

Using a satisfying model of experimenter decision-making to guide finite-sample inference for compromised experiments: Online Appendix

JAMES J. HECKMAN[†] AND GANESH KARAPAKULA[‡]

[†]*Center for the Economics of Human Development, The University of Chicago, Chicago, IL 60637, USA.*

Email: jjh@uchicago.edu

[‡]*Yale University, New Haven, CT 06511, USA.*

Email: ganesh.karapakula@yale.edu

APPENDIX S1. COMPARISON OF INFERENCE FROM ALTERNATIVE APPROACHES

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

In all of the following tables, 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-programme 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 nonresponse and imbalance in pre-programme covariates between the experimental groups), respectively. Rows (01) to (24) contain various single *p*-values (unadjusted for multiplicity of hypotheses) corresponding to these estimates. Rows (01) to (03) provide the one-sided asymptotic *p*-values, while rows (04) to (06) provide the bootstrap *p*-values. Rows (07) to (09) provide the permutation *p*-values based on nonstudentized test statistics, while rows (10) to (12) provide those based on studentized test statistics. Rows (13) to (15) provide the worst-case maximum *p*-values based on nonstudentized test statistics, while rows (16) to (18) provide those based on studentized test statistics. Rows (19) to (21) provide the worst-case de Haan *p*-values based the nonstudentized test statistics, while rows (22) to (24) provide those based on studentized test statistics. Rows (25) to (36) contain Holm stepdown *p*-values (adjusted for multiplicity of hypotheses) corresponding to the unadjusted single worst-case *p*-values in rows (13) to (24). In the tables, the nonstudentized DIM, OLS, and AIPW test statistics are denoted by N-DIM, N-OLS, and N-AIPW, respectively, while the studentized DIM, OLS, and AIPW test statistics are denoted by S-DIM, S-OLS, and S-AIPW, respectively.

APPENDIX S2. INFERENCE ON CONVICTIONS FOR VIOLENT CRIMES

Table S2.1. Treatment effects on violent crime conviction outcomes of the Pooled participants.

Statistic	Violent misdemeanor convictions by:			Violent felony convictions by: age 30
	age 30	age 40	age 40	
(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
(iv) DIM (difference in means) estimate	-0.4714	-0.5948	-0.0950	-0.2832
(v) OLS estimate (with covariates)	-0.5783	-0.7009	-0.0565	-0.2169
(vi) AIPW (augmented IPW) estimate	-0.5300	-0.6491	-0.0561	-0.2052
(01) Asymptotic <i>p</i> -value for DIM	0.0109	0.0033	0.2301	0.0333
(02) Asymptotic <i>p</i> -value for OLS	0.0097	0.0038	0.3248	0.0676
(03) Asymptotic <i>p</i> -value for AIPW	0.0064	0.0021	0.3174	0.0664
(04) Bootstrap <i>p</i> -value for DIM	0.0021	0.0005	0.2263	0.0332
(05) Bootstrap <i>p</i> -value for OLS	0.0017	0.0006	0.3217	0.0708
(06) Bootstrap <i>p</i> -value for AIPW	0.0020	0.0010	0.3217	0.0778
(07) Permutation <i>p</i> -value for N-DIM	0.0036	0.0008	0.2648	0.0392
(08) Permutation <i>p</i> -value for N-OLS	0.0004	0.0004	0.3604	0.0704
(09) Permutation <i>p</i> -value for N-AIPW	0.0016	0.0004	0.3556	0.0792
(10) Permutation <i>p</i> -value for S-DIM	0.0036	0.0004	0.2624	0.0384
(11) Permutation <i>p</i> -value for S-OLS	0.0028	0.0004	0.3552	0.0680
(12) Permutation <i>p</i> -value for S-AIPW	0.0024	0.0008	0.3488	0.0708
(13) Worst-case maximum <i>p</i> for N-DIM	0.0124	0.0083	0.3871	0.0758
(14) Worst-case maximum <i>p</i> for N-OLS	0.0069	0.0050	0.4896	0.1425
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0112	0.0064	0.4840	0.1597
(16) Worst-case maximum <i>p</i> for S-DIM	0.0135	0.0054	0.3867	0.0792
(17) Worst-case maximum <i>p</i> for S-OLS	0.0135	0.0054	0.4836	0.1265
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0102	0.0051	0.4809	0.1376

Table S2.1. Continued

Statistic	Violent misdemeanor convictions by:		Violent felony convictions by: age 30
	age 30	age 40	
(19) Worst-case de Haan p for N-DIM	0.0286	0.0094	0.6647
(20) Worst-case de Haan p for N-OLS	0.0129	0.0050	0.6792
(21) Worst-case de Haan p for N-AIPW	0.1086	0.0073	0.6537
(22) Worst-case de Haan p for S-DIM	0.1002	0.0092	0.6691
(23) Worst-case de Haan p for S-OLS	0.1002	0.0267	0.9861
(24) Worst-case de Haan p for S-AIPW	0.0267	0.0052	0.7310
(25) Worst-case maximum p for N-DIM	0.0371	0.0330	0.3871
(26) Worst-case maximum p for N-OLS	0.0208	0.0200	0.4896
(27) Worst-case maximum p for N-AIPW	0.0336	0.0255	0.4840
(28) Worst-case maximum p for S-DIM	0.0405	0.0215	0.3867
(29) Worst-case maximum p for S-OLS	0.0405	0.0215	0.4836
(30) Worst-case maximum p for S-AIPW	0.0306	0.0204	0.4809
(31) Worst-case de Haan p for N-DIM	0.0857	0.0378	0.6647
(32) Worst-case de Haan p for N-OLS	0.0387	0.0201	0.6792
(33) Worst-case de Haan p for N-AIPW	0.3258	0.0293	0.6537
(34) Worst-case de Haan p for S-DIM	0.3005	0.0366	0.6691
(35) Worst-case de Haan p for S-OLS	0.3005	0.1066	0.9861
(36) Worst-case de Haan p for S-AIPW	0.0800	0.0208	0.7310

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above violent crime conviction outcome variables.

Table S2.2. Treatment effects on violent crime conviction outcomes of the **Male** participants.

Statistic	Violent misdemeanor convictions by:			Violent felony convictions by: age 30 age 40
	age 30	age 40	age 30	
(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	0.1562	0.3333	0.3438
(iv) DIM (difference in means) estimate	-0.4988	-0.6859	-0.1154	-0.4194
(v) OLS estimate (with covariates)	-0.6525	-0.8529	-0.0694	-0.3420
(vi) AIPW (augmented IPW) estimate	-0.5961	-0.7750	-0.0711	-0.3236
(01) Asymptotic <i>p</i> -value for DIM	0.0081	0.0019	0.2915	0.0486
(02) Asymptotic <i>p</i> -value for OLS	0.0062	0.0025	0.3742	0.0892
(03) Asymptotic <i>p</i> -value for AIPW	0.0036	0.0013	0.3618	0.0831
(04) Bootstrap <i>p</i> -value for DIM	0.0061	0.0013	0.2863	0.0439
(05) Bootstrap <i>p</i> -value for OLS	0.0052	0.0016	0.3719	0.0899
(06) Bootstrap <i>p</i> -value for AIPW	0.0044	0.0019	0.3650	0.0936
(07) Permutation <i>p</i> -value for N-DIM	0.0116	0.0040	0.3256	0.0512
(08) Permutation <i>p</i> -value for N-OLS	0.0012	0.0004	0.4016	0.0880
(09) Permutation <i>p</i> -value for N-AIPW	0.0028	0.0008	0.3956	0.1000
(10) Permutation <i>p</i> -value for S-DIM	0.0104	0.0028	0.3208	0.0508
(11) Permutation <i>p</i> -value for S-OLS	0.0072	0.0020	0.3980	0.0880
(12) Permutation <i>p</i> -value for S-AIPW	0.0060	0.0016	0.3876	0.0928
(13) Worst-case maximum <i>p</i> for N-DIM	0.0235	0.0061	0.4529	0.0973
(14) Worst-case maximum <i>p</i> for N-OLS	0.0112	0.0071	0.5311	0.1752
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0148	0.0075	0.5424	0.1804
(16) Worst-case maximum <i>p</i> for S-DIM	0.0223	0.0086	0.4521	0.1025
(17) Worst-case maximum <i>p</i> for S-OLS	0.0223	0.0100	0.5265	0.1549
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0171	0.0132	0.5342	0.1655

Table S2.2. Continued

Statistic	Violent misdemeanor convictions by:		Violent felony convictions by: age 30
	age 30	age 40	
(19) Worst-case de Haan p for N-DIM	0.1402	0.0117	1.0000
(20) Worst-case de Haan p for N-OLS	0.1430	0.0125	0.8846
(21) Worst-case de Haan p for N-AIPW	0.0378	0.0125	0.7528
(22) Worst-case de Haan p for S-DIM	0.1278	0.0310	1.0000
(23) Worst-case de Haan p for S-OLS	0.0521	0.0391	0.8638
(24) Worst-case de Haan p for S-AIPW	0.0631	0.0263	0.8923
(25) Worst-case maximum p for N-DIM	0.0706	0.0245	0.4529
(26) Worst-case maximum p for N-OLS	0.0336	0.0285	0.5311
(27) Worst-case maximum p for N-AIPW	0.0443	0.0299	0.5424
(28) Worst-case maximum p for S-DIM	0.0668	0.0346	0.4521
(29) Worst-case maximum p for S-OLS	0.0668	0.0398	0.5265
(30) Worst-case maximum p for S-AIPW	0.0528	0.0528	0.5342
(31) Worst-case de Haan p for N-DIM	0.4207	0.0470	1.0000
(32) Worst-case de Haan p for N-OLS	0.4290	0.0501	0.8846
(33) Worst-case de Haan p for N-AIPW	0.1135	0.0501	0.7528
(34) Worst-case de Haan p for S-DIM	0.3835	0.1239	1.0000
(35) Worst-case de Haan p for S-OLS	0.1564	0.1563	0.8638
(36) Worst-case de Haan p for S-AIPW	0.1894	0.1053	0.8923

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above violent crime conviction outcome variables.

Table S2.3. Treatment effects on violent crime conviction outcomes of the Female participants.

Statistic	Violent misdemeanor convictions by:			Violent felony convictions by:	
	age 30	age 40	age 30	age 30	age 40
(i) Number of observations	51	50	51	51	50
(ii) Mean of the control group	0.4231	0.4400	0.0385	0.0400	0.0400
(iii) Mean of the treatment group	0.0000	0.0000	0.0000	0.0000	0.0000
(iv) DIM (difference in means) estimate	-0.4231	-0.4400	-0.0385	-0.0400	-0.0400
(v) OLS estimate (with covariates)	-0.4900	-0.5228	-0.0407	-0.0436	-0.0436
(vi) AIPW (augmented IPW) estimate	-0.4367	-0.4713	-0.0350	-0.0380	-0.0380
(01) Asymptotic <i>p</i> -value for DIM	0.0672	0.0670	0.1602	0.1602	0.1602
(02) Asymptotic <i>p</i> -value for OLS	0.0673	0.0692	0.1662	0.1657	0.1657
(03) Asymptotic <i>p</i> -value for AIPW	0.0601	0.0543	0.1707	0.1630	0.1630
(04) Bootstrap <i>p</i> -value for DIM	0.0663	0.0649	0.1431	0.1434	0.1434
(05) Bootstrap <i>p</i> -value for OLS	0.0760	0.0762	0.1520	0.1529	0.1529
(06) Bootstrap <i>p</i> -value for AIPW	0.0770	0.0789	0.1452	0.1460	0.1460
(07) Permutation <i>p</i> -value for N-DIM	0.0508	0.0428	0.2028	0.1688	0.1688
(08) Permutation <i>p</i> -value for N-OLS	0.0228	0.0152	0.2176	0.1744	0.1744
(09) Permutation <i>p</i> -value for N-AIPW	0.0524	0.0416	0.3084	0.2616	0.2616
(10) Permutation <i>p</i> -value for S-DIM	0.0452	0.0408	0.2148	0.1992	0.1992
(11) Permutation <i>p</i> -value for S-OLS	0.0524	0.0672	0.2936	0.2456	0.2456
(12) Permutation <i>p</i> -value for S-AIPW	0.0704	0.0568	0.3724	0.3224	0.3224
(13) Worst-case maximum <i>p</i> for N-DIM	0.0691	0.0579	0.3061	0.2822	0.2822
(14) Worst-case maximum <i>p</i> for N-OLS	0.0536	0.0478	0.3202	0.2682	0.2682
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0927	0.0807	0.4021	0.3626	0.3626
(16) Worst-case maximum <i>p</i> for S-DIM	0.0609	0.0573	0.3173	0.3090	0.3090
(17) Worst-case maximum <i>p</i> for S-OLS	0.0815	0.0974	0.3991	0.3583	0.3583
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0893	0.0729	0.3811	0.3336	0.3336

Table S2.3. Continued

Statistic	Violent misdemeanor convictions by:		Violent felony convictions by: age 30
	age 30	age 40	
(19) Worst-case de Haan p for N-DIM	1.0000	0.1666	0.4288
(20) Worst-case de Haan p for N-OLS	0.9055	0.0815	0.4401
(21) Worst-case de Haan p for N-AIPW	0.5249	0.6554	0.5290
(22) Worst-case de Haan p for S-DIM	0.1856	0.2120	0.4725
(23) Worst-case de Haan p for S-OLS	1.0000	0.2870	0.6588
(24) Worst-case de Haan p for S-AIPW	0.3036	0.1446	0.7411
(25) Worst-case maximum p for N-DIM	0.2315	0.2315	0.5644
(26) Worst-case maximum p for N-OLS	0.1912	0.1912	0.5363
(27) Worst-case maximum p for N-AIPW	0.3228	0.3228	0.7251
(28) Worst-case maximum p for S-DIM	0.2291	0.2291	0.6181
(29) Worst-case maximum p for S-OLS	0.3258	0.3258	0.7166
(30) Worst-case maximum p for S-AIPW	0.2916	0.2916	0.6713
(31) Worst-case de Haan p for N-DIM	1.0000	0.6665	1.0000
(32) Worst-case de Haan p for N-OLS	1.0000	0.3260	1.0000
(33) Worst-case de Haan p for N-AIPW	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	0.7424	0.7424	0.8693
(35) Worst-case de Haan p for S-OLS	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	0.9109	0.5785	0.9109

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above violent crime conviction outcome variables.

APPENDIX S3. INFERENCE ON STANFORD–BINET (SB) IQ SCORES

Table S3.1. 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
(i) Number of observations	123	93	120	119	117	117	114
(ii) Mean of the control group	83.3	83.5	86.3	87.1	86.9	86.8	84.6
(iii) Mean of the treatment group	95.5	94.9	91.2	91.7	88.1	87.7	85.0
(iv) DIM (difference in means) estimate	12.2	11.4	4.9	4.7	1.3	0.9	0.4
(v) OLS estimate (with covariates)	10.8	11.0	4.2	3.9	1.1	-0.1	0.1
(vi) AIPW (augmented IPW) estimate	10.8	11.2	4.0	4.0	-0.3	-0.4	-0.1
(01) Asymptotic <i>p</i> -value for DIM	0.0000	0.0000	0.0116	0.0189	0.3024	0.3404	0.4327
(02) Asymptotic <i>p</i> -value for OLS	0.0000	0.0000	0.0120	0.0367	0.3132	0.4769	0.4780
(03) Asymptotic <i>p</i> -value for AIPW	0.0000	0.0000	0.0097	0.0254	0.4433	0.4111	0.4846
(04) Bootstrap <i>p</i> -value for DIM	0.0000	0.0000	0.0079	0.0091	0.2815	0.3339	0.4294
(05) Bootstrap <i>p</i> -value for OLS	0.0000	0.0000	0.0070	0.0173	0.2928	0.4755	0.4776
(06) Bootstrap <i>p</i> -value for AIPW	0.0000	0.0000	0.0062	0.0157	0.4417	0.4144	0.4857
(07) Permutation <i>p</i> -value for N-DIM	0.0004	0.0008	0.0104	0.0156	0.3408	0.3728	0.4360
(08) Permutation <i>p</i> -value for N-OLS	0.0004	0.0004	0.0140	0.0292	0.3752	0.4248	0.4932
(09) Permutation <i>p</i> -value for N-AIPW	0.0004	0.0004	0.0200	0.0352	0.3912	0.3704	0.4724
(10) Permutation <i>p</i> -value for S-DIM	0.0004	0.0008	0.0116	0.0164	0.3452	0.3716	0.4372
(11) Permutation <i>p</i> -value for S-OLS	0.0004	0.0004	0.0136	0.0392	0.3828	0.4264	0.4936
(12) Permutation <i>p</i> -value for S-AIPW	0.0004	0.0004	0.0152	0.0416	0.3936	0.3680	0.4712
(13) Worst-case maximum <i>p</i> for N-DIM	0.0025	0.0025	0.0164	0.0204	0.2737	0.3612	0.4863
(14) Worst-case maximum <i>p</i> for N-OLS	0.0025	0.0025	0.0289	0.0441	0.3478	0.5985	0.6568
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0025	0.0054	0.0289	0.0446	0.6691	0.5604	0.5034
(16) Worst-case maximum <i>p</i> for S-DIM	0.0025	0.0025	0.0178	0.0247	0.2819	0.3612	0.4863
(17) Worst-case maximum <i>p</i> for S-OLS	0.0025	0.0025	0.0285	0.0493	0.3539	0.5985	0.6568
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0025	0.0054	0.0268	0.0490	0.6691	0.5595	0.5021

Table S3.1. Continued

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
(19) Worst-case de Haan p for N-DIM	0.0026	0.0026	0.0828	0.0403	0.4045	0.5331	0.6096
(20) Worst-case de Haan p for N-OLS	0.0026	0.0026	0.0606	0.1694	0.7476	1.0000	1.0000
(21) Worst-case de Haan p for N-AIPW	0.0026	0.0059	0.0787	0.1228	0.8346	0.9882	0.6967
(22) Worst-case de Haan p for S-DIM	0.0026	0.0026	0.0362	0.1582	0.4535	0.5003	0.6016
(23) Worst-case de Haan p for S-OLS	0.0026	0.0026	0.0905	0.1303	0.4347	1.0000	1.0000
(24) Worst-case de Haan p for S-AIPW	0.0026	0.0059	0.1713	0.1295	0.9418	1.0000	0.5616
(25) Worst-case maximum p for N-DIM	0.0174	0.0174	0.0821	0.0821	0.8211	0.8211	0.8211
(26) Worst-case maximum p for N-OLS	0.0174	0.0174	0.1445	0.1765	1.0000	1.0000	1.0000
(27) Worst-case maximum p for N-AIPW	0.0174	0.0326	0.1445	0.1786	1.0000	1.0000	1.0000
(28) Worst-case maximum p for S-DIM	0.0174	0.0174	0.0889	0.0989	0.8458	0.8458	0.8458
(29) Worst-case maximum p for S-OLS	0.0174	0.0174	0.1427	0.1970	1.0000	1.0000	1.0000
(30) Worst-case maximum p for S-AIPW	0.0174	0.0326	0.1340	0.1960	1.0000	1.0000	1.0000
(31) Worst-case de Haan p for N-DIM	0.0181	0.0181	0.3312	0.2015	1.0000	1.0000	1.0000
(32) Worst-case de Haan p for N-OLS	0.0181	0.0181	0.3032	0.6776	1.0000	1.0000	1.0000
(33) Worst-case de Haan p for N-AIPW	0.0181	0.0356	0.2934	0.4912	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	0.0181	0.0181	0.1811	0.6327	1.0000	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	0.0181	0.0181	0.4524	0.5210	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	0.0181	0.0356	0.6851	0.6475	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all seven of the above Stanford–Binet IQ score variables.

Table S3.2. 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
(i) Number of observations	72	54	72	71	67	71	71
(ii) Mean of the control group	83.1	84.8	85.8	87.7	89.1	89.0	86.0
(iii) Mean of the treatment group	94.9	95.4	91.5	91.1	88.3	88.4	83.7
(iv) DIM (difference in means) estimate	11.8	10.6	5.7	3.4	-0.7	-0.6	-2.3
(v) OLS estimate (with covariates)	9.1	9.1	3.5	1.7	-2.2	-3.2	-4.0
(vi) AIPW (augmented IPW) estimate	9.0	9.2	3.1	1.6	-3.8	-4.2	-4.7
(01) Asymptotic <i>p</i> -value for DIM	0.0000	0.0001	0.0189	0.0854	0.4034	0.4088	0.1784
(02) Asymptotic <i>p</i> -value for OLS	0.0000	0.0003	0.0478	0.2015	0.1932	0.1114	0.0593
(03) Asymptotic <i>p</i> -value for AIPW	0.0000	0.0000	0.0557	0.2040	0.0512	0.0398	0.0225
(04) Bootstrap <i>p</i> -value for DIM	0.0000	0.0001	0.0160	0.0808	0.4020	0.4114	0.1814
(05) Bootstrap <i>p</i> -value for OLS	0.0000	0.0001	0.0338	0.1951	0.1959	0.1134	0.0671
(06) Bootstrap <i>p</i> -value for AIPW	0.0000	0.0002	0.0512	0.2143	0.0719	0.0577	0.0412
(07) Permutation <i>p</i> -value for N-DIM	0.0004	0.0008	0.0212	0.0812	0.3768	0.3932	0.1880
(08) Permutation <i>p</i> -value for N-OLS	0.0004	0.0004	0.0680	0.2140	0.1696	0.1052	0.0600
(09) Permutation <i>p</i> -value for N-AIPW	0.0004	0.0020	0.1004	0.2388	0.0652	0.0504	0.0272
(10) Permutation <i>p</i> -value for S-DIM	0.0004	0.0008	0.0208	0.0784	0.3772	0.3912	0.1860
(11) Permutation <i>p</i> -value for S-OLS	0.0004	0.0012	0.0560	0.2008	0.1560	0.1108	0.0640
(12) Permutation <i>p</i> -value for S-AIPW	0.0004	0.0004	0.0712	0.2104	0.0556	0.0472	0.0292
(13) Worst-case maximum <i>p</i> for N-DIM	0.0026	0.0053	0.0289	0.0858	0.5501	0.5649	0.2529
(14) Worst-case maximum <i>p</i> for N-OLS	0.0025	0.0053	0.0704	0.2291	0.3226	0.1943	0.1071
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0052	0.0092	0.1006	0.2111	0.1697	0.1167	0.0648
(16) Worst-case maximum <i>p</i> for S-DIM	0.0035	0.0053	0.0289	0.0858	0.5501	0.5635	0.2529
(17) Worst-case maximum <i>p</i> for S-OLS	0.0025	0.0074	0.0700	0.2099	0.3066	0.2054	0.1147
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0052	0.0050	0.0838	0.1905	0.1519	0.1147	0.0678

Table S3.2. Continued

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
(19) Worst-case de Haan p for N-DIM	0.0032	0.1314	0.0940	0.1793	0.7093	0.9011	0.4790
(20) Worst-case de Haan p for N-OLS	0.0026	0.0216	0.2025	0.7967	0.9460	0.3635	0.1739
(21) Worst-case de Haan p for N-AIPW	0.0701	0.1048	0.4004	0.5044	0.2243	0.2760	0.0949
(22) Worst-case de Haan p for S-DIM	0.0051	0.1314	0.0975	0.2259	0.7234	0.9429	0.3615
(23) Worst-case de Haan p for S-OLS	0.0026	0.0077	0.1732	0.4357	0.4821	0.2462	0.2182
(24) Worst-case de Haan p for S-AIPW	0.0060	0.0059	0.2362	0.3701	0.1894	0.3800	0.1012
(25) Worst-case maximum p for N-DIM	0.0179	0.0319	0.1447	0.3433	1.0000	1.0000	0.7588
(26) Worst-case maximum p for N-OLS	0.0174	0.0317	0.3520	0.5828	0.5828	0.5828	0.4285
(27) Worst-case maximum p for N-AIPW	0.0366	0.0551	0.4024	0.4024	0.4024	0.4024	0.3242
(28) Worst-case maximum p for S-DIM	0.0246	0.0319	0.1447	0.3433	1.0000	1.0000	0.7588
(29) Worst-case maximum p for S-OLS	0.0174	0.0447	0.2501	0.6163	0.6163	0.6163	0.4589
(30) Worst-case maximum p for S-AIPW	0.0348	0.0348	0.3391	0.3440	0.3440	0.3440	0.3391
(31) Worst-case de Haan p for N-DIM	0.0222	0.6571	0.5643	0.7170	1.0000	1.0000	1.0000
(32) Worst-case de Haan p for N-OLS	0.0181	0.1296	0.8696	1.0000	1.0000	1.0000	0.8696
(33) Worst-case de Haan p for N-AIPW	0.4904	0.5692	0.8972	0.8972	0.8972	0.8972	0.5692
(34) Worst-case de Haan p for S-DIM	0.0358	0.6571	0.5848	0.9034	1.0000	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	0.0181	0.0462	0.8662	0.8727	0.8727	0.8727	0.8727
(36) Worst-case de Haan p for S-AIPW	0.0415	0.0415	0.7574	0.7574	0.7574	0.7574	0.5062

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all seven of the above Stanford–Binet IQ score variables.

Table S3.3. 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
(i) Number of observations	51	39	48	48	50	46	43
(ii) Mean of the control group	83.7	81.7	87.2	86.0	83.6	83.0	81.8
(iii) Mean of the treatment group	96.4	94.3	90.9	92.5	87.8	86.7	86.8
(iv) DIM (difference in means) estimate	12.7	12.7	3.8	6.5	4.2	3.7	5.0
(v) OLS estimate (with covariates)	13.3	13.5	5.1	7.1	4.9	4.7	6.5
(vi) AIPW (augmented IPW) estimate	13.4	14.2	5.3	7.3	4.7	4.8	6.5
(01) Asymptotic <i>p</i> -value for DIM	0.00000	0.00023	0.1221	0.04444	0.1247	0.1466	0.0918
(02) Asymptotic <i>p</i> -value for OLS	0.00000	0.00119	0.05114	0.0427	0.0962	0.0670	0.0317
(03) Asymptotic <i>p</i> -value for AIPW	0.00000	0.00088	0.0365	0.0313	0.1144	0.0633	0.0277
(04) Bootstrap <i>p</i> -value for DIM	0.00000	0.00013	0.1139	0.0181	0.0937	0.1281	0.0703
(05) Bootstrap <i>p</i> -value for OLS	0.00000	0.00007	0.0398	0.0173	0.0637	0.0528	0.0257
(06) Bootstrap <i>p</i> -value for AIPW	0.00000	0.00066	0.0281	0.0154	0.0896	0.0679	0.0323
(07) Permutation <i>p</i> -value for N-DIM	0.0004	0.0044	0.1188	0.0444	0.1360	0.1560	0.0940
(08) Permutation <i>p</i> -value for N-OLS	0.0004	0.0048	0.0676	0.0344	0.1108	0.1100	0.0404
(09) Permutation <i>p</i> -value for N-AIPW	0.0004	0.0040	0.0604	0.0364	0.1276	0.1112	0.0508
(10) Permutation <i>p</i> -value for S-DIM	0.0004	0.0052	0.1232	0.0468	0.1416	0.1608	0.0936
(11) Permutation <i>p</i> -value for S-OLS	0.0004	0.0036	0.0632	0.0524	0.1244	0.0928	0.0396
(12) Permutation <i>p</i> -value for S-AIPW	0.0004	0.0064	0.0636	0.0564	0.1704	0.1128	0.0596
(13) Worst-case maximum <i>p</i> for N-DIM	0.0025	0.0175	0.1215	0.0689	0.1513	0.1982	0.1357
(14) Worst-case maximum <i>p</i> for N-OLS	0.0025	0.0169	0.0870	0.0590	0.1410	0.1593	0.1322
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0025	0.0234	0.0807	0.0661	0.1784	0.1548	0.1808
(16) Worst-case maximum <i>p</i> for S-DIM	0.0025	0.0208	0.1285	0.0796	0.1514	0.2102	0.1301
(17) Worst-case maximum <i>p</i> for S-OLS	0.0053	0.0212	0.0808	0.0831	0.1632	0.1354	0.1215
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0053	0.0238	0.0799	0.0952	0.2080	0.1578	0.1840

Table S3.3. Continued

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
(19) Worst-case de Haan p for N-DIM	0.0026	0.1063	0.1996	0.1198	0.2354	0.2707	0.2583
(20) Worst-case de Haan p for N-OLS	0.0026	0.8784	0.2124	0.2118	0.2828	0.5234	0.1652
(21) Worst-case de Haan p for N-AIPW	0.0026	0.2050	0.2584	0.2121	0.3290	0.3581	1.0000
(22) Worst-case de Haan p for S-DIM	0.0026	0.0635	0.3200	0.3847	0.6181	0.3197	0.7997
(23) Worst-case de Haan p for S-OLS	0.0055	0.0854	0.1752	0.2639	0.7208	0.2805	0.2821
(24) Worst-case de Haan p for S-AIPW	0.0055	0.0503	0.4790	0.1950	0.2832	0.2873	0.4267
(25) Worst-case maximum p for N-DIM	0.0174	0.1048	0.4858	0.3445	0.4858	0.4858	0.4858
(26) Worst-case maximum p for N-OLS	0.0174	0.1017	0.3480	0.2951	0.3967	0.3967	0.3967
(27) Worst-case maximum p for N-AIPW	0.0174	0.1404	0.3305	0.3305	0.4643	0.4643	0.4643
(28) Worst-case maximum p for S-DIM	0.0174	0.1246	0.5141	0.3982	0.5141	0.5141	0.5141
(29) Worst-case maximum p for S-OLS	0.0369	0.1274	0.4041	0.4041	0.4041	0.4041	0.4041
(30) Worst-case maximum p for S-AIPW	0.0369	0.1550	0.3997	0.3997	0.4734	0.4734	0.4734
(31) Worst-case de Haan p for N-DIM	0.0181	0.6378	0.7983	0.6378	0.7983	0.7983	0.7983
(32) Worst-case de Haan p for N-OLS	0.0181	1.0000	1.0000	1.0000	1.0000	1.0000	0.9909
(33) Worst-case de Haan p for N-AIPW	0.0181	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	0.0181	0.3810	1.0000	1.0000	1.0000	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	0.0384	0.5124	0.8762	1.0000	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	0.0387	0.3020	1.0000	0.9749	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all seven of the above Stanford–Binet IQ score variables.

APPENDIX S4. INFERENCE ON ACHIEVEMENT TEST OUTCOMES

Table S4.1. 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
(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
(iv) DIM (difference in means)	6.2990	6.4920	9.5892	10.141	12.232
(v) OLS estimate (with 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
(01) Asymptotic <i>p</i> -value for DIM	0.0070	0.0197	0.0004	0.0003	0.0023
(02) Asymptotic <i>p</i> -value for OLS	0.0403	0.0867	0.0019	0.0014	0.0130
(03) Asymptotic <i>p</i> -value for AIPW	0.0448	0.0590	0.0013	0.0009	0.0095
(04) Bootstrap <i>p</i> -value for DIM	0.0043	0.0145	0.0001	0.0001	0.0009
(05) Bootstrap <i>p</i> -value for OLS	0.0339	0.0801	0.0013	0.0008	0.0082
(06) Bootstrap <i>p</i> -value for AIPW	0.0636	0.0772	0.0022	0.0011	0.0170
(07) Permutation <i>p</i> -value for N-DIM	0.0104	0.0240	0.0016	0.0012	0.0036
(08) Permutation <i>p</i> -value for N-OLS	0.0404	0.1060	0.0036	0.0032	0.0092
(09) Permutation <i>p</i> -value for N-AIPW	0.0712	0.0952	0.0056	0.0036	0.0176
(10) Permutation <i>p</i> -value for S-DIM	0.0100	0.0212	0.0012	0.0008	0.0032
(11) Permutation <i>p</i> -value for S-OLS	0.0430	0.0988	0.0024	0.0020	0.0108
(12) Permutation <i>p</i> -value for S-AIPW	0.0700	0.0920	0.0028	0.0028	0.0140
(13) Worst-case max. <i>p</i> for N-DIM	0.0304	0.0541	0.0089	0.0111	0.0106
(14) Worst-case max. <i>p</i> for N-OLS	0.0734	0.1409	0.0126	0.0123	0.0242
(15) Worst-case max. <i>p</i> for N-AIPW	0.0951	0.1121	0.0195	0.0165	0.0318
(16) Worst-case max. <i>p</i> for S-DIM	0.0320	0.0436	0.0083	0.0056	0.0102
(17) Worst-case max. <i>p</i> for S-OLS	0.0796	0.1351	0.0156	0.0132	0.0255
(18) Worst-case max. <i>p</i> for S-AIPW	0.0872	0.1116	0.0156	0.0117	0.0251

Table S4.1. Continued

(19) Worst-case de Haan p for N-DIM	0.0639	0.1386	0.0574	0.2847	0.0878
(20) Worst-case de Haan p for N-OLS	0.2056	0.2557	0.0434	0.0360	0.1602
(21) Worst-case de Haan p ; N-AIPW	0.3433	0.4750	0.0819	0.0472	0.1659
(22) Worst-case de Haan p for S-DIM	0.0578	0.0984	0.0574	0.2847	0.0186
(23) Worst-case de Haan p for S-OLS	0.1514	0.2009	1.0000	0.0346	0.2330
(24) Worst-case de Haan p for S-AIPW	0.2201	0.2904	0.0574	0.0227	0.2808
(25) Worst-case max. p for N-DIM	0.0607	0.0607	0.0447	0.0447	0.0447
(26) Worst-case max. p for N-OLS	0.1468	0.1468	0.0616	0.0616	0.0725
(27) Worst-case max. p for N-AIPW	0.1901	0.1901	0.0824	0.0824	0.0954
(28) Worst-case max. p for S-DIM	0.0640	0.0640	0.0331	0.0278	0.0331
(29) Worst-case max. p for S-OLS	0.1593	0.1593	0.0660	0.0660	0.0766
(30) Worst-case max. p for S-AIPW	0.1743	0.1743	0.0625	0.0587	0.0752
(31) Worst-case de Haan p for N-DIM	0.2868	0.2868	0.2868	0.2868	0.2868
(32) Worst-case de Haan p for N-OLS	0.4807	0.4807	0.1798	0.1798	0.4807
(33) Worst-case de Haan p ; N-AIPW	0.6866	0.6866	0.3277	0.2358	0.4976
(34) Worst-case de Haan p for S-DIM	0.2294	0.2294	0.2294	0.2847	0.0930
(35) Worst-case de Haan p for S-OLS	0.6058	0.6058	1.0000	0.1732	0.6058
(36) Worst-case de Haan p for S-AIPW	0.6602	0.6602	0.2295	0.1137	0.6602

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all five of the above California Achievement Test component score variables.

Table S4.2. 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
(i) Number of observations	55	55	55	55	55
(ii) Mean of the control group	9,000	8,1071	6,5357	6,9643	11,536
(iii) Mean of the treatment group	13,926	16,000	14,333	15,556	18,519
(iv) DIM (difference in means)	4,9259	7,8929	7,7976	8,5913	6,9828
(v) OLS estimate (with 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
(01) Asymptotic <i>p</i> -value for DIM	0.0748	0.0589	0.0236	0.0193	0.0952
(02) Asymptotic <i>p</i> -value for OLS	0.2638	0.2814	0.0947	0.0707	0.2478
(03) Asymptotic <i>p</i> -value for AIPW	0.2957	0.2410	0.0815	0.0538	0.2652
(04) Bootstrap <i>p</i> -value for DIM	0.0712	0.0453	0.0168	0.0133	0.0840
(05) Bootstrap <i>p</i> -value for OLS	0.2656	0.2760	0.0899	0.0658	0.2372
(06) Bootstrap <i>p</i> -value for AIPW	0.3221	0.2629	0.0995	0.0638	0.2865
(07) Permutation <i>p</i> -value for N-DIM	0.0768	0.0664	0.0276	0.0176	0.0944
(08) Permutation <i>p</i> -value for N-OLS	0.2456	0.2836	0.1100	0.0880	0.2200
(09) Permutation <i>p</i> -value for N-AIPW	0.3012	0.2600	0.1184	0.0896	0.2564
(10) Permutation <i>p</i> -value for S-DIM	0.0752	0.0664	0.0260	0.0156	0.0928
(11) Permutation <i>p</i> -value for S-OLS	0.2716	0.2932	0.1180	0.0836	0.2492
(12) Permutation <i>p</i> -value for S-AIPW	0.3112	0.2608	0.1076	0.0712	0.2600
(13) Worst-case max. <i>p</i> for N-DIM	0.1449	0.1242	0.0729	0.0554	0.1144
(14) Worst-case max. <i>p</i> for N-OLS	0.2817	0.3155	0.1617	0.1346	0.2532
(15) Worst-case max. <i>p</i> for N-AIPW	0.3200	0.2671	0.1763	0.1520	0.2679
(16) Worst-case max. <i>p</i> for S-DIM	0.1444	0.1185	0.0686	0.0464	0.1137
(17) Worst-case max. <i>p</i> for S-OLS	0.3033	0.3230	0.1552	0.1244	0.2685
(18) Worst-case max. <i>p</i> for S-AIPW	0.3253	0.2722	0.1764	0.1234	0.2741

Table S4.2. Continued

Statistic	CAT reading age 14	CAT arithmetic age 14	CAT language age 14	CAT mechanics age 14	CAT spelling age 14
(19) Worst-case de Haan p for N-DIM	0.3541	0.2534	0.1158	0.1163	0.1980
(20) Worst-case de Haan p for N-OLS	0.8308	0.4729	0.2333	0.2364	0.5789
(21) Worst-case de Haan p : N-AIPW	0.5141	0.5172	0.4745	0.2668	0.4986
(22) Worst-case de Haan p for S-DIM	0.1975	0.3746	0.1592	0.1307	0.2247
(23) Worst-case de Haan p for S-OLS	0.4292	0.5221	0.4189	0.2937	0.5058
(24) Worst-case de Haan p for S-AIPW	0.4273	0.3434	0.3482	0.2364	0.4587
(25) Worst-case max. p for N-DIM	0.3432	0.3432	0.2917	0.2771	0.3432
(26) Worst-case max. p for N-OLS	0.7597	0.7597	0.6728	0.6728	0.7597
(27) Worst-case max. p for N-AIPW	0.8012	0.8012	0.7598	0.7598	0.8012
(28) Worst-case max. p for S-DIM	0.3410	0.3410	0.2743	0.2320	0.3410
(29) Worst-case max. p for S-OLS	0.8056	0.8056	0.6219	0.6219	0.8056
(30) Worst-case max. p for S-AIPW	0.8167	0.8167	0.7058	0.6171	0.8167
(31) Worst-case de Haan p for N-DIM	0.5941	0.5941	0.5791	0.5791	0.5941
(32) Worst-case de Haan p for N-OLS	1.0000	1.0000	1.0000	1.0000	1.0000
(33) Worst-case de Haan p : N-AIPW	1.0000	1.0000	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	0.6537	0.6537	0.6537	0.6537	0.6537
(35) Worst-case de Haan p for S-OLS	1.0000	1.0000	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all five of the above California Achievement Test component score variables.

Table S4.3. 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
(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
(iv) DIM (difference in means)	8.0556	4.9293	11.621	11.803	18.778
(v) OLS estimate (with 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
(01) Asymptotic <i>p</i> -value for DIM	0.0139	0.0621	0.0023	0.0030	0.0017
(02) Asymptotic <i>p</i> -value for OLS	0.0110	0.0067	0.0095	0.0007	0.0019
(03) Asymptotic <i>p</i> -value for AIPW	0.0130	0.0102	0.0099	0.0014	0.0017
(04) Bootstrap <i>p</i> -value for DIM	0.0087	0.0499	0.0011	0.0013	0.0007
(05) Bootstrap <i>p</i> -value for OLS	0.0081	0.0069	0.0004	0.0005	0.0018
(06) Bootstrap <i>p</i> -value for AIPW	0.0128	0.0138	0.0013	0.0015	0.0042
(07) Permutation <i>p</i> -value for N-DIM	0.0160	0.0656	0.0048	0.0052	0.0040
(08) Permutation <i>p</i> -value for N-OLS	0.0276	0.0284	0.0020	0.0036	0.0048
(09) Permutation <i>p</i> -value for N-AIPW	0.0356	0.0540	0.0064	0.0056	0.0068
(10) Permutation <i>p</i> -value for S-DIM	0.0132	0.0596	0.0024	0.0032	0.0036
(11) Permutation <i>p</i> -value for S-OLS	0.0156	0.0084	0.0012	0.0016	0.0028
(12) Permutation <i>p</i> -value for S-AIPW	0.0268	0.0284	0.0044	0.0072	0.0064
(13) Worst-case max. <i>p</i> for N-DIM	0.0348	0.1106	0.0134	0.0154	0.0144
(14) Worst-case max. <i>p</i> for N-OLS	0.0573	0.0719	0.0130	0.0151	0.0189
(15) Worst-case max. <i>p</i> for N-AIPW	0.0704	0.1202	0.0179	0.0179	0.0273
(16) Worst-case max. <i>p</i> for S-DIM	0.0358	0.1046	0.0113	0.0137	0.0115
(17) Worst-case max. <i>p</i> for S-OLS	0.0440	0.0266	0.0100	0.0112	0.0149
(18) Worst-case max. <i>p</i> for S-AIPW	0.0561	0.0624	0.0168	0.0211	0.0176

Table S4.3. Continued

Statistic	CAT reading age 14	CAT arithmetic age 14	CAT language age 14	CAT mechanics age 14	CAT spelling age 14
(19) Worst-case de Haan p for N-DIM	0.1001	0.2008	0.0328	0.0264	0.0661
(20) Worst-case de Haan p for N-OLS	0.1213	0.1409	0.0551	0.0587	0.0571
(21) Worst-case de Haan p : N-AIPW	0.1631	0.2323	0.2163	0.0606	0.0894
(22) Worst-case de Haan p for S-DIM	0.1081	0.1704	0.0328	0.1974	0.0434
(23) Worst-case de Haan p for S-OLS	0.1330	0.0483	0.0252	0.0250	0.0271
(24) Worst-case de Haan p for S-AIPW	0.1125	0.0731	0.0524	0.0606	0.0234
(25) Worst-case max. p for N-DIM	0.0696	0.1106	0.0671	0.0671	0.0671
(26) Worst-case max. p for N-OLS	0.1147	0.1147	0.0652	0.0652	0.0652
(27) Worst-case max. p for N-AIPW	0.1408	0.1408	0.0895	0.0895	0.0895
(28) Worst-case max. p for S-DIM	0.0715	0.1046	0.0566	0.0566	0.0566
(29) Worst-case max. p for S-OLS	0.0533	0.0533	0.0500	0.0500	0.0500
(30) Worst-case max. p for S-AIPW	0.1122	0.1122	0.0842	0.0842	0.0842
(31) Worst-case de Haan p for N-DIM	0.2001	0.2008	0.1321	0.1321	0.1984
(32) Worst-case de Haan p for N-OLS	0.2757	0.2757	0.2757	0.2757	0.2757
(33) Worst-case de Haan p : N-AIPW	0.4894	0.4894	0.4894	0.3030	0.3576
(34) Worst-case de Haan p for S-DIM	0.3244	0.3407	0.1640	0.3407	0.1736
(35) Worst-case de Haan p for S-OLS	0.1330	0.1250	0.1250	0.1250	0.1250
(36) Worst-case de Haan p for S-AIPW	0.2097	0.2097	0.2097	0.2097	0.1272

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all five of the above California Achievement Test component score variables.

APPENDIX S5. INFERENCE ON EDUCATION OUTCOMES

Table S5.1. Treatment effects on education outcomes of the **Pooled** participants.

Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19
(i) Number of observations	123	123	121	77
(ii) Mean of the control group	0.4000	0.2308	11.079	1.7051
(iii) Mean of the treatment group	0.6379	0.3276	11.534	2.0826
(iv) DIM (difference in means) estimate	0.2379	0.0968	0.4551	0.3775
(v) OLS estimate (with covariates)	0.2425	0.1215	0.5667	0.3339
(vi) AIPW (augmented IPW) estimate	0.2449	0.1173	0.5497	0.3763
(01) Asymptotic <i>p</i> -value for DIM	0.0066	0.1073	0.0456	0.0122
(02) Asymptotic <i>p</i> -value for OLS	0.0069	0.0712	0.0289	0.0419
(03) Asymptotic <i>p</i> -value for AIPW	0.0051	0.0723	0.0226	0.0094
(04) Bootstrap <i>p</i> -value for DIM	0.0012	0.1113	0.0246	0.0034
(05) Bootstrap <i>p</i> -value for OLS	0.0015	0.0716	0.0138	0.0163
(06) Bootstrap <i>p</i> -value for AIPW	0.0028	0.0887	0.0264	0.0438
(07) Permutation <i>p</i> -value for N-DIM	0.0072	0.1228	0.0576	0.0120
(08) Permutation <i>p</i> -value for N-OLS	0.0064	0.0716	0.0220	0.0252
(09) Permutation <i>p</i> -value for N-AIPW	0.0052	0.0740	0.0264	0.0192
(10) Permutation <i>p</i> -value for S-DIM	0.0088	0.1192	0.0620	0.0124
(11) Permutation <i>p</i> -value for S-OLS	0.0072	0.0848	0.0404	0.0424
(12) Permutation <i>p</i> -value for S-AIPW	0.0092	0.0888	0.0432	0.0216
(13) Worst-case maximum <i>p</i> for N-DIM	0.0212	0.1821	0.0521	0.0358
(14) Worst-case maximum <i>p</i> for N-OLS	0.0220	0.1283	0.0361	0.0679
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0258	0.1551	0.0458	0.0679
(16) Worst-case maximum <i>p</i> for S-DIM	0.0251	0.1733	0.0603	0.0387
(17) Worst-case maximum <i>p</i> for S-OLS	0.0299	0.1406	0.0629	0.0897
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0356	0.1685	0.0653	0.0734

Table S5.1. Continued

Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19
(19) Worst-case de Haan p for N-DIM	0.0433	0.3634	0.0972	0.1215
(20) Worst-case de Haan p for N-OLS	0.2024	0.2779	1.0000	0.1321
(21) Worst-case de Haan p for N-AIPW	0.0716	0.1955	0.3247	0.1244
(22) Worst-case de Haan p for S-DIM	0.0520	0.3104	0.1630	0.0906
(23) Worst-case de Haan p for S-OLS	0.1802	0.3265	0.1009	0.5868
(24) Worst-case de Haan p for S-AIPW	0.0893	0.2323	0.1372	0.1608
(25) Worst-case maximum p for N-DIM	0.0849	0.1821	0.1075	0.1075
(26) Worst-case maximum p for N-OLS	0.0880	0.1358	0.1083	0.1358
(27) Worst-case maximum p for N-AIPW	0.1031	0.1551	0.1375	0.1375
(28) Worst-case maximum p for S-DIM	0.1006	0.1733	0.1207	0.1162
(29) Worst-case maximum p for S-OLS	0.1198	0.1886	0.1886	0.1886
(30) Worst-case maximum p for S-AIPW	0.1424	0.1959	0.1959	0.1959
(31) Worst-case de Haan p for N-DIM	0.1734	0.3634	0.2915	0.2915
(32) Worst-case de Haan p for N-OLS	0.6071	0.6071	1.0000	0.5282
(33) Worst-case de Haan p for N-AIPW	0.2866	0.3911	0.3911	0.3733
(34) Worst-case de Haan p for S-DIM	0.2080	0.3260	0.3260	0.2719
(35) Worst-case de Haan p for S-OLS	0.5405	0.6531	0.4035	0.6531
(36) Worst-case de Haan p for S-AIPW	0.3570	0.4116	0.4116	0.4116

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above education variables.

Table S5.2. Treatment effects on education outcomes of the Male participants.

Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19
(i) Number of observations	72	72	72	47
(ii) Mean of the control group	0.5128	0.3333	11.282	1.7942
(iii) Mean of the treatment group	0.4848	0.3939	11.364	1.8138
(iv) DIM (difference in means) estimate	-0.0280	0.0606	0.0816	0.0196
(v) OLS estimate (with covariates)	0.0210	0.0716	0.1442	-0.0252
(vi) AIPW (augmented IPW) estimate	0.0148	0.0706	0.0874	-0.0354
(01) Asymptotic <i>p</i> -value for DIM	0.4112	0.2932	0.3996	0.4637
(02) Asymptotic <i>p</i> -value for OLS	0.4386	0.2843	0.3367	0.4593
(03) Asymptotic <i>p</i> -value for AIPW	0.4550	0.2762	0.3902	0.4366
(04) Bootstrap <i>p</i> -value for DIM	0.4013	0.2925	0.3948	0.4580
(05) Bootstrap <i>p</i> -value for OLS	0.4278	0.2797	0.3246	0.4512
(06) Bootstrap <i>p</i> -value for AIPW	0.4540	0.2886	0.3901	0.4336
(07) Permutation <i>p</i> -value for N-DIM	0.3936	0.3080	0.4244	0.4832
(08) Permutation <i>p</i> -value for N-OLS	0.4672	0.2760	0.3568	0.4352
(09) Permutation <i>p</i> -value for N-AIPW	0.4824	0.2776	0.4196	0.4288
(10) Permutation <i>p</i> -value for S-DIM	0.3760	0.2976	0.4216	0.4828
(11) Permutation <i>p</i> -value for S-OLS	0.4700	0.2960	0.3652	0.4396
(12) Permutation <i>p</i> -value for S-AIPW	0.4868	0.2932	0.4240	0.4328
(13) Worst-case maximum <i>p</i> for N-DIM	0.6567	0.3688	0.3587	0.5132
(14) Worst-case maximum <i>p</i> for N-OLS	0.5365	0.3357	0.3902	0.5529
(15) Worst-case maximum <i>p</i> for N-AIPW	0.5633	0.3414	0.4456	0.5225
(16) Worst-case maximum <i>p</i> for S-DIM	0.6373	0.3582	0.3526	0.5132
(17) Worst-case maximum <i>p</i> for S-OLS	0.5393	0.3535	0.3962	0.5535
(18) Worst-case maximum <i>p</i> for S-AIPW	0.5651	0.3619	0.4583	0.5267

Table S5.2. Continued

Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19
(19) Worst-case de Haan p for N-DIM	1.0000	0.4664	0.7137	1.0000
(20) Worst-case de Haan p for N-OLS	0.6911	0.4052	0.4444	1.0000
(21) Worst-case de Haan p for N-AIPW	0.8732	0.6451	0.7984	1.0000
(22) Worst-case de Haan p for S-DIM	0.8990	0.4612	0.5578	0.8153
(23) Worst-case de Haan p for S-OLS	0.6606	0.7614	0.6140	0.8626
(24) Worst-case de Haan p for S-AIPW	0.7190	0.4378	0.6282	0.6983
(25) Worst-case maximum p for N-DIM	1.0000	1.0000	1.0000	1.0000
(26) Worst-case maximum p for N-OLS	1.0000	1.0000	1.0000	1.0000
(27) Worst-case maximum p for N-AIPW	1.0000	1.0000	1.0000	1.0000
(28) Worst-case maximum p for S-DIM	1.0000	1.0000	1.0000	1.0000
(29) Worst-case maximum p for S-OLS	1.0000	1.0000	1.0000	1.0000
(30) Worst-case maximum p for S-AIPW	1.0000	1.0000	1.0000	1.0000
(31) Worst-case de Haan p for N-DIM	1.0000	1.0000	1.0000	1.0000
(32) Worst-case de Haan p for N-OLS	1.0000	1.0000	1.0000	1.0000
(33) Worst-case de Haan p for N-AIPW	1.0000	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	1.0000	1.0000	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	1.0000	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above education variables.

Table S5.3. Treatment effects on education outcomes of the Female participants.

Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19
(i) Number of observations	51	51	49	30
(ii) Mean of the control group	0.2308	0.0769	10.750	1.5269
(iii) Mean of the treatment group	0.8400	0.2400	11.760	2.4147
(iv) DIM (difference in means) estimate	0.6092	0.1631	1.0100	0.8878
(v) OLS estimate (with covariates)	0.5541	0.1735	1.1648	0.9098
(vi) AIPW (augmented IPW) estimate	0.5699	0.1832	1.2024	0.9575
(01) Asymptotic <i>p</i> -value for DIM	0.0000	0.0450	0.0055	0.0000
(02) Asymptotic <i>p</i> -value for OLS	0.0000	0.0337	0.0085	0.0002
(03) Asymptotic <i>p</i> -value for AIPW	0.0000	0.0286	0.0023	0.0000
(04) Bootstrap <i>p</i> -value for DIM	0.0000	0.0529	0.0030	0.0000
(05) Bootstrap <i>p</i> -value for OLS	0.0000	0.0430	0.0076	0.0003
(06) Bootstrap <i>p</i> -value for AIPW	0.0000	0.0494	0.0106	0.0155
(07) Permutation <i>p</i> -value for N-DIM	0.0004	0.0624	0.0088	0.0004
(08) Permutation <i>p</i> -value for N-OLS	0.0004	0.0512	0.0036	0.0004
(09) Permutation <i>p</i> -value for N-AIPW	0.0004	0.0412	0.0020	0.0004
(10) Permutation <i>p</i> -value for S-DIM	0.0004	0.0572	0.0144	0.0004
(11) Permutation <i>p</i> -value for S-OLS	0.0004	0.0396	0.0164	0.0004
(12) Permutation <i>p</i> -value for S-AIPW	0.0004	0.0420	0.0120	0.0004
(13) Worst-case maximum <i>p</i> for N-DIM	0.0037	0.1341	0.0275	0.0078
(14) Worst-case maximum <i>p</i> for N-OLS	0.0037	0.1201	0.0172	0.0132
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0037	0.1085	0.0172	0.0101
(16) Worst-case maximum <i>p</i> for S-DIM	0.0037	0.1085	0.0297	0.0086
(17) Worst-case maximum <i>p</i> for S-OLS	0.0046	0.0951	0.0374	0.0135
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0054	0.1056	0.0345	0.0119

Table S5.3. Continued

Statistic	HS graduate age 19	Vocational training age 40	Highest grade age 19	Grade point avg. age 19
(19) Worst-case de Haan p for N-DIM	0.0236	0.9242	0.0698	0.0129
(20) Worst-case de Haan p for N-OLS	0.0236	0.3124	0.0852	0.0166
(21) Worst-case de Haan p for N-AIPW	0.0236	0.5274	0.1024	0.0247
(22) Worst-case de Haan p for S-DIM	0.0236	0.1872	0.0585	0.0151
(23) Worst-case de Haan p for S-OLS	0.0135	0.4660	0.0981	0.0401
(24) Worst-case de Haan p for S-AIPW	0.0058	0.2630	0.0935	0.0164
(25) Worst-case maximum p for N-DIM	0.0148	0.1341	0.0551	0.0234
(26) Worst-case maximum p for N-OLS	0.0146	0.1201	0.0397	0.0397
(27) Worst-case maximum p for N-AIPW	0.0148	0.1085	0.0345	0.0303
(28) Worst-case maximum p for S-DIM	0.0148	0.1085	0.0593	0.0259
(29) Worst-case maximum p for S-OLS	0.0184	0.0951	0.0748	0.0404
(30) Worst-case maximum p for S-AIPW	0.0218	0.1056	0.0690	0.0357
(31) Worst-case de Haan p for N-DIM	0.0709	0.9242	0.1395	0.0517
(32) Worst-case de Haan p for N-OLS	0.0709	0.3124	0.1704	0.0662
(33) Worst-case de Haan p for N-AIPW	0.0945	0.5274	0.2047	0.0945
(34) Worst-case de Haan p for S-DIM	0.0709	0.1872	0.1169	0.0603
(35) Worst-case de Haan p for S-OLS	0.0539	0.4660	0.1963	0.1204
(36) Worst-case de Haan p for S-AIPW	0.0232	0.2630	0.1871	0.0493

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above education variables.

APPENDIX S6. INFERENCE ON CRIME OUTCOMES

Table S6.1. Treatment effects on crime outcomes of the Pooled participants.

Statistic	Nonjuvenile arrests age 40	Crime cost age 40	Total charges age 40	Nonvictimless charges age 40
(i) Number of observations	123	123	123	123
(ii) Mean of the control group	8.8000	582.94	10,000	1,9692
(iii) Mean of the treatment group	5.1724	251.18	6,0862	0.8621
(iv) DIM (difference in means)	-3.6276	-331.76	-3,9138	-1.1072
(v) OLS estimate (with covariates)	-3.0631	-342.71	-3,2673	-0.9580
(vi) AIPW (augmented IPW) estimate	-3.0832	-285.47	-3,3760	-0.9484
(01) Asymptotic <i>p</i> -value for DIM	0.0149	0.0505	0.0218	0.0106
(02) Asymptotic <i>p</i> -value for OLS	0.0331	0.0583	0.0441	0.0193
(03) Asymptotic <i>p</i> -value for AIPW	0.0231	0.0686	0.0295	0.0142
(04) Bootstrap <i>p</i> -value for DIM	0.0087	0.0532	0.0149	0.0113
(05) Bootstrap <i>p</i> -value for OLS	0.0162	0.0549	0.0265	0.0167
(06) Bootstrap <i>p</i> -value for AIPW	0.0132	0.0639	0.0195	0.0137
(07) Permutation <i>p</i> -value for N-DIM	0.0176	0.0688	0.0236	0.0116
(08) Permutation <i>p</i> -value for N-OLS	0.0348	0.0696	0.0500	0.0244
(09) Permutation <i>p</i> -value for N-AIPW	0.0340	0.1132	0.0448	0.0284
(10) Permutation <i>p</i> -value for S-DIM	0.0184	0.0776	0.0252	0.0108
(11) Permutation <i>p</i> -value for S-OLS	0.0408	0.0808	0.0572	0.0220
(12) Permutation <i>p</i> -value for S-AIPW	0.0368	0.1028	0.0448	0.0212
(13) Worst-case max, <i>p</i> for N-DIM	0.0332	0.0950	0.0494	0.0454
(14) Worst-case max, <i>p</i> for N-OLS	0.0552	0.0949	0.0663	0.0609
(15) Worst-case max, <i>p</i> for N-AIPW	0.0573	0.1325	0.0718	0.0661
(16) Worst-case max, <i>p</i> for S-DIM	0.0398	0.1036	0.0528	0.0428
(17) Worst-case max, <i>p</i> for S-OLS	0.0672	0.0997	0.0754	0.0578
(18) Worst-case max, <i>p</i> for S-AIPW	0.0608	0.1230	0.0773	0.0588

Table S6.1. Continued

Statistic	Nonjuvenile arrests age 40	Crime cost age 40	Total charges age 40	Nonvictimless charges age 40
(19) Worst-case de Haan p for N-DIM	0.2530	0.1975	0.2181	0.1415
(20) Worst-case de Haan p for N-OLS	0.1199	0.1622	1.0000	0.1195
(21) Worst-case de Haan p : N-AIPW	0.0891	0.2152	0.1322	0.1649
(22) Worst-case de Haan p for S-DIM	0.1406	0.2588	0.2705	0.0997
(23) Worst-case de Haan p for S-OLS	0.1677	0.2118	0.8130	0.2019
(24) Worst-case de Haan p for S-AIPW	0.1962	0.3460	0.2066	0.1808
(25) Worst-case max. p for N-DIM	0.1328	0.1361	0.1361	0.1361
(26) Worst-case max. p for N-OLS	0.2209	0.2209	0.2209	0.2209
(27) Worst-case max. p for N-AIPW	0.2291	0.2291	0.2291	0.2291
(28) Worst-case max. p for S-DIM	0.1591	0.1591	0.1591	0.1591
(29) Worst-case max. p for S-OLS	0.2313	0.2313	0.2313	0.2313
(30) Worst-case max. p for S-AIPW	0.2353	0.2353	0.2353	0.2353
(31) Worst-case de Haan p for N-DIM	0.5924	0.5924	0.5924	0.5658
(32) Worst-case de Haan p for N-OLS	0.4780	0.4780	1.0000	0.4780
(33) Worst-case de Haan p : N-AIPW	0.3563	0.3967	0.3967	0.3967
(34) Worst-case de Haan p for S-DIM	0.4219	0.5175	0.5175	0.3988
(35) Worst-case de Haan p for S-OLS	0.6708	0.6708	0.8130	0.6708
(36) Worst-case de Haan p for S-AIPW	0.7231	0.7231	0.7231	0.7231

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above crime variables.

Table S6.2. Treatment effects on crime outcomes of the Male participants.

Statistic	Nonjuvenile arrests age 40	Crime cost age 40	Total charges age 40	Nonvictimless charges age 40
(i) Number of observations	72	72	72	72
(ii) Mean of the control group	11.718	77.90	13.385	3.0769
(iii) Mean of the treatment group	7.4545	424.68	9,0000	1.4848
(iv) DIM (difference in means)	-4.2634	-351.22	-4.3846	-1.5921
(v) OLS estimate (with covariates)	-3.8991	-399.30	-4.0113	-1.4684
(vi) AIPW (augmented IPW) estimate	-3.8947	-313.26	-4.1316	-1.4442
(01) Asymptotic <i>p</i> -value for DIM	0.0394	0.1328	0.0058	0.0233
(02) Asymptotic <i>p</i> -value for OLS	0.0615	0.1135	0.0928	0.0391
(03) Asymptotic <i>p</i> -value for AIPW	0.0461	0.1376	0.0678	0.0274
(04) Bootstrap <i>p</i> -value for DIM	0.0287	0.1269	0.0519	0.0168
(05) Bootstrap <i>p</i> -value for OLS	0.0439	0.1159	0.0731	0.0295
(06) Bootstrap <i>p</i> -value for AIPW	0.0368	0.1361	0.0579	0.0238
(07) Permutation <i>p</i> -value for N-DIM	0.0468	0.1576	0.0720	0.0268
(08) Permutation <i>p</i> -value for N-OLS	0.0668	0.1276	0.0996	0.0432
(09) Permutation <i>p</i> -value for N-AIPW	0.0652	0.1944	0.0904	0.0464
(10) Permutation <i>p</i> -value for S-DIM	0.0464	0.1600	0.0736	0.0256
(11) Permutation <i>p</i> -value for S-OLS	0.0816	0.1468	0.1076	0.0412
(12) Permutation <i>p</i> -value for S-AIPW	0.0668	0.1764	0.0920	0.0372
(13) Worst-case max. <i>p</i> for N-DIM	0.0702	0.1744	0.1088	0.0702
(14) Worst-case max. <i>p</i> for N-OLS	0.0871	0.1723	0.1226	0.0785
(15) Worst-case max. <i>p</i> for N-AIPW	0.0892	0.2262	0.1242	0.0907
(16) Worst-case max. <i>p</i> for S-DIM	0.0713	0.1746	0.1136	0.0688
(17) Worst-case max. <i>p</i> for S-OLS	0.0984	0.1731	0.1369	0.0864
(18) Worst-case max. <i>p</i> for S-AIPW	0.0951	0.2024	0.1216	0.0856

Table S6.2. Continued

Statistic	Nonjuvenile arrests age 40	Crime cost age 40	Total charges age 40	Nonvictimless charges age 40
(19) Worst-case de Haan p for N-DIM	0.2921	0.3513	0.3040	0.1961
(20) Worst-case de Haan p for N-OLS	0.5222	0.3202	0.2944	0.3251
(21) Worst-case de Haan p : N-AIPW	1.0000	0.3519	0.3295	0.2386
(22) Worst-case de Haan p for S-DIM	0.3823	0.3494	0.2610	0.3433
(23) Worst-case de Haan p for S-OLS	0.5534	0.3288	0.8165	0.2420
(24) Worst-case de Haan p for S-AIPW	1.0000	0.3880	0.2598	0.2792
(25) Worst-case max. p for N-DIM	0.2808	0.2808	0.2808	0.2808
(26) Worst-case max. p for N-OLS	0.3140	0.3140	0.3140	0.3140
(27) Worst-case max. p for N-AIPW	0.3569	0.3569	0.3569	0.3569
(28) Worst-case max. p for S-DIM	0.2752	0.2752	0.2752	0.2752
(29) Worst-case max. p for S-OLS	0.3457	0.3457	0.3457	0.3457
(30) Worst-case max. p for S-AIPW	0.3424	0.3424	0.3424	0.3424
(31) Worst-case de Haan p for N-DIM	0.8764	0.8764	0.8764	0.7845
(32) Worst-case de Haan p for N-OLS	1.0000	1.0000	1.0000	1.0000
(33) Worst-case de Haan p : N-AIPW	1.0000	0.9885	0.9885	0.9543
(34) Worst-case de Haan p for S-DIM	1.0000	1.0000	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	1.0000	0.9865	1.0000	0.9681
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above crime variables.

Table S6.3. Treatment effects on crime outcomes of the Female participants.

Statistic	Nonjuvenile arrests age 40	Crime cost age 40	Total charges age 40	Nonvictimless charges age 40
(i) Number of observations	51	51	51	51
(ii) Mean of the control group	4.4231	293.50	4.9231	0.3077
(iii) Mean of the treatment group	2.1600	22.165	2.2400	0.0400
(iv) DIM (difference in means)	-2.2631	-271.33	-2.6831	-0.2677
(v) OLS estimate (with covariates)	-2.1265	-283.60	-2.4812	-0.2749
(vi) AIPW (augmented IPW) estimate	-1.9376	-246.24	-2.3094	-0.2486
(01) Asymptotic <i>p</i> -value for DIM	0.0457	0.1365	0.0300	0.0339
(02) Asymptotic <i>p</i> -value for OLS	0.0691	0.1451	0.0497	0.0391
(03) Asymptotic <i>p</i> -value for AIPW	0.0657	0.1475	0.0439	0.0365
(04) Bootstrap <i>p</i> -value for DIM	0.0382	0.1198	0.0236	0.0286
(05) Bootstrap <i>p</i> -value for OLS	0.0702	0.1313	0.0494	0.0322
(06) Bootstrap <i>p</i> -value for AIPW	0.0795	0.1227	0.0528	0.0263
(07) Permutation <i>p</i> -value for N-DIM	0.0356	0.0648	0.0264	0.0408
(08) Permutation <i>p</i> -value for N-OLS	0.0616	0.1560	0.0436	0.0500
(09) Permutation <i>p</i> -value for N-AIPW	0.0868	0.2276	0.0640	0.0736
(10) Permutation <i>p</i> -value for S-DIM	0.0380	0.0188	0.0268	0.0364
(11) Permutation <i>p</i> -value for S-OLS	0.0624	0.0412	0.0440	0.0452
(12) Permutation <i>p</i> -value for S-AIPW	0.0880	0.2436	0.0580	0.0612
(13) Worst-case max. <i>p</i> for N-DIM	0.1044	0.1249	0.0841	0.0745
(14) Worst-case max. <i>p</i> for N-OLS	0.1358	0.2389	0.1131	0.0873
(15) Worst-case max. <i>p</i> for N-AIPW	0.1624	0.3335	0.1476	0.1175
(16) Worst-case max. <i>p</i> for S-DIM	0.1245	0.0601	0.1005	0.0677
(17) Worst-case max. <i>p</i> for S-OLS	0.1531	0.0886	0.1135	0.0733
(18) Worst-case max. <i>p</i> for S-AIPW	0.1695	0.2524	0.1526	0.0906

Table S6.3. Continued

Statistic	Nonjuvenile arrests age 40	Crime cost age 40	Total charges age 40	Nonvictimless charges age 40
(19) Worst-case de Haan p for N-DIM	0.3559	0.2163	0.2465	0.1558
(20) Worst-case de Haan p for N-OLS	0.2201	0.9147	0.2629	0.3694
(21) Worst-case de Haan p : N-AIPW	0.4444	0.5385	0.2208	0.4745
(22) Worst-case de Haan p for S-DIM	0.1625	0.0983	0.1661	0.2141
(23) Worst-case de Haan p for S-OLS	0.5326	0.1872	0.2109	0.3151
(24) Worst-case de Haan p for S-AIPW	0.4424	0.5508	0.2963	0.2574
(25) Worst-case max. p for N-DIM	0.2978	0.2978	0.2978	0.2978
(26) Worst-case max. p for N-OLS	0.3492	0.3492	0.3492	0.3492
(27) Worst-case max. p for N-AIPW	0.4700	0.4700	0.4700	0.4700
(28) Worst-case max. p for S-DIM	0.2403	0.2403	0.2403	0.2403
(29) Worst-case max. p for S-OLS	0.2931	0.2931	0.2931	0.2931
(30) Worst-case max. p for S-AIPW	0.4578	0.4578	0.4578	0.3625
(31) Worst-case de Haan p for N-DIM	0.6490	0.6490	0.6490	0.6233
(32) Worst-case de Haan p for N-OLS	0.8804	0.9147	0.8804	0.8804
(33) Worst-case de Haan p : N-AIPW	1.0000	1.0000	0.8834	1.0000
(34) Worst-case de Haan p for S-DIM	0.4874	0.3932	0.4874	0.4874
(35) Worst-case de Haan p for S-OLS	0.7490	0.7490	0.7490	0.7490
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all four of the above crime variables.

APPENDIX S7. INFERENCE ON EMPLOYMENT OUTCOMES AT AGE 19

Table S7.1. Treatment effects on employment outcomes (at age 19) of the Pooled participants.

Statistic	Currently employed age 19	Unemp. last year age 19	Unemp. last year age 19	Jobless months age 19
(i) Number of observations	123	123	123	112
(ii) Mean of the control group	0.3077	0.3077	0.3077	6.0175
(iii) Mean of the treatment group	0.5000	0.2414	0.2414	5.2545
(iv) DIM (difference in means) estimate	0.1923	-0.0663	-0.0663	-0.7630
(v) OLS estimate (with covariates)	0.1810	-0.0979	-0.0979	-1.2753
(vi) AIPW (augmented IPW) estimate	0.2093	-0.0868	-0.0868	-0.9401
(01) Asymptotic <i>p</i> -value for DIM	0.0194	0.2183	0.2183	0.3284
(02) Asymptotic <i>p</i> -value for OLS	0.0319	0.1471	0.1471	0.2513
(03) Asymptotic <i>p</i> -value for AIPW	0.0109	0.1613	0.1613	0.2946
(04) Bootstrap <i>p</i> -value for DIM	0.0105	0.1888	0.1888	0.3173
(05) Bootstrap <i>p</i> -value for OLS	0.0195	0.1152	0.1152	0.2330
(06) Bootstrap <i>p</i> -value for AIPW	0.0102	0.1518	0.1518	0.3186
(07) Permutation <i>p</i> -value for N-DIM	0.0180	0.2256	0.2256	0.3116
(08) Permutation <i>p</i> -value for N-OLS	0.0216	0.1256	0.1256	0.2260
(09) Permutation <i>p</i> -value for N-AIPW	0.0120	0.1564	0.1564	0.2912
(10) Permutation <i>p</i> -value for S-DIM	0.0200	0.2216	0.2216	0.3116
(11) Permutation <i>p</i> -value for S-OLS	0.0324	0.1492	0.1492	0.2500
(12) Permutation <i>p</i> -value for S-AIPW	0.0164	0.1804	0.1804	0.3016
(13) Worst-case maximum <i>p</i> for N-DIM	0.0923	0.2527	0.2527	0.3500
(14) Worst-case maximum <i>p</i> for N-OLS	0.1034	0.2034	0.2034	0.3442
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0849	0.2386	0.2386	0.4288
(16) Worst-case maximum <i>p</i> for S-DIM	0.0950	0.2502	0.2502	0.3500
(17) Worst-case maximum <i>p</i> for S-OLS	0.1184	0.2286	0.2286	0.3636
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0896	0.2567	0.2567	0.4503

Table S7.1. Continued

Statistic	Currently employed age 19	Unemp. last year age 19	Unemp. last year age 19	Jobless months age 19
(19) Worst-case de Haan p for N-DIM	0.3214	0.7560		0.8027
(20) Worst-case de Haan p for N-OLS	0.5687	0.5237		0.4588
(21) Worst-case de Haan p for N-AIPW	0.2872	0.4063		0.6057
(22) Worst-case de Haan p for S-DIM	0.3337	1.0000		0.8027
(23) Worst-case de Haan p for S-OLS	0.1894	0.4870		0.6318
(24) Worst-case de Haan p for S-AIPW	0.4314	0.4279		0.7359
(25) Worst-case maximum p for N-DIM	0.2770	0.5054		0.5054
(26) Worst-case maximum p for N-OLS	0.3101	0.4068		0.4068
(27) Worst-case maximum p for N-AIPW	0.2547	0.4772		0.4772
(28) Worst-case maximum p for S-DIM	0.2851	0.5004		0.5004
(29) Worst-case maximum p for S-OLS	0.3551	0.4571		0.4571
(30) Worst-case maximum p for S-AIPW	0.2689	0.5134		0.5134
(31) Worst-case de Haan p for N-DIM	0.9641	1.0000		1.0000
(32) Worst-case de Haan p for N-OLS	1.0000	1.0000		1.0000
(33) Worst-case de Haan p for N-AIPW	0.8617	0.8617		0.8617
(34) Worst-case de Haan p for S-DIM	1.0000	1.0000		1.0000
(35) Worst-case de Haan p for S-OLS	0.5682	0.9740		0.9740
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000		1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 19.

Table S7.2. Treatment effects on employment outcomes (at age 19) of the **Male** participants.

Statistic	Currently employed age 19	Unemp. last year age 19	Jobless months age 19
(i) Number of observations	72	72	70
(ii) Mean of the control group	0.4103	0.1282	3.8158
(iii) Mean of the treatment group	0.5455	0.2424	5.2812
(iv) DIM (difference in means) estimate	0.1352	0.1142	1.4655
(v) OLS estimate (with covariates)	0.1265	0.0881	1.4042
(vi) AIPW (augmented IPW) estimate	0.1471	0.1021	1.3665
(01) Asymptotic <i>p</i> -value for DIM	0.1340	0.1277	0.2378
(02) Asymptotic <i>p</i> -value for OLS	0.1700	0.2261	0.2673
(03) Asymptotic <i>p</i> -value for AIPW	0.1263	0.1817	0.2572
(04) Bootstrap <i>p</i> -value for DIM	0.1181	0.1003	0.2104
(05) Bootstrap <i>p</i> -value for OLS	0.1527	0.1985	0.2422
(06) Bootstrap <i>p</i> -value for AIPW	0.1315	0.1827	0.2501
(07) Permutation <i>p</i> -value for N-DIM	0.1364	0.1456	0.2592
(08) Permutation <i>p</i> -value for N-OLS	0.1444	0.2148	0.2840
(09) Permutation <i>p</i> -value for N-AIPW	0.1152	0.1800	0.2848
(10) Permutation <i>p</i> -value for S-DIM	0.1384	0.1408	0.2638
(11) Permutation <i>p</i> -value for S-OLS	0.1572	0.2512	0.2952
(12) Permutation <i>p</i> -value for S-AIPW	0.1292	0.2148	0.2928
(13) Worst-case maximum <i>p</i> for N-DIM	0.2775	0.2532	0.3280
(14) Worst-case maximum <i>p</i> for N-OLS	0.2915	0.2898	0.3145
(15) Worst-case maximum <i>p</i> for N-AIPW	0.2725	0.2494	0.3203
(16) Worst-case maximum <i>p</i> for S-DIM	0.2763	0.2471	0.3161
(17) Worst-case maximum <i>p</i> for S-OLS	0.3101	0.3194	0.3314
(18) Worst-case maximum <i>p</i> for S-AIPW	0.2999	0.2861	0.3243

Table S7.2. Continued

Statistic	Currently employed age 19	Unemp. last year age 19	Jobless months age 19
(19) Worst-case de Haan p for N-DIM	1.0000	0.4170	0.5675
(20) Worst-case de Haan p for N-OLS	0.4625	0.4628	1.0000
(21) Worst-case de Haan p for N-AIPW	0.5502	0.5881	0.7425
(22) Worst-case de Haan p for S-DIM	0.7335	0.5214	0.5369
(23) Worst-case de Haan p for S-OLS	0.4794	0.5422	1.0000
(24) Worst-case de Haan p for S-AIPW	0.4666	0.6064	1.0000
(25) Worst-case maximum p for N-DIM	0.7595	0.7595	0.7595
(26) Worst-case maximum p for N-OLS	0.8695	0.8695	0.8695
(27) Worst-case maximum p for N-AIPW	0.7481	0.7481	0.7481
(28) Worst-case maximum p for S-DIM	0.7413	0.7413	0.7413
(29) Worst-case maximum p for S-OLS	0.9303	0.9303	0.9303
(30) Worst-case maximum p for S-AIPW	0.8583	0.8583	0.8583
(31) Worst-case de Haan p for N-DIM	1.0000	1.0000	1.0000
(32) Worst-case de Haan p for N-OLS	1.0000	1.0000	1.0000
(33) Worst-case de Haan p for N-AIPW	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	1.0000	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 19.

Table S7.3. Treatment effects on employment outcomes (at age 19) of the Female participants.

Statistic	Currently employed age 19	Unemp. last year age 19	Jobless months age 19
(i) Number of observations	51	51	42
(ii) Mean of the control group	0.1538	0.5769	10.421
(iii) Mean of the treatment group	0.4400	0.2400	5.2174
(iv) DIM (difference in means) estimate	0.2862	-0.3369	-5.2037
(v) OLS estimate (with covariates)	0.2618	-0.3526	-4.3561
(vi) AIPW (augmented IPW) estimate	0.2971	-0.3536	-4.1965
(01) Asymptotic <i>p</i> -value for DIM	0.0144	0.0083	0.0514
(02) Asymptotic <i>p</i> -value for OLS	0.0216	0.0058	0.0896
(03) Asymptotic <i>p</i> -value for AIPW	0.0054	0.0029	0.0723
(04) Bootstrap <i>p</i> -value for DIM	0.0061	0.0048	0.0492
(05) Bootstrap <i>p</i> -value for OLS	0.0130	0.0043	0.0981
(06) Bootstrap <i>p</i> -value for AIPW	0.0048	0.0033	0.1386
(07) Permutation <i>p</i> -value for N-DIM	0.0156	0.0120	0.0528
(08) Permutation <i>p</i> -value for N-OLS	0.0264	0.0088	0.0784
(09) Permutation <i>p</i> -value for N-AIPW	0.0136	0.0100	0.0908
(10) Permutation <i>p</i> -value for S-DIM	0.0164	0.0152	0.0600
(11) Permutation <i>p</i> -value for S-OLS	0.0352	0.0120	0.0988
(12) Permutation <i>p</i> -value for S-AIPW	0.0152	0.0104	0.1140
(13) Worst-case maximum <i>p</i> for N-DIM	0.0594	0.0328	0.0783
(14) Worst-case maximum <i>p</i> for N-OLS	0.0855	0.0261	0.1419
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0641	0.0261	0.1469
(16) Worst-case maximum <i>p</i> for S-DIM	0.0562	0.0441	0.0838
(17) Worst-case maximum <i>p</i> for S-OLS	0.0863	0.0353	0.1579
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0578	0.0341	0.1886

Table S7.3. Continued

Statistic	Currently employed age 19	Unemp. last year age 19	Unemp. last year age 19	Jobless months age 19
(19) Worst-case de Haan p for N-DIM	0.0934	0.0815		0.2264
(20) Worst-case de Haan p for N-OLS	0.2711	0.0901		0.2328
(21) Worst-case de Haan p for N-AIPW	0.0807	0.0716		0.2345
(22) Worst-case de Haan p for S-DIM			0.1037	0.2354
(23) Worst-case de Haan p for S-OLS	1.0000	0.0657		0.4185
(24) Worst-case de Haan p for S-AIPW	0.2187	0.1878		0.4619
(25) Worst-case maximum p for N-DIM	0.1188	0.0983		0.1188
(26) Worst-case maximum p for N-OLS	0.1710	0.0782		0.1710
(27) Worst-case maximum p for N-AIPW	0.1282	0.0782		0.1469
(28) Worst-case maximum p for S-DIM	0.1323	0.1323		0.1323
(29) Worst-case maximum p for S-OLS	0.1727	0.1060		0.1727
(30) Worst-case maximum p for S-AIPW	0.1155	0.1022		0.1886
(31) Worst-case de Haan p for N-DIM	0.2446	0.2446		0.2446
(32) Worst-case de Haan p for N-OLS	0.4657	0.2703		0.4657
(33) Worst-case de Haan p for N-AIPW	0.2148	0.2148		0.2345
(34) Worst-case de Haan p for S-DIM	0.2697	0.2697		0.2697
(35) Worst-case de Haan p for S-OLS	1.0000	0.1971		0.8371
(36) Worst-case de Haan p for S-AIPW	0.5635	0.5635		0.5635

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 19.

APPENDIX S8. INFERENCE ON EMPLOYMENT OUTCOMES AT AGE 27

Table S8.1. Treatment effects on employment outcomes (at age 27) of the Pooled participants.

Statistic	Currently employed age 27	Unemp. last year age 27	Unemp. last year age 27	Jobless months age 27
(i) Number of observations	116	120	116	9,393.4
(ii) Mean of the control group	0.5574	0.3968	0.2456	5,636.4
(iii) Mean of the treatment group	0.6909			
(iv) DIM (difference in means) estimate	0.1335	-0.1512	-0.1512	-3.7571
(v) OLS estimate (with covariates)	0.1580	-0.1554	-0.1554	-3.4659
(vi) AIPW (augmented IPW) estimate	0.1410	-0.1589	-0.1589	-2.8021
(01) Asymptotic <i>p</i> -value for DIM	0.0854	0.0620	0.0365	
(02) Asymptotic <i>p</i> -value for OLS	0.0609	0.0693	0.0548	
(03) Asymptotic <i>p</i> -value for AIPW	0.0531	0.0535	0.0681	
(04) Bootstrap <i>p</i> -value for DIM	0.0590	0.0337	0.0177	
(05) Bootstrap <i>p</i> -value for OLS	0.0389	0.0384	0.0326	
(06) Bootstrap <i>p</i> -value for AIPW	0.0574	0.0373	0.0687	
(07) Permutation <i>p</i> -value for N-DIM	0.0780	0.0612	0.0280	
(08) Permutation <i>p</i> -value for N-OLS	0.0420	0.0504	0.0348	
(09) Permutation <i>p</i> -value for N-AIPW	0.0608	0.0552	0.0752	
(10) Permutation <i>p</i> -value for S-DIM	0.0796	0.0584	0.0300	
(11) Permutation <i>p</i> -value for S-OLS	0.0596	0.0708	0.0538	
(12) Permutation <i>p</i> -value for S-AIPW	0.0644	0.0696	0.0796	
(13) Worst-case maximum <i>p</i> for N-DIM	0.0859	0.1233	0.0702	
(14) Worst-case maximum <i>p</i> for N-OLS	0.0813	0.1244	0.0957	
(15) Worst-case maximum <i>p</i> for N-AIPW	0.1087	0.1395	0.1632	
(16) Worst-case maximum <i>p</i> for S-DIM	0.0872	0.1206	0.0824	
(17) Worst-case maximum <i>p</i> for S-OLS	0.1061	0.1577	0.1233	
(18) Worst-case maximum <i>p</i> for S-AIPW	0.1156	0.1604	0.1750	

Table S8.1. Continued

Statistic	Currently employed age 27	Unemp. last year age 27	Jobless months age 27
(19) Worst-case de Haan p for N-DIM	0.2298	0.2166	0.3282
(20) Worst-case de Haan p for N-OLS	0.1134	0.2965	0.2902
(21) Worst-case de Haan p for N-AIPW	0.2778	0.1761	0.8723
(22) Worst-case de Haan p for S-DIM	0.2298	0.2049	0.2355
(23) Worst-case de Haan p for S-OLS	0.1664	0.3416	1.0000
(24) Worst-case de Haan p for S-AIPW	0.4616	0.2959	0.7862
(25) Worst-case maximum p for N-DIM	0.2105	0.2105	0.2105
(26) Worst-case maximum p for N-OLS	0.2440	0.2440	0.2440
(27) Worst-case maximum p for N-AIPW	0.3260	0.3260	0.3260
(28) Worst-case maximum p for S-DIM	0.2473	0.2473	0.2473
(29) Worst-case maximum p for S-OLS	0.3182	0.3182	0.3182
(30) Worst-case maximum p for S-AIPW	0.3469	0.3469	0.3469
(31) Worst-case de Haan p for N-DIM	0.6497	0.6497	0.6497
(32) Worst-case de Haan p for N-OLS	0.3401	0.5805	0.5805
(33) Worst-case de Haan p for N-AIPW	0.5556	0.5283	0.8723
(34) Worst-case de Haan p for S-DIM	0.6146	0.6146	0.6146
(35) Worst-case de Haan p for S-OLS	0.4992	0.6833	1.0000
(36) Worst-case de Haan p for S-AIPW	0.9232	0.8877	0.9232

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 27.

Table S8.2. Treatment effects on employment outcomes (at age 27) of the **Male** participants.

Statistic	Currently employed age 27	Unemp. last year age 27	Jobless months age 27
(i) Number of observations	69	72	69
(ii) Mean of the control group	0.5641	0.3077	8.7949
(iii) Mean of the treatment group	0.6000	0.2424	5.1333
(iv) DIM (difference in means) estimate	0.0359	-0.0653	-3.6615
(v) OLS estimate (with covariates)	0.0839	-0.0682	-3.7638
(vi) AIPW (augmented IPW) estimate	0.0889	-0.0810	-3.8675
(01) Asymptotic <i>p</i> -value for DIM	0.3851	0.2778	0.0631
(02) Asymptotic <i>p</i> -value for OLS	0.2525	0.2832	0.0648
(03) Asymptotic <i>p</i> -value for AIPW	0.2156	0.2238	0.0438
(04) Bootstrap <i>p</i> -value for DIM	0.3819	0.2731	0.0585
(05) Bootstrap <i>p</i> -value for OLS	0.2477	0.2721	0.0566
(06) Bootstrap <i>p</i> -value for AIPW	0.2259	0.2190	0.0430
(07) Permutation <i>p</i> -value for N-DIM	0.4112	0.2984	0.0680
(08) Permutation <i>p</i> -value for N-OLS	0.2644	0.2868	0.0644
(09) Permutation <i>p</i> -value for N-AIPW	0.2476	0.2464	0.0580
(10) Permutation <i>p</i> -value for S-DIM	0.3988	0.2884	0.0668
(11) Permutation <i>p</i> -value for S-OLS	0.2764	0.2976	0.0736
(12) Permutation <i>p</i> -value for S-AIPW	0.2452	0.2388	0.0588
(13) Worst-case maximum <i>p</i> for N-DIM	0.3448	0.3594	0.0822
(14) Worst-case maximum <i>p</i> for N-OLS	0.3395	0.3953	0.1166
(15) Worst-case maximum <i>p</i> for N-AIPW	0.3272	0.3516	0.1049
(16) Worst-case maximum <i>p</i> for S-DIM	0.3304	0.3393	0.0866
(17) Worst-case maximum <i>p</i> for S-OLS	0.3484	0.4051	0.1382
(18) Worst-case maximum <i>p</i> for S-AIPW	0.3335	0.3488	0.1115

Table S8.2. Continued

Statistic	Currently employed age 27	Unemp. last year age 27	Unemp.	Jobless months age 27
(19) Worst-case de Haan p for N-DIM	0.3948	0.4518		0.1549
(20) Worst-case de Haan p for N-OLS	0.5398	0.6816		0.2943
(21) Worst-case de Haan p for N-AIPW	0.8452	0.4526		0.5338
(22) Worst-case de Haan p for S-DIM	0.6281	0.4793		0.3266
(23) Worst-case de Haan p for S-OLS	0.5446	0.6322		0.3767
(24) Worst-case de Haan p for S-AIPW	0.8446	0.5882		0.3548
(25) Worst-case maximum p for N-DIM	0.6896	0.6896		0.2467
(26) Worst-case maximum p for N-OLS	0.6790	0.6790		0.3498
(27) Worst-case maximum p for N-AIPW	0.6543	0.6543		0.3147
(28) Worst-case maximum p for S-DIM	0.6608	0.6608		0.2599
(29) Worst-case maximum p for S-OLS	0.6967	0.6967		0.4145
(30) Worst-case maximum p for S-AIPW	0.6670	0.6670		0.3344
(31) Worst-case de Haan p for N-DIM	0.7895	0.7895		0.4648
(32) Worst-case de Haan p for N-OLS	1.0000	1.0000		0.8828
(33) Worst-case de Haan p for N-AIPW	1.0000	1.0000		1.0000
(34) Worst-case de Haan p for S-DIM	0.9799	0.9799		0.9799
(35) Worst-case de Haan p for S-OLS	1.0000	1.0000		1.0000
(36) Worst-case de Haan p for S-AIPW	1.0000	1.0000		1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 27.

Table S8.3. Treatment effects on employment outcomes (at age 27) of the Female participants.

Statistic	Currently employed age 27	Unemp. last year age 27	Jobless months age 27
(i) Number of observations	47	48	47
(ii) Mean of the control group	0.5455	0.5417	10.455
(iii) Mean of the treatment group	0.8000	0.2500	6.2400
(iv) DIM (difference in means) estimate	0.2545	-0.2917	-4.2145
(v) OLS estimate (with covariates)	0.2341	-0.2839	-2.3964
(vi) AIPW (augmented IPW) estimate	0.2145	-0.2689	-1.2981
(01) Asymptotic <i>p</i> -value for DIM	0.0456	0.0309	0.1040
(02) Asymptotic <i>p</i> -value for OLS	0.0627	0.0528	0.2449
(03) Asymptotic <i>p</i> -value for AIPW	0.0471	0.0523	0.3338
(04) Bootstrap <i>p</i> -value for DIM	0.0199	0.0125	0.0669
(05) Bootstrap <i>p</i> -value for OLS	0.0386	0.0298	0.2218
(06) Bootstrap <i>p</i> -value for AIPW	0.0604	0.0457	0.3449
(07) Permutation <i>p</i> -value for N-DIM	0.0384	0.0284	0.0740
(08) Permutation <i>p</i> -value for N-OLS	0.0452	0.0416	0.1832
(09) Permutation <i>p</i> -value for N-AIPW	0.0584	0.0532	0.2856
(10) Permutation <i>p</i> -value for S-DIM	0.0416	0.0352	0.0828
(11) Permutation <i>p</i> -value for S-OLS	0.0608	0.0604	0.2076
(12) Permutation <i>p</i> -value for S-AIPW	0.0648	0.0728	0.2916
(13) Worst-case maximum <i>p</i> for N-DIM	0.0760	0.0908	0.1900
(14) Worst-case maximum <i>p</i> for N-OLS	0.0789	0.1088	0.3472
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0977	0.1256	0.4807
(16) Worst-case maximum <i>p</i> for S-DIM	0.0760	0.0962	0.1970
(17) Worst-case maximum <i>p</i> for S-OLS	0.1011	0.1429	0.3767
(18) Worst-case maximum <i>p</i> for S-AIPW	0.1048	0.1721	0.4821

Table S8.3. Continued

Statistic	Currently employed age 27	Unemp. last year age 27	Unemp.	Jobless months age 27
(19) Worst-case de Haan p for N-DIM	0.2472	0.2652	0.2991	0.2991
(20) Worst-case de Haan p for N-OLS	0.8273	0.2472	0.5399	0.5399
(21) Worst-case de Haan p for N-AIPW	0.2444	0.2558	0.7448	0.7448
(22) Worst-case de Haan p for S-DIM	0.1810	0.1323	0.2889	0.2889
(23) Worst-case de Haan p for S-OLS	0.2534	0.3629	0.5004	0.5004
(24) Worst-case de Haan p for S-AIPW	0.2521	1.0000	0.7544	0.7544
(25) Worst-case maximum p for N-DIM	0.2281	0.2281	0.2281	0.2281
(26) Worst-case maximum p for N-OLS	0.2368	0.2368	0.3472	0.3472
(27) Worst-case maximum p for N-AIPW	0.2932	0.2932	0.4807	0.4807
(28) Worst-case maximum p for S-DIM	0.2281	0.2281	0.2281	0.2281
(29) Worst-case maximum p for S-OLS	0.3033	0.3033	0.3767	0.3767
(30) Worst-case maximum p for S-AIPW	0.3143	0.3443	0.4821	0.4821
(31) Worst-case de Haan p for N-DIM	0.7416	0.7416	0.7416	0.7416
(32) Worst-case de Haan p for N-OLS	1.0000	0.7417	1.0000	1.0000
(33) Worst-case de Haan p for N-AIPW	0.7331	0.7331	0.7448	0.7448
(34) Worst-case de Haan p for S-DIM	0.3969	0.3969	0.3969	0.3969
(35) Worst-case de Haan p for S-OLS	0.7601	0.7601	0.7601	0.7601
(36) Worst-case de Haan p for S-AIPW	0.7562	1.0000	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 27.

APPENDIX S9. INFERENCE ON EMPLOYMENT OUTCOMES AT AGE 40

Table S9.1. Treatment effects on employment outcomes (at age 40) of the Pooled participants.

Statistic	Currently employed age 40	Unemp. last year age 40	Unemp. last year age 40	Jobless months age 40
(i) Number of observations	112	119	112	
(ii) Mean of the control group	0.6207	0.4426	8.5862	
(iii) Mean of the treatment group	0.7593	0.2759	5.7963	
(iv) DIM (difference in means) estimate	0.1386	-0.1668	-2.7899	
(v) OLS estimate (with covariates)	0.1714	-0.1603	-2.9605	
(vi) AIPW (augmented IPW) estimate	0.1489	-0.1638	-2.7613	
(01) Asymptotic <i>p</i> -value for DIM	0.0541	0.0270	0.0591	
(02) Asymptotic <i>p</i> -value for OLS	0.0349	0.0413	0.0523	
(03) Asymptotic <i>p</i> -value for AIPW	0.0443	0.0252	0.0509	
(04) Bootstrap <i>p</i> -value for DIM	0.0441	0.0252	0.0577	
(05) Bootstrap <i>p</i> -value for OLS	0.0250	0.0388	0.0537	
(06) Bootstrap <i>p</i> -value for AIPW	0.0487	0.0354	0.0707	
(07) Permutation <i>p</i> -value for N-DIM	0.0628	0.0320	0.0520	
(08) Permutation <i>p</i> -value for N-OLS	0.0288	0.0352	0.0368	
(09) Permutation <i>p</i> -value for N-AIPW	0.0576	0.0344	0.0504	
(10) Permutation <i>p</i> -value for S-DIM	0.0556	0.0316	0.0508	
(11) Permutation <i>p</i> -value for S-OLS	0.0440	0.0528	0.0504	
(12) Permutation <i>p</i> -value for S-AIPW	0.0600	0.0432	0.0596	
(13) Worst-case maximum <i>p</i> for N-DIM	0.0621	0.0369	0.0716	
(14) Worst-case maximum <i>p</i> for N-OLS	0.0528	0.0641	0.0805	
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0924	0.0722	0.1229	
(16) Worst-case maximum <i>p</i> for S-DIM	0.0609	0.0376	0.0792	
(17) Worst-case maximum <i>p</i> for S-OLS	0.0632	0.0835	0.0963	
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0987	0.0896	0.1243	

Table S9.1. Continued

Statistic	Currently employed age 40	Unemp. last year age 40	Unemp. last year age 40	Jobless months age 40
(19) Worst-case de Haan p for N-DIM	0.1357	0.1028	0.1159	0.1159
(20) Worst-case de Haan p for N-OLS	0.1642	0.4096	0.2693	0.2693
(21) Worst-case de Haan p for N-AIPW	0.1733	0.1391	0.1904	0.1904
(22) Worst-case de Haan p for S-DIM	0.2962	0.1159	0.1366	0.1366
(23) Worst-case de Haan p for S-OLS	0.1906	0.1971	0.3685	0.3685
(24) Worst-case de Haan p for S-AIPW	0.2624	0.1879	0.2152	0.2152
(25) Worst-case maximum p for N-DIM	0.1242	0.1108	0.1242	0.1242
(26) Worst-case maximum p for N-OLS	0.1584	0.1584	0.1584	0.1584
(27) Worst-case maximum p for N-AIPW	0.2166	0.2166	0.2166	0.2166
(28) Worst-case maximum p for S-DIM	0.1218	0.1129	0.1218	0.1218
(29) Worst-case maximum p for S-OLS	0.1896	0.1896	0.1896	0.1896
(30) Worst-case maximum p for S-AIPW	0.2688	0.2688	0.2688	0.2688
(31) Worst-case de Haan p for N-DIM	0.3083	0.3083	0.3083	0.3083
(32) Worst-case de Haan p for N-OLS	0.4926	0.5387	0.5387	0.5387
(33) Worst-case de Haan p for N-AIPW	0.4173	0.4173	0.4173	0.4173
(34) Worst-case de Haan p for S-DIM	0.3478	0.3478	0.3478	0.3478
(35) Worst-case de Haan p for S-OLS	0.5719	0.5719	0.5719	0.5719
(36) Worst-case de Haan p for S-AIPW	0.5638	0.5638	0.5638	0.5638

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 40.

Table S9.2. Treatment effects on employment outcomes (at age 40) of the **Male** participants.

Statistic	Currently employed age 40	Unemp. last year age 40	Unemp. last year age 40	Jobless months age 40
(i) Number of observations	66	72	66	
(ii) Mean of the control group	0.5000	0.4615	0.4615	10.750
(iii) Mean of the treatment group	0.7000	0.3636	0.3636	7.2333
(iv) DIM (difference in means) estimate	0.2000	-0.0979	-0.0979	-3.5167
(v) OLS estimate (with covariates)	0.2879	-0.1388	-0.1388	-4.7120
(vi) AIPW (augmented IPW) estimate	0.2655	-0.1425	-0.1425	-4.7578
(01) Asymptotic <i>p</i> -value for DIM	0.0520	0.1959	0.1959	0.0765
(02) Asymptotic <i>p</i> -value for OLS	0.0119	0.1173	0.1173	0.0285
(03) Asymptotic <i>p</i> -value for AIPW	0.0089	0.0843	0.0843	0.0154
(04) Bootstrap <i>p</i> -value for DIM	0.0363	0.1970	0.1970	0.0733
(05) Bootstrap <i>p</i> -value for OLS	0.0064	0.1159	0.1159	0.0256
(06) Bootstrap <i>p</i> -value for AIPW	0.0075	0.0957	0.0957	0.0200
(07) Permutation <i>p</i> -value for N-DIM	0.0676	0.2004	0.2004	0.0784
(08) Permutation <i>p</i> -value for N-OLS	0.0112	0.1020	0.1020	0.0188
(09) Permutation <i>p</i> -value for N-AIPW	0.0256	0.0920	0.0920	0.0188
(10) Permutation <i>p</i> -value for S-DIM	0.0620	0.1948	0.1948	0.0776
(11) Permutation <i>p</i> -value for S-OLS	0.0164	0.1208	0.1208	0.0272
(12) Permutation <i>p</i> -value for S-AIPW	0.0204	0.0912	0.0912	0.0188
(13) Worst-case maximum <i>p</i> for N-DIM	0.0612	0.1737	0.1737	0.0786
(14) Worst-case maximum <i>p</i> for N-OLS	0.0360	0.1748	0.1748	0.0666
(15) Worst-case maximum <i>p</i> for N-AIPW	0.0490	0.1695	0.1695	0.0609
(16) Worst-case maximum <i>p</i> for S-DIM	0.0595	0.1694	0.1694	0.0779
(17) Worst-case maximum <i>p</i> for S-OLS	0.0490	0.1945	0.1945	0.0840
(18) Worst-case maximum <i>p</i> for S-AIPW	0.0484	0.1695	0.1695	0.0650

Table S9.2. Continued

Statistic	Currently employed age 40	Unemp. last year age 40	Jobless months age 40
(19) Worst-case de Haan p for N-DIM	0.1106	0.5550	0.1053
(20) Worst-case de Haan p for N-OLS	0.1276	1.0000	0.2673
(21) Worst-case de Haan p for N-AIPW	0.0941	0.3893	0.1553
(22) Worst-case de Haan p for S-DIM	0.1149	1.0000	0.1220
(23) Worst-case de Haan p for S-OLS	0.0876	0.4670	0.7203
(24) Worst-case de Haan p for S-AIPW	0.0971	0.5219	0.1341
(25) Worst-case maximum p for N-DIM	0.1835	0.1835	0.1835
(26) Worst-case maximum p for N-OLS	0.1080	0.1748	0.1333
(27) Worst-case maximum p for N-AIPW	0.1470	0.1695	0.1470
(28) Worst-case maximum p for S-DIM	0.1784	0.1784	0.1784
(29) Worst-case maximum p for S-OLS	0.1470	0.1945	0.1681
(30) Worst-case maximum p for S-AIPW	0.1451	0.1695	0.1451
(31) Worst-case de Haan p for N-DIM	0.3158	0.5550	0.3158
(32) Worst-case de Haan p for N-OLS	0.3828	1.0000	0.5347
(33) Worst-case de Haan p for N-AIPW	0.2823	0.3893	0.3105
(34) Worst-case de Haan p for S-DIM	0.3446	1.0000	0.3446
(35) Worst-case de Haan p for S-OLS	0.2627	0.9339	0.9339
(36) Worst-case de Haan p for S-AIPW	0.2912	0.5219	0.2912

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 40.

Table S9.3. Treatment effects on employment outcomes (at age 40) of the Female participants.

Statistic	Currently employed age 40	Unemp. last year age 40	Jobless months age 40
(i) Number of observations	46	47	46
(ii) Mean of the control group	0.8182	0.4091	5.0455
(iii) Mean of the treatment group	0.8333	0.1600	4.0000
(iv) DIM (difference in means) estimate	0.0152	-0.2491	-1.0455
(v) OLS estimate (with covariates)	-0.0242	-0.2096	-0.0475
(vi) AIPW (augmented IPW) estimate	-0.0157	-0.1938	0.0574
(01) Asymptotic <i>p</i> -value for DIM	0.4474	0.0308	0.3341
(02) Asymptotic <i>p</i> -value for OLS	0.4346	0.0789	0.4931
(03) Asymptotic <i>p</i> -value for AIPW	0.4536	0.0807	0.4910
(04) Bootstrap <i>p</i> -value for DIM	0.4419	0.0245	0.3268
(05) Bootstrap <i>p</i> -value for OLS	0.4309	0.0839	0.4934
(06) Bootstrap <i>p</i> -value for AIPW	0.4586	0.1079	0.4927
(07) Permutation <i>p</i> -value for N-DIM	0.4380	0.0300	0.3104
(08) Permutation <i>p</i> -value for N-OLS	0.4580	0.0728	0.4520
(09) Permutation <i>p</i> -value for N-AIPW	0.4812	0.0960	0.4712
(10) Permutation <i>p</i> -value for S-DIM	0.4356	0.0412	0.3120
(11) Permutation <i>p</i> -value for S-OLS	0.4692	0.1008	0.4524
(12) Permutation <i>p</i> -value for S-AIPW	0.4912	0.1324	0.4700
(13) Worst-case maximum <i>p</i> for N-DIM	0.4883	0.0797	0.4408
(14) Worst-case maximum <i>p</i> for N-OLS	0.6181	0.1146	0.5855
(15) Worst-case maximum <i>p</i> for N-AIPW	0.6331	0.1382	0.6326
(16) Worst-case maximum <i>p</i> for S-DIM	0.4818	0.1000	0.4408
(17) Worst-case maximum <i>p</i> for S-OLS	0.6261	0.1505	0.5855
(18) Worst-case maximum <i>p</i> for S-AIPW	0.6385	0.1892	0.6326

Table S9.3. Continued

Statistic	Currently employed age 40	Unemp. last year age 40	Unemp. last year age 40	Jobless months age 40
(19) Worst-case de Haan p for N-DIM	1.0000	0.1172	0.5519	0.5519
(20) Worst-case de Haan p for N-OLS	0.9359	0.1834	0.7339	0.7339
(21) Worst-case de Haan p for N-AIPW	0.8691	0.3366	1.0000	1.0000
(22) Worst-case de Haan p for S-DIM	1.0000	0.2254	0.5519	0.5519
(23) Worst-case de Haan p for S-OLS	0.8841	0.4333	0.7287	0.7287
(24) Worst-case de Haan p for S-AIPW	1.0000	0.3070	1.0000	1.0000
(25) Worst-case maximum p for N-DIM	0.8816	0.2392	0.8816	0.8816
(26) Worst-case maximum p for N-OLS	1.0000	0.3438	1.0000	1.0000
(27) Worst-case maximum p for N-AIPW	1.0000	0.4145	1.0000	1.0000
(28) Worst-case maximum p for S-DIM	0.8816	0.3001	0.8816	0.8816
(29) Worst-case maximum p for S-OLS	1.0000	0.4514	1.0000	1.0000
(30) Worst-case maximum p for S-AIPW	1.0000	0.5675	1.0000	1.0000
(31) Worst-case de Haan p for N-DIM	1.0000	0.3517	1.0000	1.0000
(32) Worst-case de Haan p for N-OLS	1.0000	0.5503	1.0000	1.0000
(33) Worst-case de Haan p for N-AIPW	1.0000	1.0000	1.0000	1.0000
(34) Worst-case de Haan p for S-DIM	1.0000	0.6761	1.0000	1.0000
(35) Worst-case de Haan p for S-OLS	1.0000	1.0000	1.0000	1.0000
(36) Worst-case de Haan p for S-AIPW	1.0000	0.9211	1.0000	1.0000

Note: See Online Appendix S1 for additional notes on the table rows. The block of variables used for Holm stepdown adjustment consists of all three of the above employment variables measured at age 40.