Appendix for

Using a Satisficing Model of Experimenter Decision-Making to Guide Finite-Sample Inference for Compromised Experiments*

Ganesh Karapakula — Yale University

James J. Heckman — University of Chicago

*Ganesh Karapakula: Yale University; email: ganesh.karapakula@yale.edu. James J. Heckman: Center for the Economics of Human Development, 1126 East 59th Street, Chicago, IL 60637; phone: 773-702-0634; fax: 773-702-8490; email: jjh@uchicago.edu. We thank Juan Pantano and Azeem Shaikh for comments on early drafts of this paper. We thank Cheryl Polk and Lawrence Schweinhart of the HighScope Educational Research Foundation for their assistance in data acquisition, sharing historical documentation, and their longstanding partnership with the Center for the Economics of Human Development. This research was supported in part by: the Buffett Early Childhood Fund; NIH Grants R01AG042390, R01AG05334301, and 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.

Contents

A Background and Eligibility Criteria of the Perry Program

The HighScope Perry Preschool Project was carried out in five waves between fall 1962 and fall 1965 near a public school called the Perry Elementary School in Ypsilanti, a small city near Detroit in Michigan. Data collection took place at the baseline age of 3 years and through surveys that were administered annually till age 15. The participants were additionally followed up around ages 19, 27, 40, and 55. Attrition continued to be low through 40, with about 83% of the 123 participants surveyed. About 12% of the participants were deceased at the latest follow-up,[1] and the rest of the attrited could not be located for the interview. Our estimators of treatment effects account for this attrition. Various measures were obtained over the years, including information on education, crime, and other economic outcomes. Intensity of the program was low relative to several later early education programs.[2] Starting at age 3, treatment in the following two years included preschool for 2.5 hours per day on weekdays during the academic year. Another major component of the program consisted of 1.5-hour weekly home visits by the Perry teachers to promote parental engagement with the child.[3] The Perry curriculum fostered active child-centered learning through intensive interactions between the children and program teachers (Schweinhart et al., 1993; Weikart et al., 1978).

Door-to-door canvassing and referrals were used to survey and identify disadvantaged families among those of the Perry Elementary School students. To be eligible for participation in the Perry Preschool Project, the children had to (i) be African-American; (ii) have low Stanford–Binet IQ scores at baseline;[4] and (iii) be socioeconomically disadvantaged according to an index of socioeconomic status based on employment and education levels of the parents as well as the number of persons per room at home. The Perry families were more disadvantaged relative to a majority of African-American families at that time in the United States. However, the Perry families were by and large representative of a substantial fraction of the underprivileged African-American population (Heckman et al., 2010).

Even when compared with the children living in the area surrounding the Perry Elementary School, the Perry participants were especially disadvantaged (Heckman et al., 2010). Since the parents of all children eligible for the program participated in the study (Weikart et al., 1978),

[1] There is negligible difference in mortality between treatments and controls (Heckman and Karapakula, 2020).

[2] An example is the Carolina Abecedarian Project (see, e.g., Campbell et al., 2002). For a discussion and comparison of the intensity of several such programs, see Cunha et al. (2006) and Elango et al. (2015).

[3] Those in the treatment group of the first entry cohort (wave 0) were provided the intervention for only one year, starting at age 4, and thus were an exception. In our estimation of treatment effects, we pool all five cohorts, even though the lower program intensity in the first cohort might in principle attenuate the magnitudes of the effects downward.

[4] The initial eligibility criteria specified that the IQs, as measured by the Stanford–Binet IQ test according to 1960s norming, be between 70 and 85, which was one standard deviation below the population average. However, in practice, the IQ range was 61 to 88. Only about two-thirds of the participants had IQs in the range specified initially.

issues of noncompliance are not a concern. As there were no substitutes to the Perry program, such as Head Start, available when the Perry experiment was implemented, control group contamination is also not a problem in our experimental setting.

B Exchanges Were Not Based on Consecutive IQ Scores

We use Perry data from wave 4 as an example to conclude that the exchanges were not necessarily between consecutively ranked pairs. In wave 4, there were 19 participants, excluding any younger siblings in the program. The IQs of these 19 people were: 61, 71, 75, 76, 76, 76, 78, 78, 79, 79, 80, 80, 81, 82, 83, 83, 83, 85, 88, involving many ties. Regardless of which method was used to break the ties, from a pure ranking procedure the staffers would have obtained two initial groups: one with IQs $\{61, 75, 76, 78, 79, 80, 81, 83, 83, 88\}$ and another group with IQs $\{71, 76, 76, 78, 79, 80, 82, 83, 85\}$. The final observed treatment group has IQs in the set: $\{61, 75, 76, 78, 80, 81, 83, 83, 83, 88\}$. Note that the person with IQ 79 is replaced by a person with IQ 83. The final observed control group has IQs in the set: $\{71, 76, 76, 78, 79, 79, 80, 82, 85\}$. Note that the person with IQ 83 is replaced by a person with IQ 79. These are the same as the initial treatment and control groups, since there were no transfers in the fifth step of the protocol, as explained in Example 3 of the paper. Thus, we can conclude that an exchange happened between participants with IQs 79 and 83, who do not comprise a consecutively ranked pair. Thus, after the IQ rank ordering, the exchanges between the two initial groups were not always between consecutively ranked IQ pairs. Thus, the Perry staffers did not strictly implement a matched pair design.

C Comparison of Inference from Alternative Approaches

The following series of appendices report the inference from a variety of inferential methods applied to the Perry sample. The general message is clear. Standard large-sample inferential methods produce overly optimistic inferences for the small Perry sample. The methods advocated in this paper are more conservative. Nonetheless, our design-based worst-case (least favorable) methods find statistically significant treatment effects for some outcomes at the usual significance levels.

In these appendices, for each outcome we include the conventional p-values (i.e., asymptotic, bootstrap, and permutation p-values) and design-based p-values (i.e., worst-case maximum and worst-case de Haan p-values) associated with each of the DIM (difference-in-means), OLS (ordinary least squares), and AIPW (augmented inverse probability weighting) estimators of treatment effects. We also include permutation and worst-case p-values based on both nonstudentized and studentized test statistics. In addition, we include stepdown versions of the worst-case p-values based on the Holm procedure.

D Inference on Cumulative Convictions for Violent Crimes

Table D-1: Treatment Effects on Violent Crime Conviction Outcomes of the **Pooled** Participants

	Statistic	Cumulative violent misdemeanor convictions age 30	Cumulative violent misdemeanor convictions age 40	Cumulative violent felony convictions age 30	Cumulative violent felony convictions age 40
Summary	(i) Number of observations	123	120	123	120
	(ii) Mean of the control group	0.5231	0.6825	0.2846	0.4762
	(iii) Mean of the treatment group	0.0517	0.0877	0.1897	0.1930
Estimates	(iv) DIM (difference in means) estimate	-0.4714	-0.5948	-0.0950	-0.2832
	(v) OLS estimate conditional on covariates	-0.5783	-0.7009	-0.0565	-0.2169
	(vi) AIPW (augmented IPW) estimate	-0.5300	-0.6491	-0.0561	-0.2052
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0109**	**0.0033**	0.2301	**0.0333**
	(02) Asymptotic p-value for OLS estimate	**0.0097**	**0.0038**	0.3248	**0.0676**
	(03) Asymptotic p-value for AIPW estimate	**0.0064**	**0.0021**	0.3174	**0.0664**
	(04) Bootstrap p-value for DIM estimate	**0.0021**	**0.0005**	0.2263	**0.0332**
	(05) Bootstrap p-value for OLS estimate	**0.0017**	**0.0006**	0.3217	**0.0708**
	(06) Bootstrap p-value for AIPW estimate	**0.0020**	**0.0010**	0.3217	**0.0778**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0036**	**0.0008**	0.2648	**0.0392**
	(08) Permutation p-value for Nonstudentized OLS	**0.0004**	**0.0004**	0.3604	**0.0704**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0016**	**0.0004**	0.3556	**0.0792**
	(10) Permutation p-value for Studentized DIM	**0.0036**	**0.0004**	0.2624	**0.0384**
	(11) Permutation p-value for Studentized OLS	**0.0028**	**0.0004**	0.3552	**0.0680**
	(12) Permutation p-value for Studentized AIPW	**0.0024**	**0.0008**	0.3488	**0.0708**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0122**	**0.0051**	0.4086	**0.0720**
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0093**	**0.0025**	0.4956	0.1488
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0135**	**0.0122**	0.4873	0.1633
	(16) Worst-case maximum p-value for Studentized DIM	**0.0122**	**0.0053**	0.4057	**0.0708**
	(17) Worst-case maximum p-value for Studentized OLS	**0.0122**	**0.0103**	0.4922	0.1443
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0099**	**0.0133**	0.4820	0.1543
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0132**	**0.0065**	0.4679	0.1118
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0096**	**0.0027**	0.5418	0.2043
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0142**	**0.0131**	0.5400	0.1894
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0154**	**0.0065**	0.5318	**0.0900**
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0154**	**0.0154**	0.5839	0.1542
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0154**	**0.0181**	0.5078	0.1754
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	**0.0367**	**0.0205**	0.4086	0.1441
	(26) Worst-case maximum p-value for Nonstudentized OLS	**0.0280**	**0.0100**	0.4956	0.2977
	(27) Worst-case maximum p-value for Nonstudentized AIPW	**0.0490**	**0.0490**	0.4873	0.3266
	(28) Worst-case maximum p-value for Studentized DIM	**0.0367**	**0.0211**	0.4057	0.1415
	(29) Worst-case maximum p-value for Studentized OLS	**0.0411**	**0.0411**	0.4922	0.2887
	(30) Worst-case maximum p-value for Studentized AIPW	**0.0394**	**0.0398**	0.4820	0.3086
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	**0.0395**	**0.0259**	0.4679	0.2235
	(32) Worst-case de Haan p-value for Nonstudentized OLS	**0.0289**	**0.0107**	0.5418	0.4086
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0523**	**0.0523**	0.5400	0.3787
	(34) Worst-case de Haan p-value for Studentized DIM	**0.0463**	**0.0259**	0.5318	0.1800
	(35) Worst-case de Haan p-value for Studentized OLS	**0.0617**	**0.0617**	0.5839	0.3084
	(36) Worst-case de Haan p-value for Studentized AIPW	**0.0617**	**0.0617**	0.5078	0.3507

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) $-$ (03) provide the one-sided asymptotic p-values, while rows (04) $-$ (06) provide the bootstrap p-values. Rows (07) $-$ (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) $-$ (12) provide those based on studentized test statistics. Rows (13) $-$ (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) $-$ (18) provide those based on studentized test statistics. Rows (19) $-$ (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) $-$ (24) provide those based on studentized test statistics. Rows (25) $-$ (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) $-$ (24). The block of variables used for Holm stepdown adjustment consists of all four of the above violent crime conviction outcome variables.

Table D-2: Treatment Effects on Violent Crime Conviction Outcomes of the **Male** Participants

	Statistic	Cumulative violent misdemeanor convictions age 30	Cumulative violent misdemeanor convictions age 40	Cumulative violent felony convictions age 30	Cumulative violent felony convictions age 40
Summary	(i) Number of observations	72	70	72	70
	(ii) Mean of the control group	0.5897	0.8421	0.4487	0.7632
	(iii) Mean of the treatment group	0.0909	0.1563	0.3333	0.3438
Estimates	(iv) DIM (difference in means) estimate	-0.4988	-0.6859	-0.1154	-0.4194
	(v) OLS estimate conditional on covariates	-0.6525	-0.8529	-0.0694	-0.3420
	(vi) AIPW (augmented IPW) estimate	-0.5961	-0.7750	-0.0711	-0.3236
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0081**	**0.0019**	0.2915	**0.0486**
	(02) Asymptotic p-value for OLS estimate	**0.0062**	**0.0025**	0.3742	**0.0892**
	(03) Asymptotic p-value for AIPW estimate	**0.0036**	**0.0013**	0.3618	**0.0831**
	(04) Bootstrap p-value for DIM estimate	**0.0061**	**0.0013**	0.2863	**0.0439**
	(05) Bootstrap p-value for OLS estimate	**0.0052**	**0.0016**	0.3719	**0.0899**
	(06) Bootstrap p-value for AIPW estimate	**0.0044**	**0.0019**	0.3650	**0.0936**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0116**	**0.0040**	0.3256	**0.0512**
	(08) Permutation p-value for Nonstudentized OLS	**0.0012**	**0.0004**	0.4016	**0.0880**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0028**	**0.0008**	0.3956	0.1000
	(10) Permutation p-value for Studentized DIM	**0.0104**	**0.0028**	0.3208	**0.0508**
	(11) Permutation p-value for Studentized OLS	**0.0072**	**0.0020**	0.3980	**0.0880**
	(12) Permutation p-value for Studentized AIPW	**0.0060**	**0.0016**	0.3876	**0.0928**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0220**	**0.0103**	0.4905	0.1128
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0117**	**0.0103**	0.5545	0.1804
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0167**	**0.0103**	0.5629	0.1878
	(16) Worst-case maximum p-value for Studentized DIM	**0.0199**	**0.0083**	0.4890	0.1151
	(17) Worst-case maximum p-value for Studentized OLS	**0.0220**	**0.0108**	0.5492	0.1732
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0241**	**0.0108**	0.5514	0.1821
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0313**	**0.0154**	0.5264	0.1184
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0353**	**0.0118**	0.6736	0.2170
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0347**	**0.0190**	0.5707	0.2656
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0311**	**0.0154**	0.5022	0.1203
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0371**	**0.0169**	0.6101	0.2091
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0468**	**0.0169**	0.5743	0.2728
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	**0.0660**	**0.0411**	0.4905	0.2257
	(26) Worst-case maximum p-value for Nonstudentized OLS	**0.0411**	**0.0411**	0.5545	0.3607
	(27) Worst-case maximum p-value for Nonstudentized AIPW	**0.0501**	**0.0411**	0.5629	0.3756
	(28) Worst-case maximum p-value for Studentized DIM	**0.0597**	**0.0332**	0.4890	0.2301
	(29) Worst-case maximum p-value for Studentized OLS	**0.0660**	**0.0431**	0.5492	0.3463
	(30) Worst-case maximum p-value for Studentized AIPW	**0.0724**	**0.0431**	0.5514	0.3641
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	**0.0940**	**0.0617**	0.5264	0.2368
	(32) Worst-case de Haan p-value for Nonstudentized OLS	0.1058	**0.0473**	0.6736	0.4341
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.1041	**0.0762**	0.5707	0.5312
	(34) Worst-case de Haan p-value for Studentized DIM	**0.0934**	**0.0617**	0.5022	0.2406
	(35) Worst-case de Haan p-value for Studentized OLS	0.1113	**0.0675**	0.6101	0.4181
	(36) Worst-case de Haan p-value for Studentized AIPW	0.1403	**0.0675**	0.5743	0.5457

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows $(01) - (03)$ provide the one-sided asymptotic p-values, while rows $(04) - (06)$ provide the bootstrap p-values. Rows $(07) - (09)$ provide the permutation p-values based on nonstudentized test statistics, while rows $(10) - (12)$ provide those based on studentized test statistics. Rows $(13) - (15)$ provide the worst-case maximum p-values based on nonstudentized test statistics, while rows $(16) - (18)$ provide those based on studentized test statistics. Rows $(19) - (21)$ provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows $(22) - (24)$ provide those based on studentized test statistics. Rows $(25) - (36)$ contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows $(13) - (24)$. The block of variables used for Holm stepdown adjustment consists of all four of the above violent crime conviction outcome variables.

Table D-3: Treatment Effects on Violent Crime Conviction Outcomes of the **Female** Participants

	Statistic	Cumulative violent misdemeanor convictions age 30	Cumulative violent misdemeanor convictions age 40	Cumulative violent felony convictions age 30	Cumulative violent felony convictions age 40
Summary	(i) Number of observations	51	50	51	50
	(ii) Mean of the control group	0.4231	0.4400	0.0385	0.0400
	(iii) Mean of the treatment group	0.0000	0.0000	0.0000	0.0000
Estimates	(iv) DIM (difference in means) estimate	-0.4231	-0.4400	-0.0385	-0.0400
	(v) OLS estimate conditional on covariates	-0.4900	-0.5228	-0.0407	-0.0436
	(vi) AIPW (augmented IPW) estimate	-0.4367	-0.4713	-0.0350	-0.0380
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0672**	**0.0670**	0.1602	0.1602
	(02) Asymptotic p-value for OLS estimate	**0.0673**	**0.0692**	0.1662	0.1657
	(03) Asymptotic p-value for AIPW estimate	**0.0601**	**0.0543**	0.1707	0.1630
	(04) Bootstrap p-value for DIM estimate	**0.0663**	**0.0649**	0.1431	0.1434
	(05) Bootstrap p-value for OLS estimate	**0.0760**	**0.0762**	0.1520	0.1529
	(06) Bootstrap p-value for AIPW estimate	**0.0770**	**0.0789**	0.1452	0.1460
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0508**	**0.0428**	0.2028	0.1688
	(08) Permutation p-value for Nonstudentized OLS	**0.0228**	**0.0152**	0.2176	0.1744
	(09) Permutation p-value for Nonstudentized AIPW	**0.0524**	**0.0416**	0.3084	0.2616
	(10) Permutation p-value for Studentized DIM	**0.0452**	**0.0408**	0.2148	0.1992
	(11) Permutation p-value for Studentized OLS	**0.0524**	**0.0672**	0.2936	0.2456
	(12) Permutation p-value for Studentized AIPW	**0.0704**	**0.0568**	0.3724	0.3224
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0706**	**0.0635**	0.3190	0.2931
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0512**	**0.0404**	0.3201	0.2804
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0906**	**0.0759**	0.4065	0.3741
	(16) Worst-case maximum p-value for Studentized DIM	**0.0600**	**0.0583**	0.3418	0.3280
	(17) Worst-case maximum p-value for Studentized OLS	**0.0836**	0.1049	0.4101	0.3684
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0876**	**0.0684**	0.3909	0.3319
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	0.1030	**0.0907**	0.3794	0.4016
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0731**	**0.0507**	0.3301	0.3095
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	0.1588	**0.0926**	0.4395	0.4520
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0852**	**0.0875**	0.3626	0.3805
	(23) Worst-case de Haan p-value for Studentized OLS	0.1688	0.2187	0.5532	0.4033
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0998**	0.1429	0.4841	0.3864
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	0.2538	0.2538	0.5861	0.5861
	(26) Worst-case maximum p-value for Nonstudentized OLS	0.1616	0.1616	0.5609	0.5609
	(27) Worst-case maximum p-value for Nonstudentized AIPW	0.3036	0.3036	0.7483	0.7483
	(28) Worst-case maximum p-value for Studentized DIM	0.2330	0.2330	0.6559	0.6559
	(29) Worst-case maximum p-value for Studentized OLS	0.3344	0.3344	0.7367	0.7367
	(30) Worst-case maximum p-value for Studentized AIPW	0.2736	0.2736	0.6638	0.6638
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	0.3629	0.3629	0.7587	0.7587
	(32) Worst-case de Haan p-value for Nonstudentized OLS	0.2194	0.2028	0.6190	0.6190
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.4763	0.3704	0.8789	0.8789
	(34) Worst-case de Haan p-value for Studentized DIM	0.3408	0.3408	0.7252	0.7252
	(35) Worst-case de Haan p-value for Studentized OLS	0.6751	0.6751	0.8066	0.8066
	(36) Worst-case de Haan p-value for Studentized AIPW	0.3991	0.4288	0.7729	0.7729

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows $(01) - (03)$ provide the one-sided asymptotic p-values, while rows $(04) - (06)$ provide the bootstrap p-values. Rows $(07) - (09)$ provide the permutation p-values based on nonstudentized test statistics, while rows $(10) - (12)$ provide those based on studentized test statistics. Rows $(13) - (15)$ provide the worst-case maximum p-values based on nonstudentized test statistics, while rows $(16) - (18)$ provide those based on studentized test statistics. Rows $(19) - (21)$ provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows $(22) - (24)$ provide those based on studentized test statistics. Rows $(25) - (36)$ contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows $(13) - (24)$. The block of variables used for Holm stepdown adjustment consists of all four of the above violent crime conviction outcome variables.

E Inference on Stanford–Binet (SB) IQ Scores

Table E-1: Treatment Effects on Stanford–Binet IQ Scores of the **Pooled** Participants

	Statistic	SB IQ age 4	SB IQ age 5	SB IQ age 6	SB IQ age 7	SB IQ age 8	SB IQ age 9	SB IQ age 10
Summary	(*i*) Number of observations	123	93	120	119	117	117	114
	(*ii*) Mean of the control group	83.323	83.510	86.344	87.066	86.855	86.770	84.614
	(*iii*) Mean of the treatment group	95.534	94.932	91.250	91.724	88.109	87.714	84.982
Estimates	(*iv*) DIM (difference in means) estimate	12.211	11.422	4.9063	4.6586	1.2543	0.9438	0.3684
	(*v*) OLS estimate conditional on covariates	10.839	11.042	4.1544	3.8600	1.1056	−0.1264	0.1189
	(*vi*) AIPW (augmented IPW) estimate	10.828	11.236	3.9742	3.9689	−0.3057	−0.4453	−0.0770
Single *p*-values	(01) Asymptotic *p*-value for DIM estimate	**0.0000**	**0.0000**	**0.0116**	**0.0189**	0.3024	0.3404	0.4327
	(02) Asymptotic *p*-value for OLS estimate	**0.0000**	**0.0000**	**0.0120**	**0.0367**	0.3132	0.4769	0.4780
	(03) Asymptotic *p*-value for AIPW estimate	**0.0000**	**0.0000**	**0.0097**	**0.0254**	0.4433	0.4111	0.4846
	(04) Bootstrap *p*-value for DIM estimate	**0.0000**	**0.0000**	**0.0079**	**0.0091**	0.2815	0.3339	0.4294
	(05) Bootstrap *p*-value for OLS estimate	**0.0000**	**0.0000**	**0.0070**	**0.0173**	0.2928	0.4755	0.4776
	(06) Bootstrap *p*-value for AIPW estimate	**0.0000**	**0.0000**	**0.0062**	**0.0157**	0.4417	0.4144	0.4857
Single *p*-values	(07) Permutation *p*-value for Nonstudentized DIM	**0.0004**	**0.0008**	**0.0104**	**0.0156**	0.3408	0.3728	0.4360
	(08) Permutation *p*-value for Nonstudentized OLS	**0.0004**	**0.0004**	**0.0140**	**0.0292**	0.3752	0.4248	0.4932
	(09) Permutation *p*-value for Nonstudentized AIPW	**0.0004**	**0.0004**	**0.0200**	**0.0352**	0.3912	0.3704	0.4724
	(10) Permutation *p*-value for Studentized DIM	**0.0004**	**0.0008**	**0.0116**	**0.0164**	0.3452	0.3716	0.4372
	(11) Permutation *p*-value for Studentized OLS	**0.0004**	**0.0004**	**0.0136**	**0.0392**	0.3828	0.4264	0.4936
	(12) Permutation *p*-value for Studentized AIPW	**0.0004**	**0.0004**	**0.0152**	**0.0416**	0.3936	0.3680	0.4712
Single *p*-values	(13) Worst-case maximum *p*-value for Nonstudentized DIM	**0.0025**	**0.0025**	**0.0144**	**0.0251**	0.2735	0.3624	0.4851
	(14) Worst-case maximum *p*-value for Nonstudentized OLS	**0.0025**	**0.0025**	**0.0267**	**0.0483**	0.3546	0.5817	0.6343
	(15) Worst-case maximum *p*-value for Nonstudentized AIPW	**0.0025**	**0.0064**	**0.0295**	**0.0473**	0.6563	0.5538	0.5089
	(16) Worst-case maximum *p*-value for Studentized DIM	**0.0025**	**0.0025**	**0.0159**	**0.0283**	0.2743	0.3624	0.4869
	(17) Worst-case maximum *p*-value for Studentized OLS	**0.0025**	**0.0025**	**0.0267**	**0.0644**	0.3577	0.5817	0.6343
	(18) Worst-case maximum *p*-value for Studentized AIPW	**0.0025**	**0.0029**	**0.0271**	**0.0543**	0.6558	0.5527	0.5049
Single *p*-values	(19) Worst-case de Haan *p*-value for Nonstudentized DIM	**0.0027**	**0.0027**	**0.0228**	**0.0359**	0.2940	0.4063	0.5208
	(20) Worst-case de Haan *p*-value for Nonstudentized OLS	**0.0027**	**0.0027**	**0.0414**	**0.0586**	0.4070	0.6527	0.7045
	(21) Worst-case de Haan *p*-value for Nonstudentized AIPW	**0.0027**	**0.0083**	**0.0396**	**0.0600**	0.6878	0.6122	0.6406
	(22) Worst-case de Haan *p*-value for Studentized DIM	**0.0027**	**0.0027**	**0.0198**	**0.0488**	0.3634	0.3922	0.5235
	(23) Worst-case de Haan *p*-value for Studentized OLS	**0.0027**	**0.0027**	**0.0469**	**0.0964**	0.4134	0.6527	0.7834
	(24) Worst-case de Haan *p*-value for Studentized AIPW	**0.0027**	**0.0036**	**0.0425**	**0.0791**	0.7856	0.5871	0.6972
Holm stepdown *p*-values	(25) Worst-case maximum *p*-value for Nonstudentized DIM	**0.0174**	**0.0174**	**0.0722**	0.1005	0.8204	0.8204	0.8204
	(26) Worst-case maximum *p*-value for Nonstudentized OLS	**0.0174**	**0.0174**	0.1336	0.1934	1.0000	1.0000	1.0000
	(27) Worst-case maximum *p*-value for Nonstudentized AIPW	**0.0174**	**0.0386**	0.1473	0.1891	1.0000	1.0000	1.0000
	(28) Worst-case maximum *p*-value for Studentized DIM	**0.0174**	**0.0174**	**0.0796**	0.1132	0.8229	0.8229	0.8229
	(29) Worst-case maximum *p*-value for Studentized OLS	**0.0174**	**0.0174**	0.1336	0.2574	1.0000	1.0000	1.0000
	(30) Worst-case maximum *p*-value for Studentized AIPW	**0.0174**	**0.0174**	0.1355	0.2170	1.0000	1.0000	1.0000
Holm stepdown *p*-values	(31) Worst-case de Haan *p*-value for Nonstudentized DIM	**0.0187**	**0.0187**	0.1139	0.1435	0.8821	0.8821	0.8821
	(32) Worst-case de Haan *p*-value for Nonstudentized OLS	**0.0187**	**0.0187**	0.2068	0.2345	1.0000	1.0000	1.0000
	(33) Worst-case de Haan *p*-value for Nonstudentized AIPW	**0.0187**	**0.0498**	0.1979	0.2401	1.0000	1.0000	1.0000
	(34) Worst-case de Haan *p*-value for Studentized DIM	**0.0187**	**0.0187**	**0.0990**	0.1950	1.0000	1.0000	1.0000
	(35) Worst-case de Haan *p*-value for Studentized OLS	**0.0187**	**0.0187**	0.2345	0.3856	1.0000	1.0000	1.0000
	(36) Worst-case de Haan *p*-value for Studentized AIPW	**0.0187**	**0.0218**	0.2127	0.3165	1.0000	1.0000	1.0000

Note: Rows (*i*) through (*iii*) contain the summary statistics for the outcomes. Rows (*iv*), (*v*), and (*vi*) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single *p*-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic *p*-values, while rows (04) − (06) provide the bootstrap *p*-values. Rows (07) − (09) provide the permutation *p*-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum *p*-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan *p*-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown *p*-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case *p*-values in rows (13) − (24). The block of variables used for Holm stepdown adjustment consists of all seven of the above Stanford–Binet IQ score variables.

Table E-2: Treatment Effects on Stanford–Binet IQ Scores of the **Male** Participants

	Statistic	SB IQ age 4	SB IQ age 5	SB IQ age 6	SB IQ age 7	SB IQ age 8	SB IQ age 9	SB IQ age 10
Summary	(i) Number of observations	72	54	72	71	67	71	71
	(ii) Mean of the control group	83.077	84.793	85.821	87.711	89.054	89.026	86.026
	(iii) Mean of the treatment group	94.909	95.400	91.485	91.121	88.333	88.394	83.697
Estimates	(iv) DIM (difference in means) estimate	11.832	10.607	5.6643	3.4107	−0.7207	−0.6324	−2.3293
	(v) OLS estimate conditional on covariates	9.0606	9.0654	3.4837	1.7225	−2.2087	−3.1880	−3.9708
	(vi) AIPW (augmented IPW) estimate	8.9881	9.1672	3.0558	1.5758	−3.8293	−4.1669	−4.7216
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0000**	**0.0001**	**0.0189**	**0.0854**	0.4034	0.4088	0.1784
	(02) Asymptotic p-value for OLS estimate	**0.0000**	**0.0003**	**0.0478**	0.2015	0.1932	0.1114	**0.0593**
	(03) Asymptotic p-value for AIPW estimate	**0.0000**	**0.0000**	**0.0557**	0.2040	**0.0512**	**0.0398**	**0.0225**
	(04) Bootstrap p-value for DIM estimate	**0.0000**	**0.0001**	**0.0160**	**0.0808**	0.4020	0.4114	0.1814
	(05) Bootstrap p-value for OLS estimate	**0.0000**	**0.0001**	**0.0338**	0.1951	0.1959	0.1134	**0.0671**
	(06) Bootstrap p-value for AIPW estimate	**0.0000**	**0.0002**	**0.0512**	0.2143	**0.0719**	**0.0577**	**0.0412**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0004**	**0.0008**	**0.0212**	**0.0812**	0.3768	0.3932	0.1880
	(08) Permutation p-value for Nonstudentized OLS	**0.0004**	**0.0004**	**0.0680**	0.2140	0.1696	0.1052	**0.0600**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0004**	**0.0020**	0.1004	0.2388	**0.0652**	**0.0504**	**0.0272**
	(10) Permutation p-value for Studentized DIM	**0.0004**	**0.0008**	**0.0208**	**0.0784**	0.3772	0.3912	0.1860
	(11) Permutation p-value for Studentized OLS	**0.0004**	**0.0012**	**0.0560**	0.2008	0.1560	0.1108	**0.0640**
	(12) Permutation p-value for Studentized AIPW	**0.0004**	**0.0004**	**0.0712**	0.2104	**0.0556**	**0.0472**	**0.0292**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0048**	**0.0053**	**0.0249**	**0.0812**	0.5455	0.5617	0.2606
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0050**	**0.0061**	**0.0789**	0.2103	0.3225	0.1936	0.1036
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0050**	**0.0071**	0.1072	0.2101	0.1636	0.1308	**0.0637**
	(16) Worst-case maximum p-value for Studentized DIM	**0.0048**	**0.0076**	**0.0249**	**0.0782**	0.5480	0.5580	0.2568
	(17) Worst-case maximum p-value for Studentized OLS	**0.0052**	**0.0053**	**0.0716**	0.2058	0.3077	0.2050	0.1111
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0049**	**0.0071**	**0.0872**	0.2002	0.1461	0.1289	**0.0707**
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0053**	**0.0087**	**0.0400**	0.1045	0.6085	0.5940	0.3617
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0076**	**0.0062**	0.1113	0.2816	0.3339	0.2606	0.1456
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0056**	**0.0071**	0.1479	0.2761	0.2641	0.1439	**0.0835**
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0053**	**0.0095**	**0.0306**	0.1185	0.5743	0.6044	0.3509
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0076**	**0.0068**	0.1290	0.2358	0.3318	0.2956	0.1506
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0056**	**0.0071**	0.1229	0.2198	0.2396	0.1457	0.1022
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	**0.0333**	**0.0333**	0.1247	0.3247	1.0000	1.0000	0.7819
	(26) Worst-case maximum p-value for Nonstudentized OLS	**0.0351**	**0.0368**	0.3943	0.5809	0.5809	0.5809	0.4146
	(27) Worst-case maximum p-value for Nonstudentized AIPW	**0.0347**	**0.0425**	0.4288	0.4288	0.4288	0.4288	0.3187
	(28) Worst-case maximum p-value for Studentized DIM	**0.0333**	**0.0456**	0.1247	0.3128	1.0000	1.0000	0.7704
	(29) Worst-case maximum p-value for Studentized OLS	**0.0361**	**0.0361**	0.3580	0.6151	0.6151	0.6151	0.4446
	(30) Worst-case maximum p-value for Studentized AIPW	**0.0346**	**0.0425**	0.3534	0.3866	0.3866	0.3866	0.3534
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	**0.0368**	**0.0522**	0.1999	0.4180	1.0000	1.0000	1.0000
	(32) Worst-case de Haan p-value for Nonstudentized OLS	**0.0458**	**0.0436**	0.5567	0.7818	0.7818	0.7818	0.5823
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0394**	**0.0425**	0.5755	0.5755	0.5755	0.5755	0.4174
	(34) Worst-case de Haan p-value for Studentized DIM	**0.0368**	**0.0572**	0.1531	0.4741	1.0000	1.0000	1.0000
	(35) Worst-case de Haan p-value for Studentized OLS	**0.0479**	**0.0479**	0.6448	0.7074	0.7074	0.7074	0.6448
	(36) Worst-case de Haan p-value for Studentized AIPW	**0.0394**	**0.0425**	0.5112	0.5112	0.5112	0.5112	0.5112

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). The block of variables used for Holm stepdown adjustment consists of all seven of the above Stanford–Binet IQ score variables.

Table E-3: Treatment Effects on Stanford–Binet IQ Scores of the **Female** Participants

	Statistic	SB IQ age 4	SB IQ age 5	SB IQ age 6	SB IQ age 7	SB IQ age 8	SB IQ age 9	SB IQ age 10
Summary	(i) Number of observations	51	39	48	48	50	46	43
	(ii) Mean of the control group	83.692	81.650	87.160	86.000	83.600	83.043	81.789
	(iii) Mean of the treatment group	96.360	94.316	90.913	92.520	87.840	86.739	86.750
Estimates	(iv) DIM (difference in means) estimate	12.668	12.666	3.7530	6.5200	4.2400	3.6957	4.9605
	(v) OLS estimate conditional on covariates	13.328	13.521	5.0710	7.1003	4.8782	4.7271	6.4631
	(vi) AIPW (augmented IPW) estimate	13.425	14.157	5.2707	7.3473	4.6689	4.8087	6.4800
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0000**	**0.0023**	0.1221	**0.0444**	0.1247	0.1466	**0.0918**
	(02) Asymptotic p-value for OLS estimate	**0.0000**	**0.0019**	**0.0514**	**0.0427**	**0.0962**	**0.0670**	**0.0317**
	(03) Asymptotic p-value for AIPW estimate	**0.0000**	**0.0008**	**0.0365**	**0.0313**	0.1144	**0.0633**	**0.0277**
	(04) Bootstrap p-value for DIM estimate	**0.0000**	**0.0013**	0.1139	**0.0181**	**0.0937**	0.1281	**0.0703**
	(05) Bootstrap p-value for OLS estimate	**0.0000**	**0.0007**	**0.0398**	**0.0173**	**0.0637**	**0.0528**	**0.0257**
	(06) Bootstrap p-value for AIPW estimate	**0.0000**	**0.0006**	**0.0281**	**0.0154**	**0.0896**	**0.0679**	**0.0323**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0004**	**0.0044**	0.1188	**0.0444**	0.1360	0.1560	**0.0940**
	(08) Permutation p-value for Nonstudentized OLS	**0.0004**	**0.0048**	**0.0676**	**0.0344**	0.1108	0.1100	**0.0404**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0004**	**0.0040**	**0.0604**	**0.0364**	0.1276	0.1112	**0.0508**
	(10) Permutation p-value for Studentized DIM	**0.0004**	**0.0052**	0.1232	**0.0468**	0.1416	0.1608	**0.0936**
	(11) Permutation p-value for Studentized OLS	**0.0004**	**0.0036**	**0.0632**	**0.0524**	0.1244	**0.0928**	**0.0396**
	(12) Permutation p-value for Studentized AIPW	**0.0004**	**0.0064**	**0.0636**	**0.0564**	0.1704	0.1128	**0.0596**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0025**	**0.0125**	0.1305	**0.0729**	0.1423	0.2126	0.1367
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0025**	**0.0198**	**0.0905**	**0.0666**	0.1378	0.1869	0.1194
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0025**	**0.0227**	**0.0818**	**0.0651**	0.1585	0.1862	0.1451
	(16) Worst-case maximum p-value for Studentized DIM	**0.0052**	**0.0183**	0.1397	**0.0734**	0.1487	0.2134	0.1398
	(17) Worst-case maximum p-value for Studentized OLS	**0.0025**	**0.0185**	**0.0859**	**0.0808**	0.1512	0.1625	0.1050
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0034**	**0.0273**	**0.0820**	**0.0858**	0.2040	0.1992	0.1600
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0027**	**0.0193**	0.1629	0.1156	0.1790	0.2540	0.1971
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0027**	**0.0203**	0.1324	0.1316	0.1684	0.3027	0.1760
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0027**	**0.0271**	**0.0942**	**0.0969**	0.1886	0.3194	0.1669
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0066**	**0.0500**	0.1476	**0.0978**	0.2143	0.2272	0.1748
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0027**	**0.0257**	0.1044	0.1044	0.2096	0.2057	0.1369
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0040**	**0.0382**	**0.0959**	0.1232	0.2141	0.2628	0.1976
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	**0.0174**	**0.0748**	0.5221	0.3645	0.5221	0.5221	0.5221
	(26) Worst-case maximum p-value for Nonstudentized OLS	**0.0174**	0.1187	0.3619	0.3329	0.3619	0.3619	0.3619
	(27) Worst-case maximum p-value for Nonstudentized AIPW	**0.0174**	0.1362	0.3274	0.3256	0.4354	0.4354	0.4354
	(28) Worst-case maximum p-value for Studentized DIM	**0.0362**	0.1095	0.5589	0.3671	0.5589	0.5589	0.5589
	(29) Worst-case maximum p-value for Studentized OLS	**0.0174**	0.1113	0.4042	0.4042	0.4042	0.4042	0.4042
	(30) Worst-case maximum p-value for Studentized AIPW	**0.0239**	0.1637	0.4099	0.4099	0.4801	0.4801	0.4801
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	**0.0187**	0.1156	0.6516	0.5778	0.6516	0.6516	0.6516
	(32) Worst-case de Haan p-value for Nonstudentized OLS	**0.0187**	0.1217	0.6580	0.6580	0.6580	0.6580	0.6580
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0187**	0.1628	0.4708	0.4708	0.5007	0.5007	0.5007
	(34) Worst-case de Haan p-value for Studentized DIM	**0.0460**	0.2999	0.5904	0.4890	0.5904	0.5904	0.5904
	(35) Worst-case de Haan p-value for Studentized OLS	**0.0187**	0.1543	0.5218	0.5218	0.5218	0.5218	0.5218
	(36) Worst-case de Haan p-value for Studentized AIPW	**0.0278**	0.2290	0.4796	0.4929	0.5928	0.5928	0.5928

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows $(01) - (03)$ provide the one-sided asymptotic p-values, while rows $(04) - (06)$ provide the bootstrap p-values. Rows $(07) - (09)$ provide the permutation p-values based on nonstudentized test statistics, while rows $(10) - (12)$ provide those based on studentized test statistics. Rows $(13) - (15)$ provide the worst-case maximum p-values based on nonstudentized test statistics, while rows $(16) - (18)$ provide those based on studentized test statistics. Rows $(19) - (21)$ provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows $(22) - (24)$ provide those based on studentized test statistics. Rows $(25) - (36)$ contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows $(13) - (24)$. The block of variables used for Holm stepdown adjustment consists of all seven of the above Stanford–Binet IQ score variables.

F Inference on Achievement Test Outcomes

Table F-1: Treatment Effects on California Achievement Test Scores of the **Pooled** Participants

	Statistic	CAT reading age 14	CAT arithmetic age 14	CAT language age 14	CAT mechanics age 14	CAT spelling age 14
Summary	(i) Number of observations	95	95	95	95	95
	(ii) Mean of the control group	8.7826	7.6304	7.0435	7.6957	11.217
	(iii) Mean of the treatment group	15.082	14.122	16.633	17.837	23.449
Estimates	(iv) DIM (difference in means) estimate	6.2990	6.4920	9.5892	10.141	12.232
	(v) OLS estimate conditional on covariates	4.7799	4.2086	8.6045	9.2684	10.043
	(vi) AIPW (augmented IPW) estimate	4.1080	4.3935	7.8877	8.6518	9.4312
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0070**	**0.0197**	**0.0004**	**0.0003**	**0.0023**
	(02) Asymptotic p-value for OLS estimate	**0.0403**	**0.0867**	**0.0019**	**0.0014**	**0.0130**
	(03) Asymptotic p-value for AIPW estimate	**0.0448**	**0.0590**	**0.0013**	**0.0009**	**0.0095**
	(04) Bootstrap p-value for DIM estimate	**0.0043**	**0.0145**	**0.0001**	**0.0001**	**0.0009**
	(05) Bootstrap p-value for OLS estimate	**0.0339**	**0.0801**	**0.0013**	**0.0008**	**0.0082**
	(06) Bootstrap p-value for AIPW estimate	**0.0636**	**0.0772**	**0.0022**	**0.0011**	**0.0170**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0104**	**0.0240**	**0.0016**	**0.0012**	**0.0036**
	(08) Permutation p-value for Nonstudentized OLS	**0.0404**	0.1060	**0.0036**	**0.0032**	**0.0092**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0712**	**0.0952**	**0.0056**	**0.0036**	**0.0176**
	(10) Permutation p-value for Studentized DIM	**0.0100**	**0.0212**	**0.0012**	**0.0008**	**0.0032**
	(11) Permutation p-value for Studentized OLS	**0.0480**	**0.0988**	**0.0024**	**0.0020**	**0.0108**
	(12) Permutation p-value for Studentized AIPW	**0.0700**	**0.0920**	**0.0028**	**0.0028**	**0.0140**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0337**	**0.0645**	**0.0092**	**0.0079**	**0.0105**
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0732**	0.1610	**0.0140**	**0.0133**	**0.0247**
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0977**	0.1258	**0.0295**	**0.0221**	**0.0313**
	(16) Worst-case maximum p-value for Studentized DIM	**0.0281**	**0.0519**	**0.0056**	**0.0070**	**0.0084**
	(17) Worst-case maximum p-value for Studentized OLS	**0.0854**	0.1530	**0.0156**	**0.0128**	**0.0275**
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0920**	0.1202	**0.0154**	**0.0187**	**0.0251**
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0605**	0.1258	**0.0121**	**0.0082**	**0.0152**
	(20) Worst-case de Haan p-value for Nonstudentized OLS	0.1131	0.1960	**0.0210**	**0.0212**	**0.0315**
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	0.1105	0.1641	**0.0365**	**0.0307**	**0.0648**
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0573**	**0.0725**	**0.0105**	**0.0099**	**0.0135**
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0973**	0.1821	**0.0172**	**0.0230**	**0.0405**
	(24) Worst-case de Haan p-value for Studentized AIPW	0.1244	0.1324	**0.0195**	**0.0234**	**0.0372**
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	**0.0674**	**0.0674**	**0.0395**	**0.0395**	**0.0395**
	(26) Worst-case maximum p-value for Nonstudentized OLS	0.1465	0.1610	**0.0666**	**0.0666**	**0.0742**
	(27) Worst-case maximum p-value for Nonstudentized AIPW	0.1954	0.1954	0.1179	0.1106	0.1179
	(28) Worst-case maximum p-value for Studentized DIM	**0.0563**	**0.0563**	**0.0278**	**0.0282**	**0.0282**
	(29) Worst-case maximum p-value for Studentized OLS	0.1707	0.1707	**0.0639**	**0.0639**	**0.0826**
	(30) Worst-case maximum p-value for Studentized AIPW	0.1841	0.1841	**0.0770**	**0.0770**	**0.0770**
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	0.1211	0.1258	**0.0485**	**0.0411**	**0.0485**
	(32) Worst-case de Haan p-value for Nonstudentized OLS	0.2262	0.2262	0.1051	0.1051	0.1051
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.2210	0.2210	0.1533	0.1533	0.1943
	(34) Worst-case de Haan p-value for Studentized DIM	0.1146	0.1146	**0.0496**	**0.0496**	**0.0496**
	(35) Worst-case de Haan p-value for Studentized OLS	0.1946	0.1946	**0.0860**	**0.0919**	0.1215
	(36) Worst-case de Haan p-value for Studentized AIPW	0.2489	0.2489	**0.0974**	**0.0974**	0.1117

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). The block of variables used for Holm stepdown adjustment consists of all five of the above California Achievement Test component score variables.

Table F-2: Treatment Effects on California Achievement Test Scores of the **Male** Participants

	Statistic	CAT reading age 14	CAT arithmetic age 14	CAT language age 14	CAT mechanics age 14	CAT spelling age 14
Summary	(i) Number of observations	55	55	55	55	55
	(ii) Mean of the control group	9.0000	8.1071	6.5357	6.9643	11.536
	(iii) Mean of the treatment group	13.926	16.000	14.333	15.556	18.519
Estimates	(iv) DIM (difference in means) estimate	4.9259	7.8929	7.7976	8.5913	6.9828
	(v) OLS estimate conditional on covariates	2.4259	2.8494	5.4333	6.2562	3.9609
	(vi) AIPW (augmented IPW) estimate	1.8150	3.0948	5.0292	5.9789	3.1707
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0748**	**0.0589**	**0.0236**	**0.0193**	**0.0952**
	(02) Asymptotic p-value for OLS estimate	0.2638	0.2814	**0.0947**	**0.0707**	0.2478
	(03) Asymptotic p-value for AIPW estimate	0.2957	0.2410	**0.0815**	**0.0538**	0.2652
	(04) Bootstrap p-value for DIM estimate	**0.0712**	**0.0453**	**0.0168**	**0.0133**	**0.0840**
	(05) Bootstrap p-value for OLS estimate	0.2656	0.2760	**0.0899**	**0.0658**	0.2372
	(06) Bootstrap p-value for AIPW estimate	0.3221	0.2629	**0.0995**	**0.0638**	0.2865
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0768**	**0.0664**	**0.0276**	**0.0176**	**0.0944**
	(08) Permutation p-value for Nonstudentized OLS	0.2456	0.2836	0.1100	**0.0880**	0.2200
	(09) Permutation p-value for Nonstudentized AIPW	0.3012	0.2600	0.1184	**0.0896**	0.2564
	(10) Permutation p-value for Studentized DIM	**0.0752**	**0.0664**	**0.0260**	**0.0156**	**0.0928**
	(11) Permutation p-value for Studentized OLS	0.2716	0.2932	0.1180	**0.0836**	0.2492
	(12) Permutation p-value for Studentized AIPW	0.3112	0.2608	0.1076	**0.0712**	0.2600
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	0.1784	0.1158	**0.0773**	**0.0619**	0.1219
	(14) Worst-case maximum p-value for Nonstudentized OLS	0.3007	0.3290	0.1745	0.1343	0.2522
	(15) Worst-case maximum p-value for Nonstudentized AIPW	0.3350	0.2890	0.1782	0.1463	0.2632
	(16) Worst-case maximum p-value for Studentized DIM	0.1711	0.1191	**0.0748**	**0.0535**	0.1225
	(17) Worst-case maximum p-value for Studentized OLS	0.3143	0.3326	0.1644	0.1244	0.2720
	(18) Worst-case maximum p-value for Studentized AIPW	0.3488	0.2909	0.1771	0.1333	0.2734
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	0.1928	0.1630	0.1051	**0.0832**	0.1328
	(20) Worst-case de Haan p-value for Nonstudentized OLS	0.3873	0.3623	0.2241	0.1672	0.3426
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	0.3706	0.3456	0.2028	0.1516	0.3235
	(22) Worst-case de Haan p-value for Studentized DIM	0.1776	0.1555	0.1284	**0.0681**	0.1503
	(23) Worst-case de Haan p-value for Studentized OLS	0.3475	0.3840	0.1996	0.1390	0.3129
	(24) Worst-case de Haan p-value for Studentized AIPW	0.3823	0.3216	0.2098	0.1467	0.3016
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	0.3474	0.3474	0.3095	0.3095	0.3474
	(26) Worst-case maximum p-value for Nonstudentized OLS	0.7567	0.7567	0.6978	0.6715	0.7567
	(27) Worst-case maximum p-value for Nonstudentized AIPW	0.7897	0.7897	0.7316	0.7316	0.7897
	(28) Worst-case maximum p-value for Studentized DIM	0.3572	0.3572	0.2994	0.2673	0.3572
	(29) Worst-case maximum p-value for Studentized OLS	0.8160	0.8160	0.6577	0.6222	0.8160
	(30) Worst-case maximum p-value for Studentized AIPW	0.8202	0.8202	0.7084	0.6667	0.8202
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	0.4202	0.4202	0.4202	0.4159	0.4202
	(32) Worst-case de Haan p-value for Nonstudentized OLS	1.0000	1.0000	0.8964	0.8360	1.0000
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.9704	0.9704	0.8114	0.7580	0.9704
	(34) Worst-case de Haan p-value for Studentized DIM	0.5135	0.5135	0.5135	0.3407	0.5135
	(35) Worst-case de Haan p-value for Studentized OLS	0.9388	0.9388	0.7985	0.6952	0.9388
	(36) Worst-case de Haan p-value for Studentized AIPW	0.9047	0.9047	0.8390	0.7335	0.9047

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows $(01) - (03)$ provide the one-sided asymptotic p-values, while rows $(04) - (06)$ provide the bootstrap p-values. Rows $(07) - (09)$ provide the permutation p-values based on nonstudentized test statistics, while rows $(10) - (12)$ provide those based on studentized test statistics. Rows $(13) - (15)$ provide the worst-case maximum p-values based on nonstudentized test statistics, while rows $(16) - (18)$ provide those based on studentized test statistics. Rows $(19) - (21)$ provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows $(22) - (24)$ provide those based on studentized test statistics. Rows $(25) - (36)$ contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows $(13) - (24)$. The block of variables used for Holm stepdown adjustment consists of all five of the above California Achievement Test component score variables.

Table F-3: Treatment Effects on California Achievement Test Scores of the **Female** Participants

	Statistic	CAT reading age 14	CAT arithmetic age 14	CAT language age 14	CAT mechanics age 14	CAT spelling age 14
Summary	(*i*) Number of observations	40	40	40	40	40
	(*ii*) Mean of the control group	8.4444	6.8889	7.8333	8.8333	10.722
	(*iii*) Mean of the treatment group	16.500	11.818	19.455	20.636	29.500
Estimates	(*iv*) DIM (difference in means) estimate	8.0556	4.9293	11.621	11.803	18.778
	(*v*) OLS estimate conditional on covariates	7.9323	6.5669	12.637	13.175	18.846
	(*vi*) AIPW (augmented IPW) estimate	7.3451	6.2269	11.923	12.425	18.270
Single *p*-values	(01) Asymptotic *p*-value for DIM estimate	**0.0139**	**0.0621**	**0.0023**	**0.0030**	**0.0017**
	(02) Asymptotic *p*-value for OLS estimate	**0.0110**	**0.0067**	**0.0005**	**0.0007**	**0.0019**
	(03) Asymptotic *p*-value for AIPW estimate	**0.0130**	**0.0102**	**0.0009**	**0.0014**	**0.0017**
	(04) Bootstrap *p*-value for DIM estimate	**0.0087**	**0.0499**	**0.0011**	**0.0013**	**0.0007**
	(05) Bootstrap *p*-value for OLS estimate	**0.0081**	**0.0069**	**0.0004**	**0.0005**	**0.0018**
	(06) Bootstrap *p*-value for AIPW estimate	**0.0128**	**0.0138**	**0.0013**	**0.0015**	**0.0042**
Single *p*-values	(07) Permutation *p*-value for Nonstudentized DIM	**0.0160**	**0.0656**	**0.0048**	**0.0052**	**0.0040**
	(08) Permutation *p*-value for Nonstudentized OLS	**0.0276**	**0.0284**	**0.0020**	**0.0036**	**0.0048**
	(09) Permutation *p*-value for Nonstudentized AIPW	**0.0356**	**0.0540**	**0.0064**	**0.0056**	**0.0068**
	(10) Permutation *p*-value for Studentized DIM	**0.0132**	**0.0596**	**0.0024**	**0.0032**	**0.0036**
	(11) Permutation *p*-value for Studentized OLS	**0.0156**	**0.0084**	**0.0012**	**0.0016**	**0.0028**
	(12) Permutation *p*-value for Studentized AIPW	**0.0268**	**0.0284**	**0.0044**	**0.0072**	**0.0064**
Single *p*-values	(13) Worst-case maximum *p*-value for Nonstudentized DIM	**0.0436**	**0.0976**	**0.0191**	**0.0158**	**0.0163**
	(14) Worst-case maximum *p*-value for Nonstudentized OLS	**0.0648**	**0.0648**	**0.0155**	**0.0155**	**0.0228**
	(15) Worst-case maximum *p*-value for Nonstudentized AIPW	**0.0710**	**0.0987**	**0.0203**	**0.0208**	**0.0249**
	(16) Worst-case maximum *p*-value for Studentized DIM	**0.0413**	0.1000	**0.0111**	**0.0116**	**0.0103**
	(17) Worst-case maximum *p*-value for Studentized OLS	**0.0430**	**0.0216**	**0.0074**	**0.0101**	**0.0116**
	(18) Worst-case maximum *p*-value for Studentized AIPW	**0.0573**	**0.0544**	**0.0178**	**0.0208**	**0.0180**
Single *p*-values	(19) Worst-case de Haan *p*-value for Nonstudentized DIM	**0.0570**	0.1176	**0.0287**	**0.0253**	**0.0212**
	(20) Worst-case de Haan *p*-value for Nonstudentized OLS	**0.0785**	0.1154	**0.0191**	**0.0208**	**0.0418**
	(21) Worst-case de Haan *p*-value for Nonstudentized AIPW	**0.0954**	0.1226	**0.0240**	**0.0268**	**0.0507**
	(22) Worst-case de Haan *p*-value for Studentized DIM	**0.0771**	0.1357	**0.0235**	**0.0287**	**0.0115**
	(23) Worst-case de Haan *p*-value for Studentized OLS	**0.0452**	**0.0312**	**0.0083**	**0.0209**	**0.0174**
	(24) Worst-case de Haan *p*-value for Studentized AIPW	**0.0935**	**0.0710**	**0.0232**	**0.0269**	**0.0253**
Holm stepdown *p*-values	(25) Worst-case maximum *p*-value for Nonstudentized DIM	**0.0873**	**0.0976**	**0.0790**	**0.0790**	**0.0790**
	(26) Worst-case maximum *p*-value for Nonstudentized OLS	0.1295	0.1295	**0.0774**	**0.0774**	**0.0774**
	(27) Worst-case maximum *p*-value for Nonstudentized AIPW	0.1420	0.1420	0.1017	0.1017	0.1017
	(28) Worst-case maximum *p*-value for Studentized DIM	**0.0825**	0.1000	**0.0514**	**0.0514**	**0.0514**
	(29) Worst-case maximum *p*-value for Studentized OLS	**0.0433**	**0.0433**	**0.0372**	**0.0404**	**0.0404**
	(30) Worst-case maximum *p*-value for Studentized AIPW	0.1088	0.1088	**0.0889**	**0.0889**	**0.0889**
Holm stepdown *p*-values	(31) Worst-case de Haan *p*-value for Nonstudentized DIM	0.1139	0.1176	0.1058	0.1058	0.1058
	(32) Worst-case de Haan *p*-value for Nonstudentized OLS	0.1571	0.1571	**0.0956**	**0.0956**	0.1254
	(33) Worst-case de Haan *p*-value for Nonstudentized AIPW	0.1908	0.1908	0.1201	0.1201	0.1522
	(34) Worst-case de Haan *p*-value for Studentized DIM	0.1542	0.1542	**0.0941**	**0.0941**	**0.0577**
	(35) Worst-case de Haan *p*-value for Studentized OLS	**0.0694**	**0.0694**	**0.0415**	**0.0694**	**0.0694**
	(36) Worst-case de Haan *p*-value for Studentized AIPW	0.1421	0.1421	0.1162	0.1162	0.1162

Note: Rows (*i*) through (*iii*) contain the summary statistics for the outcomes. Rows (*iv*), (*v*), and (*vi*) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single *p*-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic *p*-values, while rows (04) − (06) provide the bootstrap *p*-values. Rows (07) − (09) provide the permutation *p*-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum *p*-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan *p*-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown *p*-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case *p*-values in rows (13) − (24). The block of variables used for Holm stepdown adjustment consists of all five of the above California Achievement Test component score variables.

G Inference on Education and Crime Outcomes

Table G-1: Treatment Effects on Education and Crime Outcomes of the **Pooled** Participants

	Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19	Non-juvenile arrests age 40	Crime cost age 40	Total charges age40	Non-victimless charges age 40
Summary	(i) Number of observations	123	123	121	77	123	123	123	123
	(ii) Mean of the control group	0.4000	0.2308	11.079	1.7051	8.8000	582.94	10.000	1.9692
	(iii) Mean of the treatment group	0.6379	0.3276	11.534	2.0826	5.1724	251.18	6.0862	0.8621
Estimates	(iv) DIM (difference in means) estimate	0.2379	0.0968	0.4551	0.3775	−3.6276	−331.76	−3.9138	−1.1072
	(v) OLS estimate conditional on covariates	0.2425	0.1215	0.5667	0.3339	−3.0631	−342.71	−3.2673	−0.9580
	(vi) AIPW (augmented IPW) estimate	0.2449	0.1173	0.5497	0.3763	−3.0832	−285.47	−3.3760	−0.9484
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0066**	0.1073	**0.0456**	**0.0122**	**0.0149**	**0.0565**	**0.0218**	**0.0106**
	(02) Asymptotic p-value for OLS estimate	**0.0069**	**0.0712**	**0.0289**	**0.0419**	**0.0331**	**0.0583**	**0.0441**	**0.0193**
	(03) Asymptotic p-value for AIPW estimate	**0.0051**	**0.0723**	**0.0226**	**0.0094**	**0.0231**	**0.0686**	**0.0295**	**0.0142**
	(04) Bootstrap p-value for DIM estimate	**0.0012**	0.1113	**0.0246**	**0.0034**	**0.0087**	**0.0532**	**0.0149**	**0.0113**
	(05) Bootstrap p-value for OLS estimate	**0.0015**	**0.0716**	**0.0138**	**0.0163**	**0.0162**	**0.0549**	**0.0265**	**0.0167**
	(06) Bootstrap p-value for AIPW estimate	**0.0028**	**0.0887**	**0.0264**	**0.0438**	**0.0132**	**0.0639**	**0.0195**	**0.0137**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0072**	0.1228	**0.0576**	**0.0120**	**0.0176**	**0.0688**	**0.0236**	**0.0116**
	(08) Permutation p-value for Nonstudentized OLS	**0.0064**	**0.0716**	**0.0220**	**0.0252**	**0.0348**	**0.0696**	**0.0500**	**0.0244**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0052**	**0.0740**	**0.0264**	**0.0192**	**0.0340**	0.1132	**0.0448**	**0.0284**
	(10) Permutation p-value for Studentized DIM	**0.0088**	0.1192	**0.0620**	**0.0124**	**0.0184**	**0.0776**	**0.0252**	**0.0108**
	(11) Permutation p-value for Studentized OLS	**0.0072**	**0.0848**	**0.0404**	**0.0424**	**0.0408**	**0.0808**	**0.0572**	**0.0220**
	(12) Permutation p-value for Studentized AIPW	**0.0092**	**0.0888**	**0.0432**	**0.0216**	**0.0368**	0.1028	**0.0448**	**0.0212**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0250**	0.1984	**0.0535**	**0.0354**	**0.0351**	**0.0879**	**0.0433**	**0.0320**
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0269**	0.1388	**0.0410**	**0.0708**	**0.0523**	0.1007	**0.0690**	**0.0528**
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0269**	0.1644	**0.0508**	**0.0565**	**0.0575**	0.1350	**0.0716**	**0.0594**
	(16) Worst-case maximum p-value for Studentized DIM	**0.0269**	0.1950	**0.0561**	**0.0382**	**0.0380**	0.1034	**0.0479**	**0.0350**
	(17) Worst-case maximum p-value for Studentized OLS	**0.0322**	0.1494	**0.0712**	**0.0908**	**0.0623**	0.1025	**0.0827**	**0.0559**
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0418**	0.1931	**0.0681**	**0.0644**	**0.0605**	0.1211	**0.0799**	**0.0501**
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0468**	0.2269	**0.0772**	**0.0570**	0.1216	0.1173	**0.0671**	**0.0705**
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0623**	0.1470	**0.0471**	**0.0837**	**0.0741**	0.1611	0.1105	**0.0710**
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0331**	0.2593	**0.0584**	**0.0833**	**0.0895**	0.2183	0.1183	0.1038
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0416**	0.2268	**0.0671**	**0.0568**	**0.0640**	0.1232	**0.0794**	**0.0473**
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0623**	0.1769	**0.0875**	0.1155	**0.0960**	0.1301	0.1299	**0.0664**
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0803**	0.2499	**0.0981**	**0.0801**	0.1055	0.1560	0.1232	**0.0755**
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	0.1002	0.1984	0.1071	0.1061	0.1278	0.1278	0.1278	0.1278
	(26) Worst-case maximum p-value for Nonstudentized OLS	0.1074	0.1417	0.1231	0.1417	0.2092	0.2092	0.2092	0.2092
	(27) Worst-case maximum p-value for Nonstudentized AIPW	0.1074	0.1644	0.1525	0.1525	0.2299	0.2299	0.2299	0.2299
	(28) Worst-case maximum p-value for Studentized DIM	0.1074	0.1950	0.1145	0.1145	0.1399	0.1399	0.1399	0.1399
	(29) Worst-case maximum p-value for Studentized OLS	0.1287	0.2137	0.2137	0.2137	0.2238	0.2238	0.2238	0.2238
	(30) Worst-case maximum p-value for Studentized AIPW	0.1674	0.1932	0.1932	0.1932	0.2006	0.2006	0.2006	0.2006
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	0.1873	0.2269	0.1873	0.1873	0.2686	0.2686	0.2686	0.2686
	(32) Worst-case de Haan p-value for Nonstudentized OLS	0.1885	0.1885	0.1885	0.1885	0.2838	0.2838	0.2838	0.2838
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.1323	0.2593	0.1751	0.1751	0.3581	0.3581	0.3581	0.3581
	(34) Worst-case de Haan p-value for Studentized DIM	0.1665	0.2268	0.1705	0.1705	0.1920	0.1920	0.1920	0.1891
	(35) Worst-case de Haan p-value for Studentized OLS	0.2490	0.2626	0.2626	0.2626	0.2879	0.2879	0.2879	0.2655
	(36) Worst-case de Haan p-value for Studentized AIPW	0.3203	0.3203	0.3203	0.3203	0.3165	0.3165	0.3165	0.3020

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). The first four outcomes (related to education) comprise one block of variables for Holm stepdown adjustment, and the last four outcomes (related to crime) comprise another block of variables for multiple testing.

Table G-2: Treatment Effects on Education and Crime Outcomes of the **Male** Participants

	Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19	Non-juvenile arrests age 40	Crime cost age 40	Total charges age40	Non-victimless charges age 40
Summary	(i) Number of observations	72	72	72	47	72	72	72	72
	(ii) Mean of the control group	0.5128	0.3333	11.282	1.7942	11.718	775.90	13.385	3.0769
	(iii) Mean of the treatment group	0.4848	0.3939	11.364	1.8138	7.4545	424.68	9.0000	1.4848
Estimates	(iv) DIM (difference in means) estimate	−0.0280	0.0606	0.0816	0.0196	−4.2634	−351.22	−4.3846	−1.5921
	(v) OLS estimate conditional on covariates	0.0210	0.0716	0.1442	−0.0252	−3.8991	−399.30	−4.0113	−1.4684
	(vi) AIPW (augmented IPW) estimate	0.0148	0.0706	0.0874	−0.0354	−3.8947	−313.26	−4.1316	−1.4442
Single p-values	(01) Asymptotic p-value for DIM estimate	0.4112	0.2932	0.3996	0.4637	**0.0394**	0.1328	**0.0658**	**0.0233**
	(02) Asymptotic p-value for OLS estimate	0.4386	0.2843	0.3367	0.4593	**0.0615**	0.1135	**0.0928**	**0.0391**
	(03) Asymptotic p-value for AIPW estimate	0.4550	0.2762	0.3902	0.4366	**0.0461**	0.1376	**0.0678**	**0.0274**
	(04) Bootstrap p-value for DIM estimate	0.4013	0.2925	0.3948	0.4580	**0.0287**	0.1269	**0.0519**	**0.0168**
	(05) Bootstrap p-value for OLS estimate	0.4278	0.2797	0.3246	0.4512	**0.0439**	0.1159	**0.0731**	**0.0295**
	(06) Bootstrap p-value for AIPW estimate	0.4540	0.2886	0.3901	0.4336	**0.0368**	0.1361	**0.0579**	**0.0238**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	0.3936	0.3080	0.4244	0.4832	**0.0468**	0.1576	**0.0720**	**0.0268**
	(08) Permutation p-value for Nonstudentized OLS	0.4672	0.2760	0.3568	0.4352	**0.0668**	0.1276	**0.0996**	**0.0432**
	(09) Permutation p-value for Nonstudentized AIPW	0.4824	0.2776	0.4196	0.4288	**0.0652**	0.1944	**0.0904**	**0.0464**
	(10) Permutation p-value for Studentized DIM	0.3760	0.2976	0.4216	0.4828	**0.0464**	0.1600	**0.0736**	**0.0256**
	(11) Permutation p-value for Studentized OLS	0.4700	0.2960	0.3652	0.4396	**0.0816**	0.1468	0.1076	**0.0412**
	(12) Permutation p-value for Studentized AIPW	0.4868	0.2932	0.4240	0.4328	**0.0668**	0.1764	**0.0920**	**0.0372**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	0.6593	0.3782	0.3696	0.5103	**0.0645**	0.1645	**0.0878**	**0.0599**
	(14) Worst-case maximum p-value for Nonstudentized OLS	0.5449	0.3819	0.4033	0.5448	**0.0881**	0.1675	0.1167	**0.0853**
	(15) Worst-case maximum p-value for Nonstudentized AIPW	0.5569	0.3823	0.4576	0.5367	**0.0913**	0.2149	0.1169	**0.0840**
	(16) Worst-case maximum p-value for Studentized DIM	0.6322	0.3674	0.3659	0.5133	**0.0701**	0.1695	**0.0946**	**0.0623**
	(17) Worst-case maximum p-value for Studentized OLS	0.5498	0.3957	0.4097	0.5448	0.1040	0.1674	0.1377	**0.0853**
	(18) Worst-case maximum p-value for Studentized AIPW	0.5607	0.3984	0.4589	0.5414	0.1019	0.2060	0.1242	**0.0755**
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	0.7578	0.4516	0.4412	0.5375	0.1076	0.1814	0.1275	**0.0631**
	(20) Worst-case de Haan p-value for Nonstudentized OLS	0.6619	0.4632	0.4457	0.5998	**0.0989**	0.2310	0.1352	0.1667
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	0.5786	0.4914	0.5251	0.6408	0.1012	0.2457	0.1334	0.1017
	(22) Worst-case de Haan p-value for Studentized DIM	0.6718	0.4281	0.4031	0.5493	**0.0896**	0.2133	0.1038	**0.0812**
	(23) Worst-case de Haan p-value for Studentized OLS	0.6560	0.4280	0.4667	0.6037	0.1218	0.2418	0.1525	0.1611
	(24) Worst-case de Haan p-value for Studentized AIPW	0.6036	0.4353	0.6118	0.6563	0.1795	0.2329	0.1789	0.1332
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	1.0000	1.0000	1.0000	1.0000	0.2394	0.2394	0.2394	0.2394
	(26) Worst-case maximum p-value for Nonstudentized OLS	1.0000	1.0000	1.0000	1.0000	0.3413	0.3413	0.3413	0.3413
	(27) Worst-case maximum p-value for Nonstudentized AIPW	1.0000	1.0000	1.0000	1.0000	0.3359	0.3359	0.3359	0.3359
	(28) Worst-case maximum p-value for Studentized DIM	1.0000	1.0000	1.0000	1.0000	0.2493	0.2493	0.2493	0.2493
	(29) Worst-case maximum p-value for Studentized OLS	1.0000	1.0000	1.0000	1.0000	0.3413	0.3413	0.3413	0.3413
	(30) Worst-case maximum p-value for Studentized AIPW	1.0000	1.0000	1.0000	1.0000	0.3057	0.3057	0.3057	0.3020
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	1.0000	1.0000	1.0000	1.0000	0.3228	0.3228	0.3228	0.2525
	(32) Worst-case de Haan p-value for Nonstudentized OLS	1.0000	1.0000	1.0000	1.0000	0.3956	0.4055	0.4055	0.4055
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	1.0000	1.0000	1.0000	1.0000	0.4046	0.4046	0.4046	0.4046
	(34) Worst-case de Haan p-value for Studentized DIM	1.0000	1.0000	1.0000	1.0000	0.3249	0.3249	0.3249	0.3249
	(35) Worst-case de Haan p-value for Studentized OLS	1.0000	1.0000	1.0000	1.0000	0.4870	0.4870	0.4870	0.4870
	(36) Worst-case de Haan p-value for Studentized AIPW	1.0000	1.0000	1.0000	1.0000	0.5366	0.5366	0.5366	0.5327

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). The first four outcomes (related to education) comprise one block of variables for Holm stepdown adjustment, and the last four outcomes (related to crime) comprise another block of variables for multiple testing.

Table G-3: Treatment Effects on Education and Crime Outcomes of the **Female** Participants

	Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19	Non-juvenile arrests age 40	Crime cost age 40	Total charges age40	Non-victimless charges age 40
Summary	(i) Number of observations	51	51	49	30	51	51	51	51
	(ii) Mean of the control group	0.2308	0.0769	10.750	1.5269	4.4231	293.50	4.9231	0.3077
	(iii) Mean of the treatment group	0.8400	0.2400	11.760	2.4147	2.1600	22.165	2.2400	0.0400
Estimates	(iv) DIM (difference in means) estimate	0.6092	0.1631	1.0100	0.8878	−2.2631	−271.33	−2.6831	−0.2677
	(v) OLS estimate conditional on covariates	0.5541	0.1735	1.1648	0.9098	−2.1265	−283.60	−2.4812	−0.2749
	(vi) AIPW (augmented IPW) estimate	0.5699	0.1832	1.2024	0.9575	−1.9376	−246.24	−2.3094	−0.2486
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0000**	**0.0450**	**0.0055**	**0.0000**	**0.0457**	0.1365	**0.0300**	**0.0339**
	(02) Asymptotic p-value for OLS estimate	**0.0000**	**0.0337**	**0.0085**	**0.0002**	**0.0691**	0.1451	**0.0497**	**0.0391**
	(03) Asymptotic p-value for AIPW estimate	**0.0000**	**0.0286**	**0.0023**	**0.0000**	**0.0657**	0.1475	**0.0439**	**0.0365**
	(04) Bootstrap p-value for DIM estimate	**0.0000**	**0.0529**	**0.0030**	**0.0000**	**0.0382**	0.1198	**0.0236**	**0.0286**
	(05) Bootstrap p-value for OLS estimate	**0.0000**	**0.0430**	**0.0076**	**0.0003**	**0.0702**	0.1313	**0.0494**	**0.0322**
	(06) Bootstrap p-value for AIPW estimate	**0.0000**	**0.0494**	**0.0106**	**0.0155**	**0.0795**	0.1227	**0.0528**	**0.0263**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0004**	**0.0624**	**0.0088**	**0.0004**	**0.0356**	**0.0648**	**0.0264**	**0.0408**
	(08) Permutation p-value for Nonstudentized OLS	**0.0004**	**0.0512**	**0.0036**	**0.0004**	**0.0616**	0.1560	**0.0436**	**0.0500**
	(09) Permutation p-value for Nonstudentized AIPW	**0.0004**	**0.0412**	**0.0020**	**0.0004**	**0.0868**	0.2276	**0.0640**	**0.0736**
	(10) Permutation p-value for Studentized DIM	**0.0004**	**0.0572**	**0.0144**	**0.0004**	**0.0380**	**0.0188**	**0.0268**	**0.0364**
	(11) Permutation p-value for Studentized OLS	**0.0004**	**0.0396**	**0.0164**	**0.0004**	**0.0624**	**0.0412**	**0.0440**	**0.0452**
	(12) Permutation p-value for Studentized AIPW	**0.0004**	**0.0420**	**0.0120**	**0.0004**	**0.0880**	0.2436	**0.0580**	**0.0612**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0025**	0.1351	**0.0196**	**0.0070**	0.1024	0.1346	**0.0824**	**0.0823**
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0025**	0.1311	**0.0156**	**0.0079**	0.1146	0.2581	0.1079	**0.0864**
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0025**	0.1246	**0.0142**	**0.0092**	0.1509	0.3400	0.1347	0.1132
	(16) Worst-case maximum p-value for Studentized DIM	**0.0050**	0.1233	**0.0249**	**0.0079**	0.1150	**0.0660**	**0.0979**	**0.0693**
	(17) Worst-case maximum p-value for Studentized OLS	**0.0052**	0.1070	**0.0325**	**0.0106**	0.1282	**0.0861**	0.1137	**0.0782**
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0051**	0.1165	**0.0284**	**0.0112**	0.1566	0.2615	0.1366	**0.0853**
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0027**	0.1920	**0.0270**	**0.0086**	0.1246	0.1415	0.1660	**0.0921**
	(20) Worst-case de Haan p-value for Nonstudentized OLS	**0.0027**	0.2024	**0.0234**	**0.0120**	0.1968	0.3237	0.1549	**0.0957**
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0027**	0.1957	**0.0264**	**0.0120**	0.1886	0.3480	0.1583	0.1244
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0075**	0.1489	**0.0474**	**0.0094**	0.1337	**0.0790**	0.1382	**0.0803**
	(23) Worst-case de Haan p-value for Studentized OLS	**0.0073**	0.1243	**0.0532**	**0.0134**	0.1562	0.1204	0.1433	**0.0915**
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0075**	0.2231	**0.0500**	**0.0381**	0.1925	0.3036	0.1407	0.1201
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	**0.0100**	0.1351	**0.0393**	**0.0210**	0.3292	0.3292	0.3292	0.3292
	(26) Worst-case maximum p-value for Nonstudentized OLS	**0.0100**	0.1311	**0.0311**	**0.0237**	0.3458	0.3458	0.3458	0.3458
	(27) Worst-case maximum p-value for Nonstudentized AIPW	**0.0100**	0.1246	**0.0285**	**0.0277**	0.4527	0.4527	0.4527	0.4527
	(28) Worst-case maximum p-value for Studentized DIM	**0.0200**	0.1233	**0.0497**	**0.0237**	0.2641	0.2641	0.2641	0.2641
	(29) Worst-case maximum p-value for Studentized OLS	**0.0208**	0.1070	**0.0651**	**0.0318**	0.3128	0.3128	0.3128	0.3128
	(30) Worst-case maximum p-value for Studentized AIPW	**0.0202**	0.1165	**0.0567**	**0.0336**	0.4098	0.4098	0.4098	0.3413
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	**0.0107**	0.1920	**0.0540**	**0.0257**	0.3738	0.3738	0.3738	0.3684
	(32) Worst-case de Haan p-value for Nonstudentized OLS	**0.0107**	0.2024	**0.0468**	**0.0359**	0.4646	0.4646	0.4646	0.3828
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0107**	0.1957	**0.0528**	**0.0359**	0.4975	0.4975	0.4975	0.4975
	(34) Worst-case de Haan p-value for Studentized DIM	**0.0299**	0.1489	**0.0948**	**0.0299**	0.3160	0.3160	0.3160	0.3160
	(35) Worst-case de Haan p-value for Studentized OLS	**0.0294**	0.1243	0.1064	**0.0401**	0.3660	0.3660	0.3660	0.3660
	(36) Worst-case de Haan p-value for Studentized AIPW	**0.0299**	0.2231	0.1142	0.1142	0.4803	0.4803	0.4803	0.4803

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). The first four outcomes (related to education) comprise one block of variables for Holm stepdown adjustment, and the last four outcomes (related to crime) comprise another block of variables for multiple testing.

H Inference on Employment Outcomes

Table H-1: Treatment Effects on Employment Outcomes of the **Pooled** Participants

	Statistic	Currently employed age 19	Unemp. last year age 19	Jobless months age 19	Currently employed age 27	Unemp. last year age 27	Jobless months age 27	Currently employed age 40	Unemp. last year age 40	Jobless months age 40
Summary	(*i*) Number of observations	123	123	112	116	120	116	112	119	112
	(*ii*) Mean of the control group	0.3077	0.3077	6.0175	0.5574	0.3968	9.3934	0.6207	0.4426	8.5862
	(*iii*) Mean of the treatment group	0.5000	0.2414	5.2545	0.6909	0.2456	5.6364	0.7593	0.2759	5.7963
Estimates	(*iv*) DIM (difference in means) estimate	0.1923	−0.0663	−0.7630	0.1335	−0.1512	−3.7571	0.1386	−0.1668	−2.7899
	(*v*) OLS estimate conditional on covariates	0.1810	−0.0979	−1.2753	0.1580	−0.1554	−3.4659	0.1714	−0.1603	−2.9605
	(*vi*) AIPW (augmented IPW) estimate	0.2093	−0.0868	−0.9401	0.1410	−0.1589	−2.8021	0.1489	−0.1638	−2.7613
Single *p*-values	(01) Asymptotic *p*-value for DIM estimate	**0.0194**	0.2183	0.3284	**0.0854**	**0.0620**	**0.0365**	**0.0541**	**0.0270**	**0.0591**
	(02) Asymptotic *p*-value for OLS estimate	**0.0319**	0.1471	0.2513	**0.0609**	**0.0693**	**0.0548**	**0.0349**	**0.0413**	**0.0523**
	(03) Asymptotic *p*-value for AIPW estimate	**0.0109**	0.1613	0.2946	**0.0531**	**0.0535**	**0.0681**	**0.0443**	**0.0252**	**0.0509**
	(04) Bootstrap *p*-value for DIM estimate	**0.0105**	0.1888	0.3173	**0.0590**	**0.0337**	**0.0177**	**0.0441**	**0.0252**	**0.0577**
	(05) Bootstrap *p*-value for OLS estimate	**0.0195**	0.1152	0.2330	**0.0389**	**0.0384**	**0.0326**	**0.0250**	**0.0388**	**0.0537**
	(06) Bootstrap *p*-value for AIPW estimate	**0.0102**	0.1518	0.3186	**0.0574**	**0.0373**	**0.0687**	**0.0487**	**0.0354**	**0.0707**
Single *p*-values	(07) Permutation *p*-value for Nonstudentized DIM	**0.0180**	0.2256	0.3116	**0.0780**	**0.0612**	**0.0280**	**0.0628**	**0.0320**	**0.0520**
	(08) Permutation *p*-value for Nonstudentized OLS	**0.0216**	0.1256	0.2260	**0.0420**	**0.0504**	**0.0348**	**0.0288**	**0.0352**	**0.0368**
	(09) Permutation *p*-value for Nonstudentized AIPW	**0.0120**	0.1564	0.2912	**0.0608**	**0.0552**	**0.0752**	**0.0576**	**0.0344**	**0.0504**
	(10) Permutation *p*-value for Studentized DIM	**0.0200**	0.2216	0.3116	**0.0796**	**0.0584**	**0.0300**	**0.0556**	**0.0316**	**0.0508**
	(11) Permutation *p*-value for Studentized OLS	**0.0324**	0.1492	0.2500	**0.0596**	**0.0708**	**0.0528**	**0.0440**	**0.0528**	**0.0504**
	(12) Permutation *p*-value for Studentized AIPW	**0.0164**	0.1804	0.3016	**0.0644**	**0.0696**	**0.0796**	**0.0600**	**0.0432**	**0.0596**
Single *p*-values	(13) Worst-case maximum *p*-value for Nonstudentized DIM	**0.0774**	0.2634	0.3600	**0.0900**	0.1166	**0.0729**	**0.0578**	**0.0517**	**0.0584**
	(14) Worst-case maximum *p*-value for Nonstudentized OLS	**0.0977**	0.2195	0.3628	**0.0902**	0.1241	**0.0977**	**0.0555**	**0.0767**	**0.0786**
	(15) Worst-case maximum *p*-value for Nonstudentized AIPW	**0.0722**	0.2441	0.4421	0.1170	0.1382	0.1421	0.1034	**0.0758**	0.1100
	(16) Worst-case maximum *p*-value for Studentized DIM	**0.0777**	0.2634	0.3600	**0.0912**	0.1152	**0.0829**	**0.0565**	**0.0521**	**0.0592**
	(17) Worst-case maximum *p*-value for Studentized OLS	0.1215	0.2369	0.3787	0.1097	0.1578	0.1224	**0.0731**	0.1032	0.1033
	(18) Worst-case maximum *p*-value for Studentized AIPW	**0.0817**	0.2687	0.4579	0.1185	0.1547	0.1613	0.1126	**0.0987**	0.1231
Single *p*-values	(19) Worst-case de Haan *p*-value for Nonstudentized DIM	0.1002	0.3369	0.3847	0.1294	0.1297	0.1769	**0.0631**	**0.0840**	**0.0942**
	(20) Worst-case de Haan *p*-value for Nonstudentized OLS	0.1221	0.2595	0.3784	0.1043	0.1709	0.1296	**0.0754**	0.1162	0.1045
	(21) Worst-case de Haan *p*-value for Nonstudentized AIPW	**0.0889**	0.3025	0.4729	0.1571	0.2218	0.1886	0.1253	**0.0840**	0.1733
	(22) Worst-case de Haan *p*-value for Studentized DIM	**0.0900**	0.3825	0.3927	0.1364	0.1339	0.1345	**0.0718**	**0.0897**	**0.0942**
	(23) Worst-case de Haan *p*-value for Studentized OLS	0.1580	0.3468	0.4086	0.1479	0.1686	0.1619	0.1264	0.1367	0.1212
	(24) Worst-case de Haan *p*-value for Studentized AIPW	0.1165	0.3252	0.4878	0.1629	0.2029	0.1826	0.1528	0.1200	0.1424
Holm stepdown *p*-values	(25) Worst-case maximum *p*-value for Nonstudentized DIM	0.2322	0.5267	0.5267	0.2188	0.2188	0.2188	0.1551	0.1551	0.1551
	(26) Worst-case maximum *p*-value for Nonstudentized OLS	0.2930	0.4391	0.4391	0.2705	0.2705	0.2705	0.1665	0.1665	0.1665
	(27) Worst-case maximum *p*-value for Nonstudentized AIPW	0.2167	0.4883	0.4883	0.3509	0.3509	0.3509	0.2273	0.2273	0.2273
	(28) Worst-case maximum *p*-value for Studentized DIM	0.2330	0.5267	0.5267	0.2488	0.2488	0.2488	0.1563	0.1563	0.1563
	(29) Worst-case maximum *p*-value for Studentized OLS	0.3645	0.4737	0.4737	0.3290	0.3290	0.3290	0.2192	0.2192	0.2192
	(30) Worst-case maximum *p*-value for Studentized AIPW	0.2451	0.5375	0.5375	0.3555	0.3555	0.3555	0.2960	0.2960	0.2960
Holm stepdown *p*-values	(31) Worst-case de Haan *p*-value for Nonstudentized DIM	0.3007	0.6739	0.6739	0.3881	0.3881	0.3881	0.1894	0.1894	0.1894
	(32) Worst-case de Haan *p*-value for Nonstudentized OLS	0.3662	0.5189	0.5189	0.3130	0.3130	0.3130	0.2262	0.2262	0.2262
	(33) Worst-case de Haan *p*-value for Nonstudentized AIPW	0.2666	0.6050	0.6050	0.4714	0.4714	0.4714	0.2521	0.2521	0.2521
	(34) Worst-case de Haan *p*-value for Studentized DIM	0.2701	0.7650	0.7650	0.4018	0.4018	0.4018	0.2155	0.2155	0.2155
	(35) Worst-case de Haan *p*-value for Studentized OLS	0.4739	0.6936	0.6936	0.4438	0.4438	0.4438	0.3637	0.3637	0.3637
	(36) Worst-case de Haan *p*-value for Studentized AIPW	0.3494	0.6504	0.6504	0.4887	0.4887	0.4887	0.3601	0.3601	0.3601

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single *p*-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic *p*-values, while rows (04) − (06) provide the bootstrap *p*-values. Rows (07) − (09) provide the permutation *p*-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum *p*-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan *p*-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown *p*-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case *p*-values in rows (13) − (24). Three blocks of outcomes are used for Holm stepdown adjustment. Each block of variables for multiple testing consists of three employment variables measured at the same age.

Table H-2: Treatment Effects on Employment Outcomes of the **Male** Participants

	Statistic	Currently employed age 19	Unemp. last year age 19	Jobless months age 19	Currently employed age 27	Unemp. last year age 27	Jobless months age 27	Currently employed age 40	Unemp. last year age 40	Jobless months age 40
Summary	(i) Number of observations	72	72	70	69	72	69	66	72	66
	(ii) Mean of the control group	0.4103	0.1282	3.8158	0.5641	0.3077	8.7949	0.5000	0.4615	10.750
	(iii) Mean of the treatment group	0.5455	0.2424	5.2813	0.6000	0.2424	5.1333	0.7000	0.3636	7.2333
Estimates	(iv) DIM (difference in means) estimate	0.1352	0.1142	1.4655	0.0359	−0.0653	−3.6615	0.2000	−0.0979	−3.5167
	(v) OLS estimate conditional on covariates	0.1265	0.0881	1.4042	0.0839	−0.0682	−3.7638	0.2879	−0.1388	−4.7120
	(vi) AIPW (augmented IPW) estimate	0.1471	0.1021	1.3665	0.0889	−0.0810	−3.8675	0.2655	−0.1425	−4.7578
Single p-values	(01) Asymptotic p-value for DIM estimate	0.1340	0.1277	0.2378	0.3851	0.2778	**0.0631**	**0.0520**	0.1959	**0.0765**
	(02) Asymptotic p-value for OLS estimate	0.1700	0.2261	0.2673	0.2525	0.2832	**0.0648**	**0.0119**	0.1173	**0.0285**
	(03) Asymptotic p-value for AIPW estimate	0.1263	0.1817	0.2572	0.2156	0.2238	**0.0438**	**0.0089**	**0.0843**	**0.0154**
	(04) Bootstrap p-value for DIM estimate	0.1181	0.1003	0.2104	0.3819	0.2731	**0.0585**	**0.0363**	0.1970	**0.0733**
	(05) Bootstrap p-value for OLS estimate	0.1527	0.1985	0.2422	0.2477	0.2721	**0.0566**	**0.0064**	0.1159	**0.0256**
	(06) Bootstrap p-value for AIPW estimate	0.1315	0.1827	0.2501	0.2259	0.2190	**0.0430**	**0.0075**	**0.0957**	**0.0200**
Single p-values	(07) Permutation p-value for Nonstudentized DIM	0.1364	0.1456	0.2592	0.4112	0.2984	**0.0680**	**0.0676**	0.2004	**0.0784**
	(08) Permutation p-value for Nonstudentized OLS	0.1444	0.2148	0.2840	0.2644	0.2868	**0.0644**	**0.0112**	0.1020	**0.0188**
	(09) Permutation p-value for Nonstudentized AIPW	0.1152	0.1800	0.2848	0.2476	0.2464	**0.0580**	**0.0256**	**0.0920**	**0.0188**
	(10) Permutation p-value for Studentized DIM	0.1384	0.1408	0.2628	0.3988	0.2884	**0.0668**	**0.0620**	0.1948	**0.0776**
	(11) Permutation p-value for Studentized OLS	0.1572	0.2512	0.2952	0.2764	0.2976	**0.0736**	**0.0164**	0.1208	**0.0272**
	(12) Permutation p-value for Studentized AIPW	0.1292	0.2148	0.2928	0.2452	0.2388	**0.0588**	**0.0204**	**0.0912**	**0.0188**
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	0.2776	0.2635	0.3316	0.3731	0.3316	**0.0908**	**0.0594**	0.1762	**0.0702**
	(14) Worst-case maximum p-value for Nonstudentized OLS	0.3017	0.3108	0.3201	0.3465	0.3811	0.1158	**0.0433**	0.1885	**0.0605**
	(15) Worst-case maximum p-value for Nonstudentized AIPW	0.2686	0.2534	0.3355	0.3325	0.3453	0.1100	**0.0504**	0.1613	**0.0537**
	(16) Worst-case maximum p-value for Studentized DIM	0.2809	0.2486	0.3291	0.3629	0.3147	**0.0965**	**0.0566**	0.1684	**0.0693**
	(17) Worst-case maximum p-value for Studentized OLS	0.3144	0.3286	0.3253	0.3699	0.3911	0.1348	**0.0471**	0.2098	**0.0790**
	(18) Worst-case maximum p-value for Studentized AIPW	0.2989	0.3050	0.3371	0.3348	0.3460	0.1193	**0.0444**	0.1629	**0.0632**
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	0.3108	0.2912	0.3754	0.4418	0.3866	0.1143	**0.0725**	0.2486	**0.0827**
	(20) Worst-case de Haan p-value for Nonstudentized OLS	0.3503	0.3140	0.3572	0.3917	0.4651	0.1550	**0.0648**	0.2044	**0.0964**
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	0.3528	0.3160	0.4284	0.3555	0.3619	0.1301	0.1013	0.1697	**0.0841**
	(22) Worst-case de Haan p-value for Studentized DIM	0.3864	0.3673	0.4019	0.4130	0.3639	0.1265	**0.0853**	0.2494	**0.0957**
	(23) Worst-case de Haan p-value for Studentized OLS	0.4433	0.3443	0.3914	0.4241	0.5520	0.2210	**0.0662**	0.2274	**0.0893**
	(24) Worst-case de Haan p-value for Studentized AIPW	0.3351	0.4203	0.4217	0.3703	0.3776	0.2030	**0.0640**	0.1671	**0.0722**
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	0.7906	0.7906	0.7906	0.6631	0.6631	0.2724	0.1782	0.1782	0.1782
	(26) Worst-case maximum p-value for Nonstudentized OLS	0.9051	0.9051	0.9051	0.6931	0.6931	0.3473	0.1298	0.1885	0.1298
	(27) Worst-case maximum p-value for Nonstudentized AIPW	0.7602	0.7602	0.7602	0.6651	0.6651	0.3300	0.1512	0.1613	0.1512
	(28) Worst-case maximum p-value for Studentized DIM	0.7459	0.7459	0.7459	0.6294	0.6294	0.2894	0.1697	0.1697	0.1697
	(29) Worst-case maximum p-value for Studentized OLS	0.9433	0.9433	0.9433	0.7397	0.7397	0.4043	0.1413	0.2098	0.1580
	(30) Worst-case maximum p-value for Studentized AIPW	0.8968	0.8968	0.8968	0.6697	0.6697	0.3580	0.1333	0.1629	0.1333
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	0.8737	0.8737	0.8737	0.7732	0.7732	0.3430	0.2174	0.2486	0.2174
	(32) Worst-case de Haan p-value for Nonstudentized OLS	0.9419	0.9419	0.9419	0.7834	0.7834	0.4651	0.1943	0.2044	0.1943
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.9481	0.9481	0.9481	0.7110	0.7110	0.3904	0.2524	0.2524	0.2524
	(34) Worst-case de Haan p-value for Studentized DIM	1.0000	1.0000	1.0000	0.7278	0.7278	0.3795	0.2559	0.2559	0.2559
	(35) Worst-case de Haan p-value for Studentized OLS	1.0000	1.0000	1.0000	0.8482	0.8482	0.6630	0.1987	0.2274	0.1987
	(36) Worst-case de Haan p-value for Studentized AIPW	1.0000	1.0000	1.0000	0.7405	0.7405	0.6089	0.1920	0.1920	0.1920

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). Three blocks of outcomes are used for Holm stepdown adjustment. Each block of variables for multiple testing consists of three employment variables measured at the same age.

Table H-3: Treatment Effects on Employment Outcomes of the **Female** Participants

	Statistic	Currently employed age 19	Unemp. last year age 19	Jobless months age 19	Currently employed age 27	Unemp. last year age 27	Jobless months age 27	Currently employed age 40	Unemp. last year age 40	Jobless months age 40
Summary	(i) Number of observations	51	51	42	47	48	47	46	47	46
	(ii) Mean of the control group	0.1538	0.5769	10.421	0.5455	0.5417	10.455	0.8182	0.4091	5.0455
	(iii) Mean of the treatment group	0.4400	0.2400	5.2174	0.8000	0.2500	6.2400	0.8333	0.1600	4.0000
Estimates	(iv) DIM (difference in means) estimate	0.2862	−0.3369	−5.2037	0.2545	−0.2917	−4.2145	0.0152	−0.2491	−1.0455
	(v) OLS estimate conditional on covariates	0.2618	−0.3526	−4.3561	0.2341	−0.2839	−2.3964	−0.0242	−0.2096	−0.0475
	(vi) AIPW (augmented IPW) estimate	0.2971	−0.3536	−4.1965	0.2145	−0.2689	−1.2981	−0.0157	−0.1938	0.0574
Single p-values	(01) Asymptotic p-value for DIM estimate	**0.0144**	**0.0083**	**0.0514**	**0.0456**	**0.0309**	0.1040	0.4474	**0.0308**	0.3341
	(02) Asymptotic p-value for OLS estimate	**0.0216**	**0.0058**	**0.0896**	**0.0627**	**0.0528**	0.2449	0.4346	**0.0789**	0.4931
	(03) Asymptotic p-value for AIPW estimate	**0.0054**	**0.0029**	**0.0723**	**0.0471**	**0.0523**	0.3328	0.4536	**0.0807**	0.4910
	(04) Bootstrap p-value for DIM estimate	**0.0061**	**0.0048**	**0.0492**	**0.0199**	**0.0125**	**0.0669**	0.4419	**0.0245**	0.3268
	(05) Bootstrap p-value for OLS estimate	**0.0130**	**0.0043**	**0.0981**	**0.0386**	**0.0298**	0.2218	0.4309	**0.0839**	0.4934
	(06) Bootstrap p-value for AIPW estimate	**0.0048**	**0.0033**	0.1386	**0.0604**	**0.0457**	0.3449	0.4586	0.1079	0.4927
Single p-values	(07) Permutation p-value for Nonstudentized DIM	**0.0156**	**0.0120**	**0.0528**	**0.0384**	**0.0284**	**0.0740**	0.4380	**0.0300**	0.3104
	(08) Permutation p-value for Nonstudentized OLS	**0.0264**	**0.0088**	**0.0784**	**0.0452**	**0.0416**	0.1832	0.4580	**0.0728**	0.4520
	(09) Permutation p-value for Nonstudentized AIPW	**0.0136**	**0.0100**	**0.0908**	**0.0584**	**0.0532**	0.2856	0.4812	**0.0960**	0.4712
	(10) Permutation p-value for Studentized DIM	**0.0164**	**0.0152**	**0.0600**	**0.0416**	**0.0352**	**0.0828**	0.4356	**0.0412**	0.3120
	(11) Permutation p-value for Studentized OLS	**0.0352**	**0.0120**	**0.0988**	**0.0608**	**0.0604**	0.2076	0.4692	0.1008	0.4524
	(12) Permutation p-value for Studentized AIPW	**0.0152**	**0.0104**	0.1140	**0.0648**	**0.0728**	0.2916	0.4912	0.1324	0.4700
Single p-values	(13) Worst-case maximum p-value for Nonstudentized DIM	**0.0603**	**0.0303**	**0.0762**	**0.0712**	**0.0969**	0.1815	0.4786	**0.0747**	0.4359
	(14) Worst-case maximum p-value for Nonstudentized OLS	**0.0785**	**0.0291**	0.1190	**0.0770**	0.1146	0.3270	0.6185	0.1294	0.5977
	(15) Worst-case maximum p-value for Nonstudentized AIPW	**0.0572**	**0.0291**	0.1521	**0.0957**	0.1312	0.4494	0.6534	0.1403	0.6114
	(16) Worst-case maximum p-value for Studentized DIM	**0.0518**	**0.0400**	**0.0807**	**0.0712**	0.1026	0.1881	0.4786	**0.0825**	0.4386
	(17) Worst-case maximum p-value for Studentized OLS	**0.0789**	**0.0356**	0.1600	**0.0963**	0.1499	0.3494	0.6242	0.1508	0.5980
	(18) Worst-case maximum p-value for Studentized AIPW	**0.0585**	**0.0313**	0.2020	**0.0960**	0.1663	0.4526	0.6550	0.1934	0.6114
Single p-values	(19) Worst-case de Haan p-value for Nonstudentized DIM	**0.0733**	**0.0579**	0.1049	**0.0889**	0.1130	0.2224	0.5159	0.1005	0.4756
	(20) Worst-case de Haan p-value for Nonstudentized OLS	0.1219	**0.0401**	0.2083	**0.0985**	0.1639	0.3744	0.6784	0.1521	0.7964
	(21) Worst-case de Haan p-value for Nonstudentized AIPW	**0.0994**	**0.0371**	0.2005	0.1518	0.1476	0.5414	0.7685	0.1670	0.6406
	(22) Worst-case de Haan p-value for Studentized DIM	**0.0705**	**0.0581**	0.1044	**0.0867**	0.1325	0.2449	0.5336	0.1250	0.4850
	(23) Worst-case de Haan p-value for Studentized OLS	0.1063	**0.0548**	0.1900	0.1510	0.1829	0.4063	0.6442	0.2208	0.6767
	(24) Worst-case de Haan p-value for Studentized AIPW	**0.0619**	**0.0377**	0.2780	0.1237	0.2242	0.7373	0.6802	0.2386	0.6379
Holm stepdown p-values	(25) Worst-case maximum p-value for Nonstudentized DIM	0.1205	**0.0909**	0.1205	0.2137	0.2137	0.2137	0.8718	0.2241	0.8718
	(26) Worst-case maximum p-value for Nonstudentized OLS	0.1571	**0.0872**	0.1571	0.2309	0.2309	0.3270	1.0000	0.3883	1.0000
	(27) Worst-case maximum p-value for Nonstudentized AIPW	0.1144	**0.0874**	0.1521	0.2870	0.2870	0.4494	1.0000	0.4208	1.0000
	(28) Worst-case maximum p-value for Studentized DIM	0.1200	0.1200	0.1200	0.2137	0.2137	0.2137	0.8773	0.2475	0.8773
	(29) Worst-case maximum p-value for Studentized OLS	0.1579	0.1067	0.1600	0.2889	0.2998	0.3494	1.0000	0.4523	1.0000
	(30) Worst-case maximum p-value for Studentized AIPW	0.1171	**0.0939**	0.2020	0.2879	0.3325	0.4526	1.0000	0.5803	1.0000
Holm stepdown p-values	(31) Worst-case de Haan p-value for Nonstudentized DIM	0.1737	0.1737	0.1737	0.2667	0.2667	0.2667	0.9511	0.3014	0.9511
	(32) Worst-case de Haan p-value for Nonstudentized OLS	0.2437	0.1202	0.2437	0.2955	0.3278	0.3744	1.0000	0.4564	1.0000
	(33) Worst-case de Haan p-value for Nonstudentized AIPW	0.1987	0.1112	0.2005	0.4429	0.4429	0.5414	1.0000	0.5011	1.0000
	(34) Worst-case de Haan p-value for Studentized DIM	0.1742	0.1742	0.1742	0.2600	0.2650	0.2650	0.9701	0.3750	0.9701
	(35) Worst-case de Haan p-value for Studentized OLS	0.2126	0.1644	0.2126	0.4530	0.4530	0.4530	1.0000	0.6625	1.0000
	(36) Worst-case de Haan p-value for Studentized AIPW	0.1237	0.1132	0.2780	0.3712	0.4485	0.7373	1.0000	0.7157	1.0000

Note: Rows (i) through (iii) contain the summary statistics for the outcomes. Rows (iv), (v), and (vi) contain the difference-in-means (DIM) estimates of treatment effects, ordinary least squares (OLS) estimates (conditional on pre-program covariates, i.e., participant's IQ, SES, gender, and mother's working status at baseline), and the augmented inverse probability weighting (AIPW) estimates (accounting for non-response and imbalance in pre-program covariates between the experimental groups), respectively. Rows (01) through (24) contain various single p-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) − (03) provide the one-sided asymptotic p-values, while rows (04) − (06) provide the bootstrap p-values. Rows (07) − (09) provide the permutation p-values based on nonstudentized test statistics, while rows (10) − (12) provide those based on studentized test statistics. Rows (13) − (15) provide the worst-case maximum p-values based on nonstudentized test statistics, while rows (16) − (18) provide those based on studentized test statistics. Rows (19) − (21) provide the worst-case de Haan p-values based the nonstudentized test statistics, while rows (22) − (24) provide those based on studentized test statistics. Rows (25) − (36) contain Holm stepdown p-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case p-values in rows (13) − (24). Three blocks of outcomes are used for Holm stepdown adjustment. Each block of variables for multiple testing consists of three employment variables measured at the same age.

References

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.

Cunha, F., J. J. Heckman, L. Lochner, and D. V. Masterov (2006). Interpreting the evidence on life cycle skill formation. *Handbook of the Economics of Education 1*, 697–812.

Elango, S., J. L. García, J. J. Heckman, and A. Hojman (2015). Early childhood education. In *Economics of Means-Tested Transfer Programs in the United States*, Volume 2, pp. 235–297. University of Chicago Press.

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

Schweinhart, L. J., H. V. Barnes, D. P. Weikart, W. Barnett, and A. Epstein (1993). *Significant benefits: The High/Scope Perry Preschool Study through age 27 (Monographs of the High/Scope Educational Research Foundation, 10)*. Ypsilanti, MI: HighScope Educational Research Foundation.

Weikart, D. P., J. T. Bond, and J. T. McNeil (1978). *The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results through fourth grade*. Number 3. Ypsilanti, MI: HighScope Educational Research Foundation.