 0, \boldsymbol{B} \in \mathcal{B}_0\right].^{33} \tag{3}$$

It is also useful to assess the average effectiveness of a program relative to attendance in an

[33] Appendix D.12.1 displays results with alternative definitions of V (i.e., different thresholds define if a child attended alternative childcare). The results are robust to the various definitions. What matters is whether any out-of-home child care is being used ($V > 0$), and not the specific value of V.

alternative childcare center for those who would choose an alternative:

$$\boldsymbol{\Delta}_a (V = 1) := \mathbb{E}\left[\boldsymbol{Y}_a^1 - \boldsymbol{Y}_a^0 | V = 1, W = 1\right] := \mathbb{E}\left[\boldsymbol{Y}_a^1 - \boldsymbol{Y}_{a,C}^0 | V = 1, \boldsymbol{B} \in \mathcal{B}_0\right]. \tag{4}$$

Random assignment to treatment does not directly identify (3) or (4). Econometric methods are required to identify these parameters. We primarily rely on matching (conditions on observables) to control for selection into home or low-quality childcare by the control group. We report results from alternative strategies in Appendix E. We characterize the determinants of choices and our strategy for controlling for selection into "H" and "C" when we discuss the empirical results in Section 4.

3.1 Summarizing Multiple Treatment Effects

ABC/CARE has rich longitudinal data on multiple outcomes over multiple periods of the life cycle. Summarizing these effects in an interpretable way is challenging.[34] Simpler, more digestible summary measures are useful for understanding our main findings. To construct these, we use combining functions that count the proportion of treatment effects that are positive by different categories of outcomes. In another paper (García et al., 2017), we monetize outcomes and estimate rates of return and cost/benefit rates. This requires making an additional layer of assumptions to extrapolate lifetime benefits, which we avoid making in this paper.

Consider a block of N_l outcomes indexed by set $Q_l = \{1, \ldots, N_l\}$. Let $j \in Q_l$ be a particular outcome within block l. Associated with it is a mean treatment effect

$$\Delta_{j,a} := \mathbb{E}\left[Y_{j,a}^1 - Y_{j,a}^0 | \boldsymbol{B} \in \mathcal{B}_0\right], j \in Q_l. \tag{5}$$

[34] Appendix D presents step-down p-values for the blocks of outcomes that are used in the cost-benefit analysis of García et al. (2017) (Lehmann and Romano, 2005 and Romano and Shaikh, 2006). We follow the algorithm in Romano and Wolf (2016).

We assume that outcomes can be ordered so that $\Delta_{j,a} > 0$ is beneficial.[35] We summarize the estimated effects of the program on outcomes within the block by the count of positive impacts within block l:

$$C_l = \sum_{j=1}^{N_l} 1(\hat{\Delta}_{j,a} > 0). \tag{6}$$

The proportion of beneficial outcomes in block l is C_l/N_l.[36]

Let $\mathcal{L}$ be the set of blocks. Under the null hypothesis of no treatment effects for all $j \in Q_l, l \in \mathcal{L}$, and assuming the validity of asymptotic approximations, C_l/N_l should be centered around $\frac{1}{2}$. We bootstrap to obtain p-values for the null for each block and over all blocks. This procedure accounts for dependence in unobservables across outcomes. It also accounts for model pretesting.[37]

We also count the beneficial treatment effects that are statistically significant in the sets of outcomes across each of the groups indexed by the set Q_l. Using a 10% significance level, on average 10% of all outcomes should be "significant" at the 10% level even if there is no treatment effect of the program. We provide evidence against both null hypotheses.[38] Combining counts across all blocks enables us to avoid (i) arbitrarily picking outcomes that have statistically significant effects—"cherry picking"; or (ii) arbitrarily selecting blocks of outcomes to correct the p-values when accounting for multiple hypothesis testing.[39,40]

[35] All but 5% of the outcomes we study can be ranked in this fashion. See Appendix D for a discussion.

[36] In our empirical application we consider all the outcomes as a block, and then different blocks grouped according to common categories—e.g., skills, health, crime.

[37] Bootstrapping allows us to account for dependence across outcomes in a general way. We adjust for pretesting by estimating a series of alternative models and computing the standard errors that account for doing so.

[38] In this case, we perform a "double bootstrap" procedure to first determine significant treatment effects at 10% level and then calculate the standard error of the count.

[39] We present p-values for these hypotheses and a number of combining functions by outcome category in Appendix D.

[40] In Appendix D we present yet another alternative. We calculate a "latent" outcome out of the set of outcomes within a block and perform inference on this latent outcome. This analysis also points to beneficial effects of the program.

4 Estimates and Tests

This section reports estimates of treatment effects by gender for the main outcomes, tests of equality of effects by gender by outcome, and combining functions. These estimates show that the program treatment effects are strong for both males and females. Treatment effects depend on the counterfactual against which they are benchmarked (home or low-quality center childcare). Treatment effects differ sharply by gender.[41]

4.1 Estimated Treatment Effects

Tables 4 and 5 present estimates of a representative selection of treatment effects for males and females respectively. The full list of treatment effects is reported in Appendix D. Column (1) of each table gives sample mean differences in outcomes between treatment and control groups. Column (2) adjusts the differences for attrition and controls for background variables. Both are estimates of the parameter defined in equation (2), but with different conditioning sets. Column (3) shows the mean difference between the full treatment-group and the control-group subjects who stay at home. Column (4) gives matching estimates for the parameter defined in equation (3).[42] Column (5) reports estimates of the parameter defined by equation (4): mean differences between the treatment-group and control-group children who attended alternatives. Column (6) gives matching estimates for the parameter of equation (4).[43]

The results for females show that ABC/CARE has substantial effects on education when

[41] See Appendix B.3 for the exact procedure we use to conduct inference.

[42] In Appendix D.1.1, we provide details on: (i) the kernel matching estimator that we use; (ii) the matching variables that we use; and (iii) a sensitivity analysis to these matching variables.

[43] We want to clarify one set of variables: those related to parental income. There are some considerations that contribute to the higher parental income at the later ages. One consideration is that the variables are asked differently, with the question at age 3.5 using categories. These categories end at greater than $14,000, which truncates the distribution. Although we fill this out using a tobit model, we cannot fully impute these higher incomes. The other consideration is that the mothers are very young, with many being enrolled in school at the child's birth. Once the children are 8, more mothers have entered the workforce which in turn increases the reported income.

comparing outcomes of the treatment-group subjects to those from the next best alternative. High school graduation increases between 13 and 25 percentage points, depending on which estimate we consider; college graduation increases 13 percentage points; and the average years of schooling increase between 1.8 and 2.1 years. Employment at age 30 increases between 8 and 13 percentage points. ABC/CARE has substantial impacts on human capital accumulation and employment. The results strengthen when we compare treatment outcomes with outcomes for those who remain at home.

The results for males are different from those for females. Treatment has substantial effects when compared to the next best alternative. The effects are positive for a variety of health indicators, including drug use and hypertension. The effects on employment and labor income are sizable. The increase in employment at age 30 ranges from 11 to 19 percentage points. Labor income at age 30 increases between $19,000 and $24,000 (2014) after treatment. The effects strengthen when comparing treatment to low-quality childcare. Separation from the mother and being placed in relatively low-quality childcare centers has more deleterious consequences for males than for females.[44]

The results hold when using other definitions of control substitution, which vary by exposure length. They remain statistically significant or are borderline statistically insignificant when computing two-tailed p-values (see Appendix D.12.2). When using step-down p-values (see Appendix D for the full set of step-down p-values), the health results remain statistically significant for males, when compared to the full control group and when compared to alternative center childcare. The education results remain statistically significant for females, when compared to the full control group and when compared to those who stay at home.

The estimated effects for females and males in Columns (3) and (5) in their respective tables are not based on adjustments for baseline covariates (matching); the estimates in (4) and (6) are. In Appendix D.1, we justify our choice of matching variables. We conduct a

[44] This is consistent with the evidence in Kottelenberg and Lehrer (2014) and Baker et al. (2015) and the analysis of Golding and Fitzgerald (2016).

Table 4: Treatment Effects on Selected Outcomes, Males

Category	Variable	Age	$\bar{Y}_C$	(1)	(2)	(3)	(4)	(5)	(6)
Parental Income	Parental Labor Income	3.5	13,505	1,036	494	73.862	1,462	123	690
				(0.374)	(0.411)	(0.474)	(0.390)	(0.479)	(0.417)
				[0.892]	[0.927]	[1.000]	[0.857]	[0.955]	[0.891]
		12	23,868	7,085	9,625	18,050	12,639	6,620	5,383
				(0.092)	**(0.020)**	**(0.038)**	**(0.074)**	**(0.098)**	(0.139)
				[0.318]	[0.192]	[0.206]	[0.425]	[0.472]	[0.564]
		15	22,985	8,488	4,495	5,540	4,805	2,885	4,345
				(0.071)	(0.221)	(0.243)	(0.264)	(0.354)	(0.296)
				[0.288]	[0.778]	[0.825]	[0.855]	[0.911]	[0.839]
		21	21,585	12,732	8,809	122	-933	10,784	10,283
				(0.005)	**(0.098)**	(0.448)	(0.456)	**(0.056)**	**(0.041)**
				[0.068]	[0.456]	[1.000]	[0.857]	[0.367]	[0.240]
Education	Graduated High School	30	0.600	0.073	0.044	0.116	0.083	0.040	0.063
				(0.262)	(0.375)	**(0.001)**	(0.346)	(0.407)	(0.317)
				[0.653]	[0.582]	[1.000]	[0.908]	[0.597]	[0.565]
	Graduated 4-year College	30	0.120	0.170	0.138	0.149	0.099	0.135	0.143
				(0.055)	(0.128)	(0.216)	(0.338)	(0.154)	(0.130)
				[0.280]	[0.505]	[0.219]	[0.896]	[0.538]	[0.473]
	Years of Education	30	12.867	0.525	0.541	1.010	0.777	0.351	0.344
				(0.151)	(0.163)	(0.998)	(0.136)	(0.280)	(0.256)
				[0.564]	[0.505]	[1.000]	[0.638]	[0.597]	[0.565]
Labor Income	Employed	30	0.700	0.119	0.196	0.108	0.040	0.237	0.261
				(0.128)	**(0.025)**	**(0.001)**	(0.383)	**(0.025)**	**(0.013)**
				[0.456]	[0.136]	[1.000]	[0.567]	[0.113]	**[0.080]**
	Labor Income	30	30,079	19,810	24,365	25,220	20,611	23,072	21,836
				(0.091)	**(0.092)**	(0.998)	(0.122)	(0.107)	**(0.094)**
				[0.357]	[0.293]	[1.000]	[0.390]	[0.339]	[0.321]
Crime	Total Felony Arrests	Mid-30s	1.370	0.196	0.685	1.523	1.340	0.481	0.188
				(0.368)	(0.183)	**(0.064)**	**(0.026)**	(0.284)	(0.410)
				[0.396]	[0.429]	[0.120]	**[0.081]**	[0.562]	[0.434]
	Total Misdemeanor Arrests	Mid-30s	1.296	-0.501	-0.244	-0.298	-0.034	-0.246	-0.507
				(0.171)	(0.289)	(0.314)	(0.422)	(0.329)	(0.168)
				[0.395]	[0.429]	[0.314]	[0.422]	[0.562]	[0.411]
Health	Self-reported drug user	Mid-30s	0.500	-0.333	-0.438	-0.673	-0.557	-0.326	-0.330
				(0.019)	**(0.002)**	**(0.000)**	**(0.000)**	**(0.039)**	**(0.023)**
				[0.092]	**[0.014]**	**[0.010]**	**[0.067]**	[0.102]	[0.112]
	Systolic Blood Pressure (mm Hg)	Mid-30s	138.071	-9.791	-13.275	14.196	14.976	-24.166	-18.559
				(0.113)	**(0.049)**	**(0.013)**	**(0.000)**	**(0.000)**	**(0.011)**
				[0.196]	**[0.086]**	[0.116]	**[0.001]**	**[0.010]**	**[0.014]**
	Diastolic Blood Pressure (mm Hg)	Mid-30s	89.214	-10.854	-14.134	-9.709	-8.741	-18.387	-13.987
				(0.032)	**(0.004)**	**(0.049)**	**(0.032)**	**(0.000)**	**(0.007)**
				[0.089]	**[0.012]**	[0.200]	[0.168]	**[0.011]**	**[0.014]**
	Hypertension	Mid-30s	0.571	-0.291	-0.377	-0.120	-0.074	-0.492	-0.434
				(0.042)	**(0.009)**	(0.302)	(0.353)	**(0.006)**	**(0.006)**
				[0.115]	**[0.036]**	[0.610]	[0.771]	**[0.018]**	**[0.014]**

Note: This table shows the treatment effects for categories of outcomes that are important for García et al. (2017). Systolic and diastolic blood pressure are measured in terms of mm Hg. Each column present estimates for the following parameters: (1) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}^0 | \boldsymbol{B} \in \mathcal{B}_0\right]$ (no adjusting for covariates); (2) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}^0 | \boldsymbol{B} \in \mathcal{B}_0\right]$ (adjusting for covariates); (3) $\mathbb{E}\left[\boldsymbol{Y}^1 | R = 1\right] - \mathbb{E}\left[\boldsymbol{Y}^0 | R = 0, V = 0\right]$ (no adjusting for covariates); (4) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}_H^0 | \boldsymbol{B} \in \mathcal{B}_0\right]$ (adjusting for covariates); (5) $\mathbb{E}\left[\boldsymbol{Y}^1 | R = 1\right] - \mathbb{E}\left[\boldsymbol{Y}^0 | R = 0, V = 1\right]$ (no adjusting for covariates); (6) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}_C^0 | \boldsymbol{B} \in \mathcal{B}_0\right]$ (adjusting for covariates). We account for the following background variables ($\boldsymbol{B}$): ABC/CARE indicator, Apgar scores at minutes 1 and 5, and the High Risk Index. We define the High Risk Index in Appendix A.2 and explain how we choose the control variables in Appendix D.1. Columns (2), (4), and (6) correct for item non-response and attrition using inverse probability weighting as we explain in Appendix B.2. Inference is based on non-parametric, one-sided p-values from the empirical bootstrap distribution. Step down p-values are reported in square brackets. We highlight point estimates significant at the 10% level. See Appendix D.12.2 for two-sided p-values.

Table 5: Treatment Effects on Selected Outcomes, Females*

Category	Variable	Age	$\bar{Y}_C$	(1)	(2)	(3)	(4)	(5)	(6)
Parental Income	Parental Labor Income	3.5	11,465	2,756	2,986	6,864	8,584	1,521	3,773
				(0.189)	(0.213)	(0.122)	**(0.045)**	(0.332)	(0.154)
				[0.447]	[0.519]	[0.249]	[0.143]	[0.624]	[0.496]
		12	20,917	13,633	19,592	28,328	26,489	15,343	18,678
				(0.054)	**(0.027)**	**(0.027)**	**(0.009)**	**(0.064)**	**(0.019)**
				[0.310]	[0.179]	[0.124]	**[0.035]**	[0.358]	[0.128]
		15	13,772	8,565	7,159	2,713	8,441	7,465	10,487
				(0.060)	(0.137)	(0.480)	(0.345)	(0.134)	**(0.064)**
				[0.310]	[0.519]	[0.709]	[0.547]	[0.504]	[0.262]
		21	20,804	5,708	8,670	45,697	25,142	6,251	3,943
				(0.136)	(0.140)	**(0.000)**	**(0.000)**	(0.224)	(0.261)
				[0.402]	[0.519]	**[0.009]**	**[0.001]**	[0.589]	[0.510]
Education	Graduated High School	30	0.529	0.253	0.131	0.553	0.595	-0.026	0.066
				(0.009)	(0.152)	**(0.003)**	**(0.000)**	(0.413)	(0.320)
				[0.072]	[0.513]	**[0.021]**	**[0.001]**	[0.698]	[0.701]
	Years of Education	30	11.794	2.143	1.843	3.861	3.923	1.163	1.409
				(0.001)	**(0.002)**	**(0.000)**	**(0.000)**	**(0.054)**	**(0.017)**
				[0.003]	**[0.033]**	**[0.010]**	**[0.001]**	[0.365]	[0.111]
Labor Income	Employed	30	0.706	0.131	0.081	0.381	0.340	-0.010	0.070
				(0.096)	(0.206)	**(0.039)**	**(0.057)**	(0.465)	(0.264)
				[0.261]	[0.501]	[0.108]	[0.131]	[0.809]	[0.660]
	Labor Income	30	23,267	2,548	1,884	15,094	13,096	-2,677	-2,122
				(0.335)	(0.382)	**(0.056)**	**(0.022)**	(0.330)	(0.363)
				[0.407]	[0.586]	[0.133]	**[0.086]**	[0.781]	[0.660]
Crime	Total Felony Arrests	Mid-30s	0.419	-0.328	-0.351	-0.944	-0.965	-0.059	0.004
				(0.077)	**(0.087)**	**(0.095)**	**(0.095)**	(0.287)	(0.500)
				[0.134]	[0.215]	[0.167]	[0.186]	[0.432]	[0.610]
	Total Misdemeanor Arrests	Mid-30s	1.161	-0.973	-0.737	-2.010	-2.451	-0.269	-0.201
				(0.057)	(0.134)	(0.134)	(0.120)	(0.273)	(0.289)
				[0.134]	[0.238]	[0.167]	[0.186]	[0.432]	[0.610]
Health	Self-reported drug user	Mid-30s	0.259	-0.033	0.004	-0.114	-0.101	0.020	0.033
				(0.381)	(0.478)	(0.273)	(0.323)	(0.443)	(0.406)
				[0.844]	[0.857]	[0.745]	[0.559]	[0.611]	[0.943]
	Systolic Blood Pressure (mm Hg)	Mid-30s	133.963	-2.899	-5.407	-0.488	-0.822	-6.239	-6.784
				(0.307)	(0.241)	(0.488)	(0.457)	(0.249)	(0.170)
				[0.418]	[0.569]	[0.832]	[0.664]	[0.578]	[0.433]
	Diastolic Blood Pressure (mm Hg)	Mid-30s	87.556	-0.002	-0.179	4.091	4.122	-1.347	-2.160
				(0.483)	(0.438)	(0.245)	(0.222)	(0.392)	(0.339)
				[0.486]	[0.643]	[0.679]	[0.659]	[0.611]	[0.569]
	Hypertension	Mid-30s	0.407	0.172	0.085	0.077	0.162	0.102	0.107
				(0.111)	(0.293)	(0.331)	(0.245)	(0.299)	(0.255)
				[0.288]	[0.643]	[0.800]	[0.664]	[0.611]	[0.569]

Note: This table shows the treatment effects for categories of outcomes that are important for García et al. (2017). Systolic and diastolic blood pressure are measured in terms of mm Hg. Each column present estimates for the following parameters: (1) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}^0 | \boldsymbol{B} \in \mathcal{B}_0\right]$ (no adjusting for covariates); (2) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}^0 | \boldsymbol{B} \in \mathcal{B}_0\right]$ (adjusting for covariates); (3) $\mathbb{E}\left[\boldsymbol{Y}^1 | R = 1\right] - \mathbb{E}\left[\boldsymbol{Y}^0 | R = 0, V = 0\right]$ (no adjusting for covariates); (4) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}^0_H | \boldsymbol{B} \in \mathcal{B}_0\right]$ (adjusting for covariates); (5) $\mathbb{E}\left[\boldsymbol{Y}^1 | R = 1\right] - \mathbb{E}\left[\boldsymbol{Y}^0 | R = 0, V = 1\right]$ (no adjusting for covariates); (6) $\mathbb{E}\left[\boldsymbol{Y}^1 - \boldsymbol{Y}^0_C | \boldsymbol{B} \in \mathcal{B}_0\right]$ (adjusting for covariates). We account for the following background variables ($\boldsymbol{B}$): ABC/CARE indicator, Apgar scores at minutes 1 and 5, and the High Risk Index. We define the High Risk Index in Appendix A.2 and explain how we choose the control variables in Appendix D.1. Columns (2), (4), and (6) correct for item non-response and attrition using inverse probability weighting as we explain in Appendix B.2. Inference is based on non-parametric, one-sided p-values from the empirical bootstrap distribution. Step down p-values are reported in square brackets. We highlight point estimates significant at the 10% level. See Appendix D.12.2 for two-sided p-values.

*For females, we do not report graduation from a four-years college because we lack of common support to compute estimates for some parameters.

thorough analysis to conclude that there is little sensitivity to the choice of these variables.[45]

We also test whether estimated treatment effects are equal across genders. Table 6 reports differences between males and females in the control and treatment groups across major categories of outcomes. Within each category there are multiple measures. Using the Wilcoxon signed-rank test, we reject equality of the distributions of the treatment and control groups.[46] The general pattern is that male results are stronger than female results in the control group. The pattern is generally reversed in the treatment group. The results are robust to adjustments for multiple hypothesis testing. Similar results emerge when we use a more complete set of outcomes (see Appendix C.2).

4.2 Estimated Combining Functions

We consider a total of 126 outcomes reported in Appendix C.2. Measures of cognitive, social-emotional, and parenting skills were collected during the intervention or while the subjects were in school. The researchers collected information on the subjects' academic performance including grade retention and special education. The adult surveys (at ages 21 and 30) cover items related to employment, post-secondary education, health, criminal activity, and family structure. When the subjects were in their mid 30s, the researchers collected administrative crime data and data from a full medical survey.

Given the large number of outcomes measured in the numerous follow-ups, reporting all treatment effects would overwhelm the reader. A companion paper, García et al. (2017), aggregates treatment effects by monetizing benefits and constructing lifetime profiles of outcomes to conduct cost/benefit and rate of return analyses. The benefits from ABC/CARE are largely driven by its effects on males. The mean benefit/cost ratio is 10.19 for males and

[45] We also present this sensitivity analysis changing the variables used to condition while estimating treatment effects and changing the variables used to construct the weights to account for attrition.

[46] We explain the procedure in its entirety in Appendix B.3.3. In summary, we bootstrap with replacement 1,000 times. For bootstrap $b \in [1, \dots, 1,000]$ we compute the treatment effect for males, $T_{M,b}$, and females, $T_{F,b}$. These treatment effects represent a pair $(T_{M,b}, T_{F,b})$. Over the distribution of these pairs, we compute the Wilcoxon signed-rank test, which compares the empirical distribution of the treatment effects for males to the empirical distribution of the treatment effects for females. We repeat this procedure for each outcome.

Table 6: Gender Differences of Treatment Effects

Category	Variable	Age	**Control Mean**				**Treatment Effect**			
			Male	Female	Difference	p -value	Male	Female	Difference	p -value
Cognitive	IQ	5	94.214	95.735	-1.521	< 0.001*	7.697	4.921	2.775	< 0.001*
		8	93.871	93.161	0.710	0.022*	4.160	5.906	-1.746	< 0.001*
	Achievement	8	93.384	91.931	1.453	0.002*	2.309	6.619	-4.311	< 0.001*
		12	90.107	87.037	3.070	< 0.001*	2.404	9.631	-7.227	< 0.001*
Social-emotional	Task orientation	1	21.143	20.694	0.448	< 0.001*	0.896	0.940	-0.044	0.696
		1.5	20.676	21.083	-0.407	< 0.001*	1.861	2.939	-1.078	< 0.001*
	Sociability	1	10.361	11.250	-0.889	< 0.001*	0.246	0.527	-0.282	< 0.001*
		1.5	11.147	11.057	0.090	0.004*	-0.710	1.074	-1.784	< 0.001*
Parental Income	Parental Labor Income	3.5	13,505	11,465	2,040	< 0.001*	1,036	2,756	-1,720	< 0.001*
		12	23,868	20,917	2,951	0.001*	7,085	13,633	-6,547	< 0.001*
		15	22,985	13,772	9,213	< 0.001*	8,488	8,565	-76.864	0.404
		21	21,585	20,804	781.402	0.934	12,732	5,708	7,024	< 0.001*
Education	Graduated High School	30	0.600	0.529	0.071	< 0.001*	0.073	0.253	-0.180	< 0.001*
	Graduated 4-year College	30	0.120	0.088	0.032	0.001*	0.170	0.134	0.036	< 0.001*
	Years of Education	30	12.867	11.794	1.073	< 0.001*	0.525	2.143	-1.618	< 0.001*
Labor Income	Employed	30	0.700	0.706	-0.006	0.348	0.119	0.131	-0.012	0.275
	Labor Income	30	30,079	23,267	6,812	< 0.001*	19,810	2,548	17,262	< 0.001*
Crime	Total Felony Arrests	Mid-30s	1.370	0.419	0.951	< 0.001*	0.196	-0.328	0.524	< 0.001*
	Total Misdemeanor Arrests	Mid-30s	1.296	1.161	0.135	< 0.001*	-0.501	-0.973	0.472	< 0.001*
Health	Self-reported drug user	Mid-30s	0.500	0.259	0.241	< 0.001*	-0.333	-0.033	-0.301	< 0.001*
	Systolic Blood Pressure (mm Hg)	Mid-30s	138.071	133.963	4.108	< 0.001*	-9.791	-2.899	-6.892	< 0.001*
	Diastolic Blood Pressure (mm Hg)	Mid-30s	89.214	87.556	1.659	0.009*	-10.854	-0.002	-10.853	< 0.001*
	Hypertension	Mid-30s	0.571	0.407	0.164	< 0.001*	-0.291	0.172	-0.464	< 0.001*

Note: This table reproduces the control-group means and estimated treatment effects reported in (1) in Tables 4 and 5. The "Difference" columns show the difference between the males' control mean (or treatment effect) and the females' control mean (or treatment effect). The p-values corresponding to each difference is from a Wilcoxon signed-rank test that compares the empirical distributions over 100 bootstrapped resamples. An asterisk indicates that the p-value remains significant after adjusting for multiple hypothesis testing. The blocks are over outcome categories reported in Appendix C.2. We report the analogous estimates for the 126 outcomes in Appendix C.2.

2.61 for females with a range of estimates all bounded above 1.

In this section, we report treatment effects by category and the proportion of statistically significant effects. We use combining functions, which count the number of positive (and significant) treatment effects by gender. We adjust for dependence across outcomes and for pretesting in constructing standard errors (see Appendix B).

The combining functions indicate that females benefit more from the program than do males, in the sense that there are more outcomes across multiple domains for which female treatments benefit relative to female controls. However, as shown in García et al. (2017), the monetary benefits of the program are greater for boys in part because it reduces their more costly criminal activity.

We find strong evidence of gender differences in education and employment. For females in comparison to alternative childcare, 40% of the education outcomes and 20% of the employment outcomes are positive and significant. These increase to 80% and 100% when compared to staying at home. The treatment advantage for males is only seen when compared to those males who attended alternative childcare centers. None of the education or employment outcomes are statistically significant and positive when compared to staying at home.

We test the hypothesis that the proportions are equal to 50%. Figure 3 shows that the proportions for both genders are statistically significantly greater than 50%. The proportion is higher for females. When considering the proportion of outcomes that are both positive and statistically significant at the 10% level, we test the hypothesis that the proportions are equal to 10%.[47] A similar pattern holds for this test as well, although the proportions are smaller. Across outcomes, the effects are stronger for males when comparing treatment to alternative childcare, but are stronger for females when comparing treatment to staying at home. This pattern holds in the individual treatment effects, the combining functions, and

[47] Our inference accounts for the dependence across outcomes, as we explain in Section 3.1 and Appendix B. Using an α-level of significance, one would expect to find by chance that α% of the treatment effects are "statistically significant," even if the null hypothesis of no effect of the program is true.

the cost/benefit analysis of García et al. (2017).

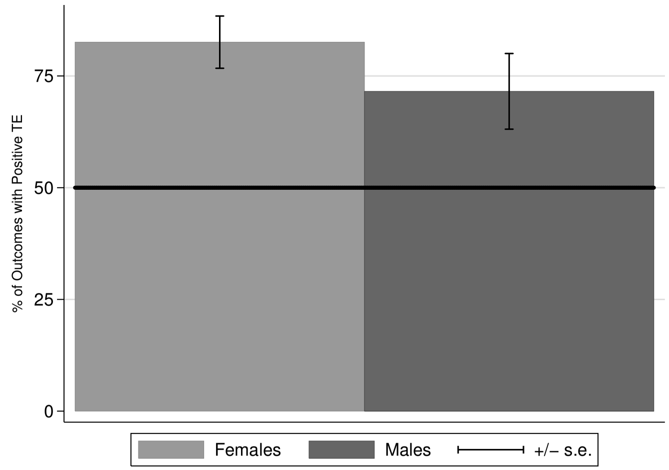
We next report estimates of the proportion of beneficial effects by outcome category and overall. This analysis is based on treatment effect (2). Figure 3 displays the results from this analysis: ABC/CARE positively impacted a large percentage of the outcomes. We show the counts for treatment compared to the next best alternative chosen by parents in Figure 3a. Proportionately more outcomes are beneficial for females, but the proportions are high for both groups and well above the benchmark of $\frac{1}{2}$. In Tables D.4 to D.12 of Appendix D, we document a large and precisely determined fraction of beneficial treatment effects well above $\frac{1}{2}$ for both genders for categories of outcomes spanning the life cycle through the mid 30s. At a 10% level of significance, 46% are statistically significant for females and 28% for males (see Figure 3b).

Figures 3c and 3d adjust the count measures used to construct Figure 3a to analyze more clearly defined counterfactuals: treatment compared to staying at home and treatment compared to alternative center childcare. These comparisons indicate that males and females benefit differently from alternatives to high-quality treatment. Compared across all categories, females benefit more from treatment when compared to staying at home (as opposed to attending alternative childcare), while males benefit more from treatment when compared to attending an alternative childcare arrangement (as opposed to staying at home).

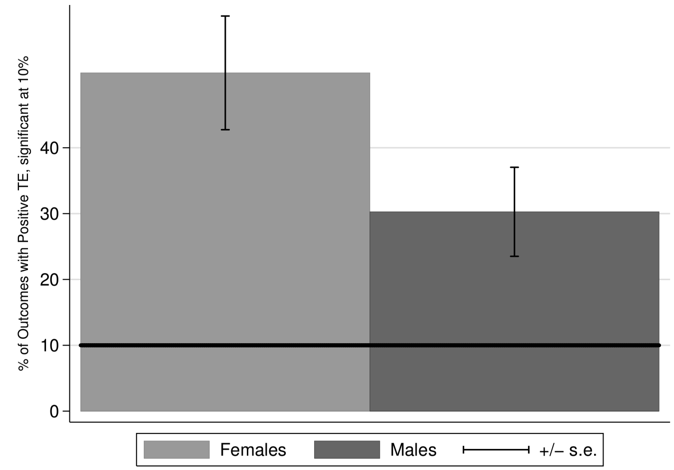
We also present estimates of the combining functions by outcome category. Figure 4 shows the estimated proportions that are statistically significantly positive at the 10% level. Consistent with the treatment effects above, control-group females tend to do better in alternative childcare than at home. This is especially true for parental income (induced by the childcare benefits of the program), IQ, education, and employment. Control-group males, on the other hand, do better at home, with more positive treatment effects compared to low-quality childcare.

Figure 3: Positively Impacted Outcomes, ABC/CARE Males and Females

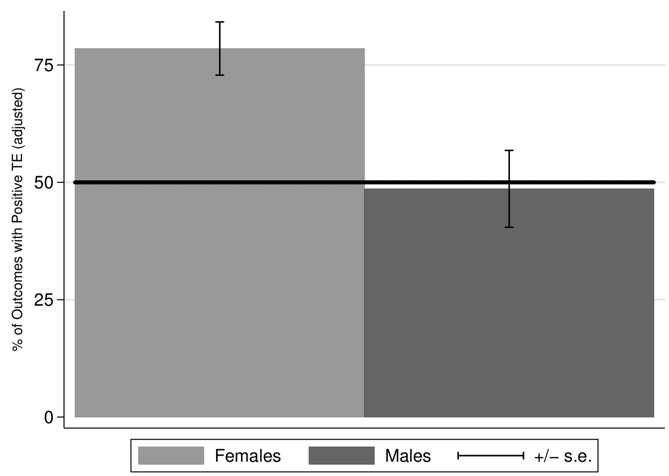
(a) Treatment vs. Next Best



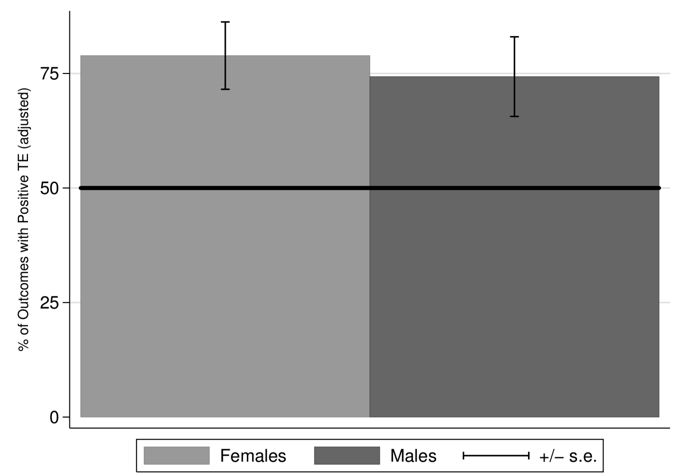
(b) Treatment vs. Next Best, Significant at 10% Level



(c) Treatment vs. Stay at Home



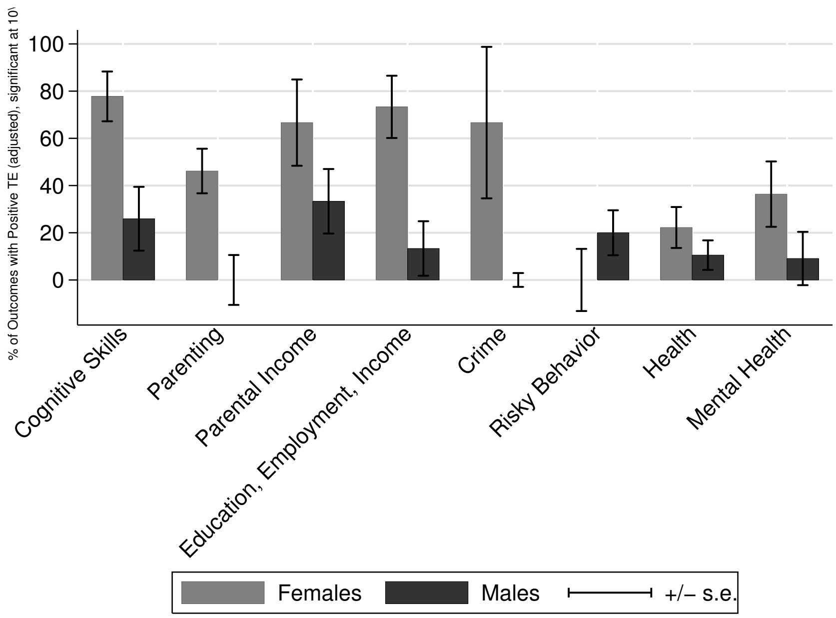
(d) Treatment vs. Alternative Childcare



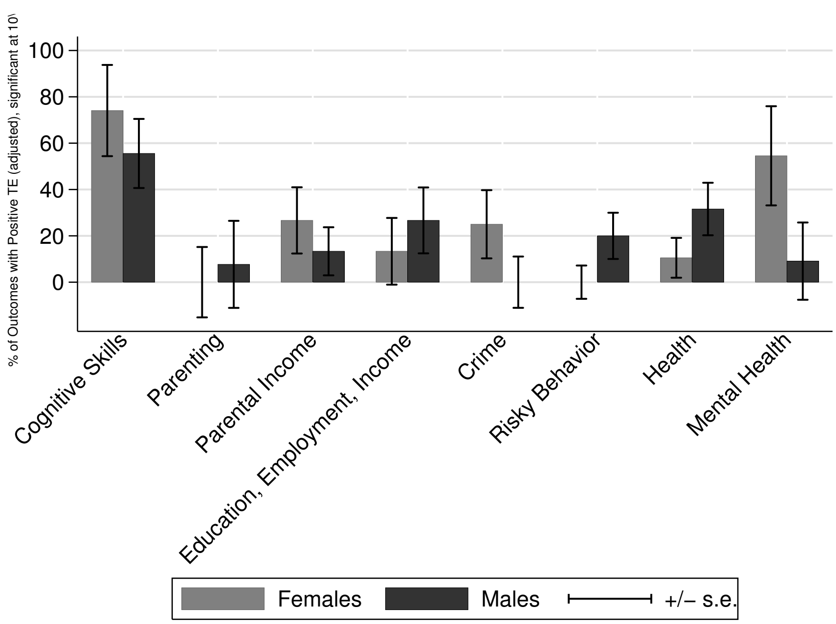
Note: Panel (a) displays the percentage of outcomes displaying a positive treatment effect, comparing treatment to the next best option. Panel (b) displays the percentage of outcomes displaying a positive and statistically significant treatment effect (10% significance level). Panel (c) displays the percentage of outcomes with a positive treatment effect, comparing treatment to staying at home. Panel (d) displays the percentage of outcomes with a positive treatment effect, comparing treatment to alternative childcare arrangements. Standard errors are based on the empirical bootstrap distribution. For Panel (b) we perform a "double bootstrap" procedure to first determine significant treatment effects at 10% level and then calculate the standard error of the count.

Figure 4: Proportion of Positively Impacted Outcomes by Category, ABC/CARE Males and Females

(a) Treatment vs. Stay at Home, Significant at 10% Level



(b) Treatment vs. Alternative Childcare, Significant at 10% Level



Note: Panel (a) displays the percentage of outcomes with a positive and statistically significant treatment effect (10% significance level), comparing treatment to staying at home. Panel (a) displays the percentage of outcomes with a positive and statistically significant treatment effect (10% significance level), comparing treatment to alternative childcare arrangements. Standard errors are based on the empirical bootstrap distribution. We perform a "double bootstrap" procedure to first determine significant treatment effects at 10% level and then calculate the standard error of the count. Standard errors with no bar indicate that the proportion is 0.

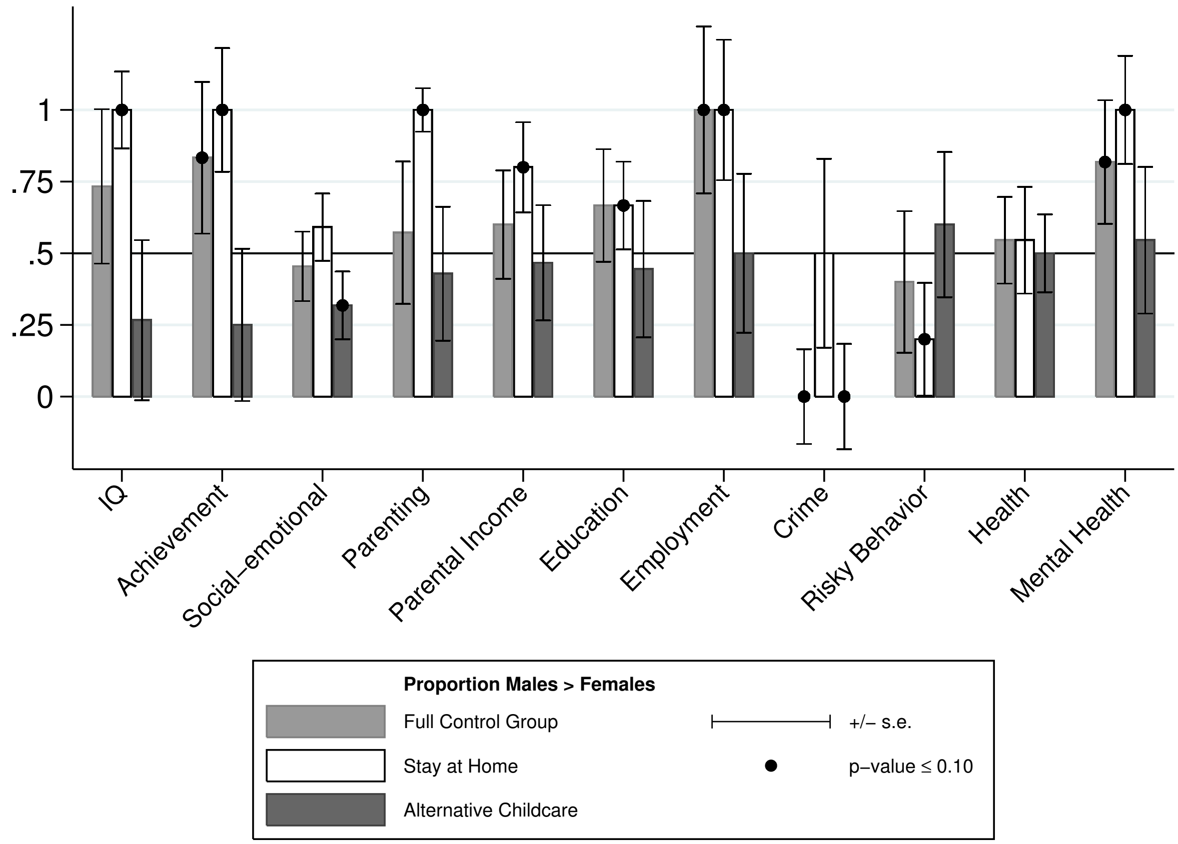
5 Explaining Gender Differences

This section discusses the sources of the estimated gender differences. A major determinant is the choice of the alternative childcare arrangements. Estimated treatment effects are very similar across genders comparing treatment to staying at home full time. Males benefit much more from treatment relative to low-quality childcare compared to treatment relative to staying at home. This result is consistent with previous research that shows (i) substantial gender differences resulting from attending low-quality childcare (Kottelenberg and Lehrer, 2014; Baker et al., 2015); and (ii) that females are less sensitive to more stressful, low-quality environments (see, e.g., Golding and Fitzgerald, 2016; Autor et al., 2015).

Table 1 shows the proportion of outcomes by category for which male outcomes exceed those of females. In Appendix C.3, we present analogous tables for staying at home and attending alternative childcare. We also partition the sample by father's presence, a potentially important moderator.[48] Figure 5 shows the proportions partitioned by alternative childcare setting. The males who stay at home do better than the females in cognitive and parenting measures, employment, and across all outcomes. Unlike the males who attend lower-quality alternative childcare, the males who stay at home have similar crime outcomes as the females. Given the important negative effects of male criminal activity, this finding highlights the magnitude of harm caused by low-quality alternative childcare for males.

[48] Recall that we condition on baseline variables to control for selection into childcare for controls. See Appendix E for a sensitivity study using other methods. They produce comparable estimates, but are less precisely estimated.

Figure 5: Proportion of Outcomes Males > Females, by Outcome Category, Partitioning by Alternative Childcare Setting



Note: These plots show the proportion of outcomes, by outcome category, for which the males' mean is larger than the females' mean. The standard errors and the p-values are computed using 1,000 bootstraps. The p-values are one-sided and test the null hypothesis that the proportion of outcomes is greater than 50% All outcome categories have higher values corresponding to socially desirable outcomes.

Appendix C reports estimates by whether or not the father is present (Figure C.1). Parenting is better for control-group males when the father is absent. While this may seem contradictory, the parenting measure includes a scale that captures the absence of punishment. Control-group males with an absent father tend to have higher parenting scores, especially for this scale, relative to control-group males with a father present. In the treatment group, the males with a present father do better than those with an absent father. This hints at a complementarity between the father's parenting and the high-quality treatment. Besides this, few clear-cut patterns emerge. Father's presence interacted with treatment also favors males for IQ. Males in the control group do better when the father is absent in employment.

Another measure of home environment is maternal locus of control which is measured when the subjects were 1.5 years old.[49] An internal locus of control, which is socially desirable, indicates feelings that the respondents can control future outcomes. In contrast, an external locus of control indicates feeling that outside forces determine future outcomes (e.g. luck). We consider pairs of opposing statements. The respondent chooses the statements that more closely align with her beliefs. One point is given for each statement the respondent selects that corresponds with an external locus of control. An example is: "In the long run people get the respect they deserve in this world" versus "Unfortunately, an individual's worth often passes unrecognized no matter how hard he tries." If the respondent selects the second statement, she has a greater perception of external locus of control. We define a respondent as having an internal locus of control if they score below the sample mean and as having an external locus of control if they score at or above the sample mean.

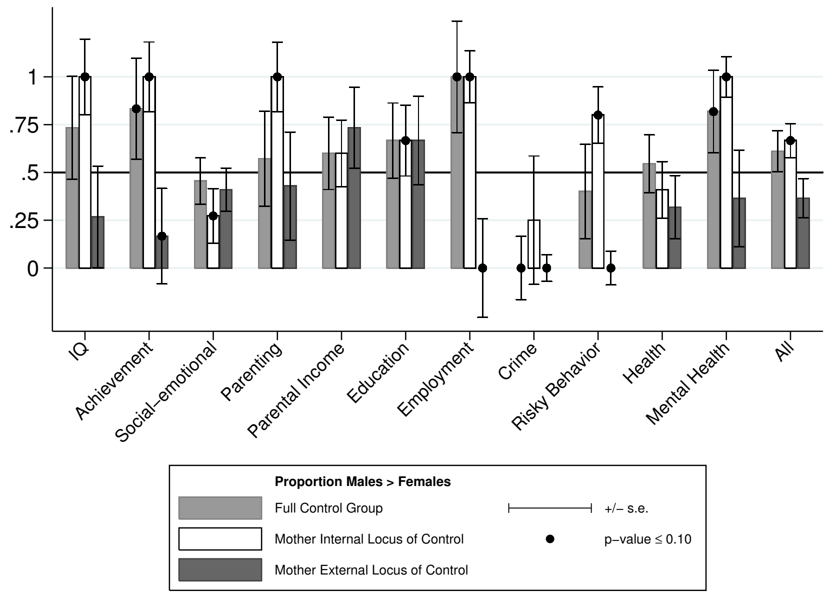
Figure 6 shows the proportion of outcomes by category that favor males classified by locus of control. In the control group, a higher internal maternal locus of control promotes better outcomes for males across all outcome categories except health. Another way to state this is that males are more negatively affected if their mothers have a higher external locus of control than are females. This is consistent with the analyses of Schore (2017) and Golding and Fitzgerald (2016). Although locus of control does not necessarily measure depression, this finding is consistent with Beeghly et al. (2017), who finds that increased maternal depression leads to worse outcomes for young males relative to females.

Maternal locus of control moderates treatment outcomes differently. Aggregating across outcomes, male treatment outcomes are not much affected by the mother's locus of control. Outcomes are affected for those in the control group. This is especially strong for mental health outcomes. The reduced differential in the treatment group suggests an important

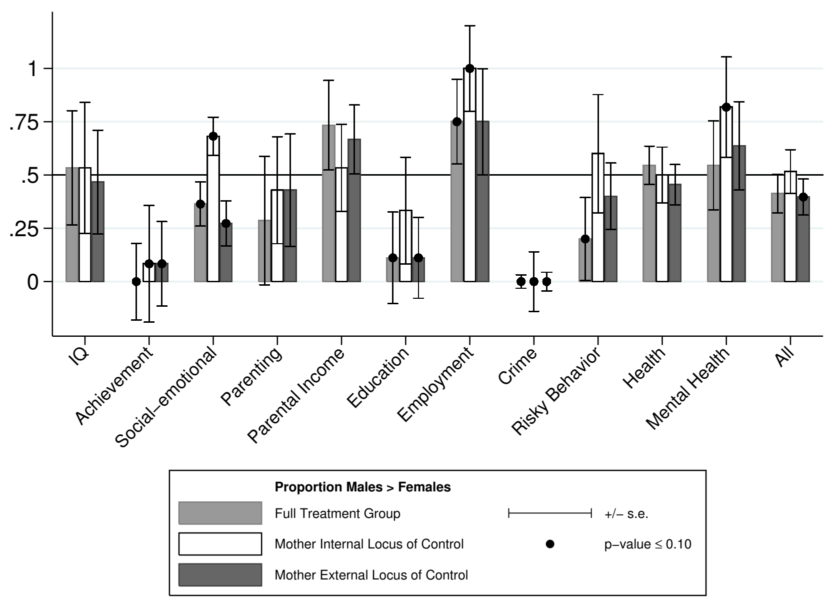
[49] See Rotter (1966). The researchers implementing ABC/CARE slightly modified the scale to be more appropriate for the population.

Figure 6: Proportion of Outcomes Males > Females, by Outcome Category, Dividing by Maternal Locus of Control

(a) Control Group



(b) Treatment Group



Note: These plots show the proportion of outcomes, by outcome category, for which the males' mean is larger than the females' mean. The standard errors and the p-values are computed using 1,000 bootstraps. The p-values are one-sided and test the null hypothesis that the proportion of outcomes is greater than 50%. All outcome categories have higher values corresponding to socially desirable outcomes.

compensating role of enriched early childhood programs.

6 Conclusion

This paper examines gender differences in the impacts of treatment of an influential early childhood program targeted to disadvantaged children. Gender gaps favoring males before treatment are altered by treatment. Control-group males tend to do better than the control-group females, especially in education and employment outcomes. After treatment, most gender gaps are narrowed or even reversed. This corresponds with the finding that a larger proportion of the treatment effects are positive and statistically significant for females than for males. ABC/CARE, in addition to improving select individual outcomes, also narrowed the male-female gap in important outcome categories.

Low-quality childcare arrangements are detrimental for boys. This finding helps explain an apparent contradiction in the literature (Baker et al. (2008, 2015); Kottelenberg and Lehrer (2014)) that childcare can harm children. Low-quality childcare harms boys but girls are robust to their childcare arrangements. This evidence is consistent with previous research that finds boys to be more vulnerable early in life than girls (Golding and Fitzgerald, 2016). Similarly, boys are more likely to be harmed by unfavorable home environments (measured by father present and maternal locus of control) than are girls.

Our analysis sounds a cautionary note about the value of early childcare programs. In the past decade, many politicians and pundits have warmly embraced early childhood programs as solutions for reducing inequality and promoting social mobility. Little attention has been paid to the quality of those programs. High-quality early childhood programs have positive effects on both boys and girls. García et al. (2017) show that they have a high economic rate of return. This study shows that low-quality early childhood programs can have harmful effects on child development, especially for boys, providing evidence on the benefits of high-quality programs.

References

Autor, D. H., D. Figlio, K. Karbownik, J. Roth, and M. Wasserman (2015, October). Family disadvantage and the gender gap in behavioral and educational outcomes. Working Paper WP-15-16, Institute for Policy Research, Northwestern University, Evanston, IL.

Baker, M., J. Gruber, and K. Milligan (2005, December). Universal childcare, maternal labor supply, and family well-being. Working Paper 11832, National Bureau of Economic Research.

Baker, M., J. Gruber, and K. Milligan (2008, August). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy 116*(4), 709–745.

Baker, M., J. Gruber, and K. Milligan (2015, September). Non-cognitive deficits and young adult outcomes: The long-run impacts of a universal child care program. Working Paper 21571, National Bureau of Economic Research.

Beeghly, M., T. Partridge, E. Tronick, M. Muzik, M. R. Mashhadi, J. L. Boeve, and J. L. Irwin (2017, January/February). Associations between early maternal depressive symptom trajectories and toddlers' felt security at 18 months: Are boys and girls at differential risk? *Infant Mental Health Journal 38*(1), 53–67.

Bernal, R. and M. P. Keane (2011, July). Child care choices and children's cognitive achievement: The case of single mothers. *Journal of Labor Economics 29*(3), 459–512.

Bertrand, M. and J. Pan (2013, January). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics 5*(1), 32–64.

Bitler, M. P., H. W. Hoynes, and T. Domina (2014). Experimental evidence on distributional effects of Head Start. Working Paper 20434, National Bureau of Economic Research.

Burchinal, M. R., F. A. Campbell, D. M. Bryant, B. H. Wasik, and C. T. Ramey (1997, October). Early intervention and mediating processes in cognitive performance of children of low-income African American families. *Child Development 68*(5), 935–954.

Campbell, F. A., G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. P. Pungello, and Y. Pan (2014). Early childhood investments substantially boost adult health. *Science 343*(6178), 1478–1485.

Campbell, F. A., E. P. Pungello, S. Miller-Johnson, M. Burchinal, and C. T. Ramey (2001, March). The development of cognitive and academic abilities: Growth curves from an early childhood educational experiment. *Developmental Psychology 37*(2), 231–242.

Campbell, F. A., C. T. Ramey, E. Pungello, J. Sparling, and S. Miller-Johnson (2002). Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science 6*(1), 42–57.

Campbell, F. A., B. Wasik, E. Pungello, M. Burchinal, O. Barbarin, K. Kainz, J. Sparling, and C. Ramey (2008). Young adult outcomes of the Abecedarian and CARE early childhood educational interventions. *Early Childhood Research Quarterly 23*(4), 452–466.

Cascio, E. U. and D. W. Schanzenbach (2013). The impacts of expanding access to high-quality preschool education. Working Paper 19735, National Bureau of Economic Research.

Clarke, S. H. and F. A. Campbell (1998). Can intervention early prevent crime later? The Abecedarian Project compared with other programs. *Early Childhood Research Quarterly 13*(2), 319–343.

Collins, A., B. D. Goodson, J. Luallen, A. R. Fountain, A. Checkoway, and Abt Associates Inc. (2010, June). Evaluation of child care subsidy strategies: Massachusetts Family Child Care study. Technical Report OPRE 2011-1, Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, Washington, DC.

Conti, G., J. J. Heckman, and R. Pinto (2016). The effects of two influential early childhood interventions on health and healthy behaviours. Economic Journal *126*(596), F28–F65.

Currie, J. (2011). Inequality at birth: Some causes and consequences. *American Economic Review 101*(3), 1–22.

Department of Health, Education, and Welfare (1968). Federal interagency day care requirements, pursuant to Sec. 522 (D) of the Economic Opportunity Act. Technical report, U.S. Department of Labor, Washington, DC.

Educare (2014). A national research agenda for early education. Technical report, Educare Learning Network Research & Evaluation Committee, Chicago, IL.

Elango, S., , J. L. García, J. J. Heckman, and A. Hojman (2016). Early childhood education. In R. A. Moffitt (Ed.), *Economics of Means-Tested Transfer Programs in the United States*, Volume 2, Chapter 4, pp. 235–297. Chicago: University of Chicago Press.

García, J. L., J. J. Heckman, A. Hojman, Y. K. Koh, J. Shea, and A. L. Ziff (2016). Documentation of full ABC/CARE treatment effects. Unpublished Manuscript, University of Chicago.

García, J. L., J. J. Heckman, D. E. Leaf, and M. J. Prados (2017). The life-cycle benefits of an influential early childhood program. University of Chicago, Department of Economics.

Golding, P. and H. E. Fitzgerald (2016). Psychology of boys at risk: Indicators from 0–5. *Infant Mental Health Journal*.

Healthy Child Manitoba (2015, April). Starting early, starting strong: A guide for play-based early learning in Manitoba: Birth to six. Technical report, Healthy Child Manitoba, Winnipeg, Manitoba.

Heckman, J. J. (1992). Randomization and social policy evaluation. In C. F. Manski and I. Garfinkel (Eds.), *Evaluating Welfare and Training Programs*, Chapter 5, pp. 201–230. Cambridge, MA: Harvard University Press.

Heckman, J. J., N. Hohmann, J. Smith, and M. Khoo (2000, May). Substitution and dropout bias in social experiments: A study of an influential social experiment. *Quarterly Journal of Economics 115*(2), 651–694.

Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010, August). Analyzing social experiments as implemented: A reexamination of the evidence from the HighScope Perry Preschool Program. *Quantitative Economics 1*(1), 1–46.

Jensen, B. and M. Nielsen (2016). Abecedarian programme, within an innovative implementation framework (APIIF). A pilot study. Website, http://pure.au.dk/portal/en/projects/abecedarian-programme-within-an-innovative-implementation-framework-apiif-a-pilot-study(8d4da9a9-f2ff-44db-87cc-1278d1e7006c).html (Accessed 8/1/2016).

Kline, P. and C. Walters (2016). Evaluating public programs with close substitutes: The case of Head Start. Quarterly Journal of Economics *131*(4), 1795–1848.

Kottelenberg, M. J. and S. F. Lehrer (2014, August). The gender effects of universal child care in Canada: Much ado about boys? Unpublished manuscript, Department of Economics, Queen's University.

Lehmann, E. and J. P. Romano (2005, June). Generalizations of the familywise error rate. *Annals of Statistics 33*(3), 1138–1154.

North Carolina General Assembly (1971). Chapter 803, House bill 100. In *North Carolina General Assembly 1971 Session*, North Carolina. An Act to protect children through licensing of day-care facilities and other limited regulation.

Ramey, C. T., D. M. Bryant, J. J. Sparling, and B. H. Wasik (1985). Project CARE: A comparison of two early intervention strategies to prevent retarded development. *Topics in Early Childhood Special Education 5*(2), 12–25.

Ramey, C. T. and F. A. Campbell (1979, February). Compensatory education for disadvantaged children. *The School Review 87*(2), 171–189.

Ramey, C. T. and F. A. Campbell (1991). Poverty, early childhood education and academic competence: The Abecedarian experiment. In A. C. Houston (Ed.), *Children in Poverty: Child Development and Public Policy*, Chapter 8, pp. 190–221. New York: Cambridge University Press.

Ramey, C. T., A. M. Collier, J. J. Sparling, F. A. Loda, F. A. Campbell, D. A. Ingram, and N. W. Finkelstein (1976). The Carolina Abecedarian Project: A longitudinal and multidisciplinary approach to the prevention of developmental retardation. In T. Tjossem (Ed.), *Intervention Strategies for High-Risk Infants and Young Children*, pp. 629–655. Baltimore, MD: University Park Press.

Ramey, C. T., M. C. Holmberg, J. H. Sparling, and A. M. Collier (1977). An introduction to the Carolina Abecedarian Project. In B. M. Caldwell and D. J. Stedman (Eds.), *Infant Education: A Guide for Helping Handicapped Children in the First Three Years*, Chapter 7, pp. 101–121. New York: Walker and Company.

Ramey, C. T., G. D. McGinness, L. Cross, A. M. Collier, and S. Barrie-Blackley (1982). The Abecedarian approach to social competence: Cognitive and linguistic intervention for disadvantaged preschoolers. In K. M. Borman (Ed.), *The Social Life of Children in a Changing Society*, Chapter 7, pp. 145–174. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ramey, C. T. and B. J. Smith (1977, January). Assessing the intellectual consequences of early intervention with high-risk infants. *American Journal of Mental Deficiency 81*(4), 318–324.

Ramey, C. T., J. J. Sparling, and S. L. Ramey (2012). *Abecedarian: The Ideas, the Approach, and the Findings* (1 ed.). Los Altos, CA: Sociometrics Corporation.

Ramey, S. L., C. T. Ramey, and R. G. Lanzi (2014). Interventions for students from impoverished environments. In J. T. Mascolo, V. C. Alfonso, and D. P. Flanagan (Eds.), *Essentials of Planning, Selecting and Tailoring Interventions for Unique Learners*, pp. 415–48. Hoboken, NJ: John Wiley & Sons, Inc.

Romano, J. P. and A. M. Shaikh (2006, August). Stepup procedures for control of generalizations of the familywise error rate. *Annals of Statistics 34*(4), 1850–1873.

Romano, J. P. and M. Wolf (2016). Efficient computation of adjusted p-values for resampling-based stepdown multiple testing. *Statistics and Probability Letters 113*, 38–40.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied 80*(1), 1–28.

Schneider, B. and S.-K. McDonald (Eds.) (2007). *Scale-Up in Education*, Volume 1: Ideas in Principle. Lanham, Maryland: Rowman & Littlefield Publishers.

Schore, A. N. (2017, January/February). All our sons: The developmental neurobiology and neuroendocrinology of boys at risk. *Infant Mental Health Journal 38*(1), 15–52.

Scull, J., J. Hattie, J. Page, J. Sparling, C. Tayler, A. King, V. Nossar, A. Piers-Blundell, and C. Ramey (2015). Building a bridge into preschool in remote Northern Territory communities.

Sparling, J. (2010). Highlights of research findings from the Abecedarian studies. Technical report, Teaching Strategies, Inc., Center on Health and Education, Georgetown University, and FPG Child Development Institute, University of North Carolina at Chapel Hill, Bethesda, MD.

Sparling, J. and I. Lewis (1979). *LearningGames for the First Three Years: A Guide to Parent/Child Play*. New York: Walker and Company.

Spiker, D., C. W. Haynes, and R. T. Gross (Eds.) (1997). *Helping Low Birth Weight, Premature Babies: The Infant Health and Development Program.* Redwood City, CA: Stanford University Press.

Wasik, B. H., C. Ramey, D. M. Bryant, and J. J. Sparling (1990, December). A longitudinal study of two early intervention strategies: Project CARE. *Child Development 61*(6), 1682–1696.

Yazejian, N. and D. M. Bryant (2012). Educare implementation study findings. Technical report, Frank Porter Graham Child Development Institute, Chapel Hill, NC.