

# Appendix: Analyzing the Short- and Long-term Effects of Early Childhood Education on Multiple Dimensions of Human Development

Jorge Luis García  
The University of Chicago

James J. Heckman  
American Bar Foundation  
The University of Chicago

Andrés Hojman  
The University of Chicago

Duncan Ermini Leaf  
University of Southern California

María José Prados  
University of Southern California

Joshua Shea  
The University of Chicago

Jake C. Torcasso  
The University of Chicago

First Draft: January 5, 2016  
This Draft: July 29, 2016

## Contents

<b>Appendix A Background</b>	<b>1</b>
A.1 Overview	1
A.2 Eligibility Criteria and Populations Served	1
A.3 Randomization Protocol and Compromises	5
A.3.1 ABC	8
A.3.2 CARE	13
A.4 Program Description and Content	15
A.4.1 Goals	15
A.4.2 Daily Schedule	15
A.4.3 Program Staff and Physical Space	16
A.4.4 Approach to Child Development	17
A.4.5 Medical Care and Nutrition	19
A.4.6 School-age Treatment	20
A.5 Control Substitution	21

A.5.1	Regulation . . . . .	23
A.5.2	Costs . . . . .	24
A.6	Data . . . . .	26
<b>Appendix B Background Variables and Tested Outcomes</b>		<b>43</b>
B.1	Controlling for Background Variables . . . . .	43
B.2	Tested Outcomes . . . . .	44
<b>Appendix C Results</b>		<b>49</b>
C.1	Combining Functions, Aggregated . . . . .	51
C.2	Combining Functions, by Category . . . . .	52
C.3	Treatment Effects for Pooled Sample . . . . .	58
C.4	Treatment Effects for Male Sample . . . . .	63
C.5	Treatment Effects for Female Sample . . . . .	68
C.6	Treatment Effects for Pooled Sample, Step Down . . . . .	73
C.7	Treatment Effects for Male Sample, Step Down . . . . .	78
C.8	Treatment Effects for Female Sample, Step Down . . . . .	83
<b>Appendix D Alternative Evaluation Methodologies</b>		<b>88</b>
D.1	Instrumental Variables . . . . .	89
D.1.1	Model and Conditions . . . . .	89
D.1.2	Instrumental Variables in Practice . . . . .	89
D.1.3	Sets of Instruments . . . . .	90
D.1.4	Specifications for the Instruments . . . . .	91
D.1.5	Functional Forms of Enrollment in Alternative Preschool . . . . .	91
D.2	Results . . . . .	92
D.2.1	Main Specification . . . . .	92
D.2.2	Varying the Sets of Instruments . . . . .	94
D.2.3	Varying the Specification of the Instruments . . . . .	96
D.2.4	Varying the Parameterization of Alternative Preschool Take-up . . . . .	100
D.3	Control Functions . . . . .	104
D.3.1	Setup . . . . .	104
D.3.2	Identification . . . . .	106
D.4	Estimates . . . . .	107
<b>Appendix E Assessing Common Critiques to ABC</b>		<b>109</b>
E.1	Treatment Effects vs. Early Differences Between Groups . . . . .	110
E.2	Fadeout of Treatment Effects . . . . .	114
E.3	Compromised Randomization . . . . .	114
<b>Appendix F Calculating the Benefit-to-cost Ratio and the IRR</b>		<b>117</b>
<b>Appendix G Accounting for Attrition</b>		<b>118</b>

<b>Appendix H</b>	<b>Details on Interpolation and Extrapolation</b>	<b>120</b>
H.1	Identifying Treatment Effects . . . . .	120
H.1.1	Fully Observed Outcomes . . . . .	120
H.1.2	Partially Observed Outcomes . . . . .	121
H.1.3	Fully Missing Outcomes . . . . .	123
H.2	Internal Rate of Return . . . . .	134
H.3	Computing the Benefit-Cost Ratio . . . . .	136
<b>Appendix I</b>	<b>Program Costs of ABC</b>	<b>137</b>
<b>Appendix J</b>	<b>Costs of Education</b>	<b>140</b>
J.1	Measuring Lifetime Educational Attainment . . . . .	141
J.2	Cost of Education . . . . .	143
J.3	Non-monetary Benefits of Education . . . . .	144
<b>Appendix K</b>	<b>Quantifying the Benefits in Crime Reduction</b>	<b>145</b>
K.1	Data Description . . . . .	145
K.1.1	ABC and CARE Datasets . . . . .	145
K.1.2	Auxiliary Datasets . . . . .	146
K.1.3	Crime Categories . . . . .	147
K.2	Methodology for Estimating Crime Costs . . . . .	148
K.2.1	Count Arrests and Sentences . . . . .	149
K.2.2	Construct Predictions . . . . .	152
K.3	Victimization Inflation . . . . .	161
K.3.1	Construction of the Total Number of Victims in the U.S. . . . .	162
K.3.2	Construction of the Total Number of Arrests in the U.S. . . . .	163
K.3.3	Victimization Inflation Factors . . . . .	163
K.3.4	Effects on Number of Crimes, After Victimization Inflation . . . . .	165
K.4	Literature on Costs of Specific Crimes . . . . .	166
K.4.1	Classifying the Costs of Crime . . . . .	167
K.4.2	Bottom-up (BU) Methodologies . . . . .	168
K.4.3	Top-down (TD) Methodologies . . . . .	168
K.4.4	Costs Used in this Study . . . . .	169
K.4.5	Timing of Effects: Incidence vs. Prevalence . . . . .	171
K.4.6	Costs of Imprisonment . . . . .	171
K.5	Effects on Costs of Crime . . . . .	171
K.5.1	Effects on Costs Before Victimization Inflation . . . . .	171
K.5.2	Effects on Costs After Victimization Inflation . . . . .	172
K.6	Sensitivity Analyses Using Alternative Cost Estimates . . . . .	174
<b>Appendix L</b>	<b>Interpolation of Parental Income</b>	<b>176</b>
<b>Appendix M</b>	<b>Extrapolation of Individual Income</b>	<b>177</b>
M.1	Data Sources . . . . .	177
M.1.1	NLSY . . . . .	177

M.1.2	CNLSY . . . . .	178
M.1.3	PSID . . . . .	179
<b>Appendix N</b>	<b>Health Outcomes in the Future America Model (FAM)</b>	<b>180</b>
N.1	Background and Description of FAM . . . . .	180
N.2	Data Sources . . . . .	184
N.2.1	PSID . . . . .	184
N.2.2	HRS . . . . .	185
N.2.3	MCBS . . . . .	186
N.2.4	MEPS . . . . .	186
N.2.5	NHANES . . . . .	187
N.2.6	ABC and CARE . . . . .	187
N.3	Methods and Analysis . . . . .	188
N.3.1	ABC and CARE Data Assumptions and Imputations . . . . .	188
N.3.2	FAM Models and Estimation . . . . .	193
N.3.3	Transition Models . . . . .	197
N.3.4	FAM simulation . . . . .	203
N.3.5	Bootstrap Sampling Strategy . . . . .	206
N.3.6	Medical Costs Before Age 30 Interview . . . . .	207
N.4	Results . . . . .	209
<b>Appendix O</b>	<b>Expanding the Analyzed Outcomes</b>	<b>216</b>
O.1	Outcomes by Category . . . . .	218
<b>Appendix P</b>	<b>Sensitivity Analysis</b>	<b>224</b>
P.1	Varying the Discount Rate . . . . .	224
P.2	Varying Deadweight Loss . . . . .	226
P.3	Varying Component Magnitudes . . . . .	230
<b>Appendix Q</b>	<b>Back-of-the-Envelope Cost Benefit Analysis of ABC</b>	<b>251</b>

## List of Figures

A.1	High-risk Index Distribution, ABC . . . . .	3
A.2	High-risk Index Distribution, CARE . . . . .	4
A.3	Randomization Protocol and Treatment Compliance, ABC . . . . .	6
A.4	Randomization Protocol and Treatment Compliance, CARE . . . . .	7
A.5	Average Number of Months in Alternative Preschool, ABC Control Group . . . . .	22
A.6	Average Number of Months in Alternative Preschool, CARE Control and Family Education Groups . . . . .	23
D.1	Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Instrumental Variables . . . . .	93
D.2	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Instrumental Variables . . . . .	94

D.3	Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables . . .	95
D.4	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables . . . . .	96
D.5	Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications . . . . .	97
D.6	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications . . . . .	98
D.7	Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications . . . . .	99
D.8	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications . . . . .	100
D.9	Effect of Center-based Childcare on IQ Scores, Accounting for an Endogenous Indicator of Take-up of Alternative Preschool . . . . .	101
D.10	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for an Endogenous Indicator of Take-up of Alternative Preschool . . . . .	102
D.11	Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool . . . . .	103
D.12	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool . . . . .	104
D.13	Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool . . . . .	108
D.14	Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool . . . . .	109
E.1	Treatment Effects on IQ, ABC . . . . .	113
H.1	Support of ABC and Auxiliary Data . . . . .	124
I.1	Primary-source Document Costs, Example . . . . .	138
I.2	Primary-source Document Costs, Personnel Wages . . . . .	139
J.1	Additional Education After Age 30 . . . . .	142
K.1	Counts of Arrests and Sentences . . . . .	152
K.2	Constructed Predictions . . . . .	161
K.3	Victims-to-Arrests Ratios by Crime . . . . .	164
K.4	Arrests-to-Sentences Ratio by Crime . . . . .	165
K.5	Effects on Number of Crimes, After Victimization Inflation . . . . .	166
K.6	Costs of Crime Before Victimization Inflation . . . . .	172
K.7	Costs of Crime After Victimization Inflation . . . . .	173
K.8	Different Cost Schedules . . . . .	175
K.9	Effect of Discounting Crime Costs . . . . .	176
N.1	Year-over-Year Excess Real Growth in Medical Costs . . . . .	205
N.2	Predicted Quality-Adjusted Life Years . . . . .	211

N.3	Predicted Total Medical Costs . . . . .	212
N.4	Predicted Survival . . . . .	214
P.1	Benefit-Cost Ratio vs. Discount Rate . . . . .	225
P.2	Internal Rate of Return vs. Deadweight Loss . . . . .	227
P.3	Benefit-Cost Ratio vs. Deadweight Loss . . . . .	229
P.4	Internal Rate of Return vs. Components, Females . . . . .	231
P.1	Internal Rate of Return vs. Components, Males . . . . .	237
P.-2	Benefit-Cost Ratio vs. Components, Females . . . . .	242
P.-5	Benefit-Cost Ratio vs. Components, Males . . . . .	247
Q.1	Benefit-to-cost Ratio and the Earnings Returns to a Kindergarten IQ Test . . . . .	253
Q.2	Benefit-to-cost Ratio and the Discounted Net Present Value of Earnings . . . . .	253

## List of Tables

A.1	Randomization Compromises, ABC . . . . .	10
A.2	Randomization Compromises, CARE . . . . .	14
A.3	Child-Staff Ratios for North Carolina, FIDCR, and Actual ABC and CARE Ratios . . . . .	24
A.4	Early Childhood Data (Part I) . . . . .	27
A.5	Early Childhood Data (Part II) . . . . .	28
A.6	Childhood and Adolescence Data (Part I) . . . . .	29
A.7	Childhood and Adolescence Data (Part II) . . . . .	30
A.8	Adulthood Data (Part I) . . . . .	31
A.9	Adulthood Data (Part II) . . . . .	32
A.10	First-phase Treatment vs. Control Groups, ABC . . . . .	34
A.11	First-phase Treatment vs. Control Groups, ABC Cohort 1 . . . . .	34
A.12	First-phase Treatment vs. Control Groups, ABC Cohort 2 . . . . .	35
A.13	First-phase Treatment vs. Control Groups, ABC Cohort 3 . . . . .	35
A.14	First-phase Treatment vs. Control Groups, ABC Cohort 4 . . . . .	36
A.15	Second-phase Treatment vs. Control Groups, ABC . . . . .	36
A.16	Observed vs. Attritted Children, ABC . . . . .	37
A.17	First-phase Treatment vs. Control Groups, Dropping Attritted Children, ABC . . . . .	37
A.18	First-phase Treatment vs. Control Groups, Subjects Completing the Health Follow-up, ABC . . . . .	38
A.19	CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare . . . . .	39
A.20	CARE Baseline Characteristics, Control vs. Family Education . . . . .	39
A.21	CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare, Cohort 5 . . . . .	40
A.22	CARE Baseline Characteristics, Control vs. Family Education, Cohort 5 . . . . .	40
A.23	CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare, Cohort 6 . . . . .	41
A.24	CARE Baseline Characteristics, Control vs. Family Education, Cohort 6 . . . . .	41

A.25 CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare Subjects Completing the Health Follow-up . . . . .	42
A.26 CARE Baseline Characteristics, Control vs. Family Education Subjects Completing the Health Follow-up . . . . .	42
B.1 Background Variables . . . . .	44
B.2 Outcome Variables . . . . .	44
C.1 Combining Functions, Pooled Sample . . . . .	51
C.2 Combining Functions, Male Sample . . . . .	51
C.3 Combining Functions, Female Sample . . . . .	51
C.4 Combining Functions by Category, Pooled Sample . . . . .	52
C.5 Combining Functions by Category   10% Significance, Pooled Sample . . . . .	53
C.6 Combining Functions by Category, Male Sample . . . . .	54
C.7 Combining Functions by Category   10% Significance, Male Sample . . . . .	55
C.8 Combining Functions by Category, Female Sample . . . . .	56
C.9 Combining Functions by Category   10% Significance, Female Sample . . . . .	57
C.10 Treatment Effects on IQ Scores, Pooled Sample . . . . .	58
C.11 Treatment Effects on Achievement Scores, Pooled Sample . . . . .	58
C.12 Treatment Effects on HOME Scores, Pooled Sample . . . . .	59
C.13 Treatment Effects on Parent Income, Pooled Sample . . . . .	59
C.14 Treatment Effects on Mother’s Employment, Pooled Sample . . . . .	59
C.15 Treatment Effects on Father at Home, Pooled Sample . . . . .	59
C.16 Treatment Effects on Education, Pooled Sample . . . . .	60
C.17 Treatment Effects on Employment and Income, Pooled Sample . . . . .	60
C.18 Treatment Effects on Crime, Pooled Sample . . . . .	60
C.19 Treatment Effects on Tobacco, Drugs, Alcohol, Pooled Sample . . . . .	60
C.20 Treatment Effects on Self-Reported Health, Pooled Sample . . . . .	61
C.21 Treatment Effects on Hypertension, Pooled Sample . . . . .	61
C.22 Treatment Effects on Cholesterol, Pooled Sample . . . . .	61
C.23 Treatment Effects on Diabetes, Pooled Sample . . . . .	61
C.24 Treatment Effects on Vitamin D Deficiency, Pooled Sample . . . . .	61
C.25 Treatment Effects on Obesity, Pooled Sample . . . . .	62
C.26 Treatment Effects on Mental Health, Pooled Sample . . . . .	62
C.27 Treatment Effects on IQ Scores, Male Sample . . . . .	63
C.28 Treatment Effects on Achievement Scores, Male Sample . . . . .	63
C.29 Treatment Effects on HOME Scores, Male Sample . . . . .	64
C.30 Treatment Effects on Parent Income, Male Sample . . . . .	64
C.31 Treatment Effects on Mother’s Employment, Male Sample . . . . .	64
C.32 Treatment Effects on Father at Home, Male Sample . . . . .	64
C.33 Treatment Effects on Education, Male Sample . . . . .	65
C.34 Treatment Effects on Employment and Income, Male Sample . . . . .	65
C.35 Treatment Effects on Crime, Male Sample . . . . .	65
C.36 Treatment Effects on Tobacco, Drugs, Alcohol, Male Sample . . . . .	65
C.37 Treatment Effects on Self-Reported Health, Male Sample . . . . .	66
C.38 Treatment Effects on Hypertension, Male Sample . . . . .	66
C.39 Treatment Effects on Cholesterol, Male Sample . . . . .	66

C.40 Treatment Effects on Diabetes, Male Sample . . . . .	66
C.41 Treatment Effects on Vitamin D Deficiency, Male Sample . . . . .	66
C.42 Treatment Effects on Obesity, Male Sample . . . . .	67
C.43 Treatment Effects on Mental Health, Male Sample . . . . .	67
C.44 Treatment Effects on IQ Scores, Female Sample . . . . .	68
C.45 Treatment Effects on Achievement Scores, Female Sample . . . . .	68
C.46 Treatment Effects on HOME Scores, Female Sample . . . . .	69
C.47 Treatment Effects on Parent Income, Female Sample . . . . .	69
C.48 Treatment Effects on Mother’s Employment, Female Sample . . . . .	69
C.49 Treatment Effects on Father at Home, Female Sample . . . . .	69
C.50 Treatment Effects on Education, Female Sample . . . . .	70
C.51 Treatment Effects on Employment and Income, Female Sample . . . . .	70
C.52 Treatment Effects on Crime, Female Sample . . . . .	70
C.53 Treatment Effects on Tobacco, Drugs, Alcohol, Female Sample . . . . .	70
C.54 Treatment Effects on Self-Reported Health, Female Sample . . . . .	71
C.55 Treatment Effects on Hypertension, Female Sample . . . . .	71
C.56 Treatment Effects on Cholesterol, Female Sample . . . . .	71
C.57 Treatment Effects on Diabetes, Female Sample . . . . .	71
C.58 Treatment Effects on Vitamin D Deficiency, Female Sample . . . . .	71
C.59 Treatment Effects on Obesity, Female Sample . . . . .	72
C.60 Treatment Effects on Mental Health, Female Sample . . . . .	72
C.61 Treatment Effects on IQ Scores, Pooled Sample . . . . .	73
C.62 Treatment Effects on Achievement Scores, Pooled Sample . . . . .	73
C.63 Treatment Effects on HOME Scores, Pooled Sample . . . . .	74
C.64 Treatment Effects on Parent Income, Pooled Sample . . . . .	74
C.65 Treatment Effects on Mother’s Employment, Pooled Sample . . . . .	74
C.66 Treatment Effects on Father at Home, Pooled Sample . . . . .	74
C.67 Treatment Effects on Education, Pooled Sample . . . . .	75
C.68 Treatment Effects on Employment and Income, Pooled Sample . . . . .	75
C.69 Treatment Effects on Crime, Pooled Sample . . . . .	75
C.70 Treatment Effects on Tobacco, Drugs, Alcohol, Pooled Sample . . . . .	75
C.71 Treatment Effects on Self-Reported Health, Pooled Sample . . . . .	76
C.72 Treatment Effects on Hypertension, Pooled Sample . . . . .	76
C.73 Treatment Effects on Cholesterol, Pooled Sample . . . . .	76
C.74 Treatment Effects on Diabetes, Pooled Sample . . . . .	76
C.75 Treatment Effects on Vitamin D Deficiency, Pooled Sample . . . . .	76
C.76 Treatment Effects on Obesity, Pooled Sample . . . . .	77
C.77 Treatment Effects on Mental Health, Pooled Sample . . . . .	77
C.78 Treatment Effects on IQ Scores, Male Sample . . . . .	78
C.79 Treatment Effects on Achievement Scores, Male Sample . . . . .	78
C.80 Treatment Effects on HOME Scores, Male Sample . . . . .	79
C.81 Treatment Effects on Parent Income, Male Sample . . . . .	79
C.82 Treatment Effects on Mother’s Employment, Male Sample . . . . .	79
C.83 Treatment Effects on Father at Home, Male Sample . . . . .	79
C.84 Treatment Effects on Education, Male Sample . . . . .	80

C.85 Treatment Effects on Employment and Income, Male Sample . . . . .	80
C.86 Treatment Effects on Crime, Male Sample . . . . .	80
C.87 Treatment Effects on Tobacco, Drugs, Alcohol, Male Sample . . . . .	80
C.88 Treatment Effects on Self-Reported Health, Male Sample . . . . .	81
C.89 Treatment Effects on Hypertension, Male Sample . . . . .	81
C.90 Treatment Effects on Cholesterol, Male Sample . . . . .	81
C.91 Treatment Effects on Diabetes, Male Sample . . . . .	81
C.92 Treatment Effects on Vitamin D Deficiency, Male Sample . . . . .	81
C.93 Treatment Effects on Obesity, Male Sample . . . . .	82
C.94 Treatment Effects on Mental Health, Male Sample . . . . .	82
C.95 Treatment Effects on IQ Scores, Female Sample . . . . .	83
C.96 Treatment Effects on Achievement Scores, Female Sample . . . . .	83
C.97 Treatment Effects on HOME Scores, Female Sample . . . . .	84
C.98 Treatment Effects on Parent Income, Female Sample . . . . .	84
C.99 Treatment Effects on Mother’s Employment, Female Sample . . . . .	84
C.100 Treatment Effects on Father at Home, Female Sample . . . . .	84
C.101 Treatment Effects on Education, Female Sample . . . . .	85
C.102 Treatment Effects on Employment and Income, Female Sample . . . . .	85
C.103 Treatment Effects on Crime, Female Sample . . . . .	85
C.104 Treatment Effects on Tobacco, Drugs, Alcohol, Female Sample . . . . .	85
C.105 Treatment Effects on Self-Reported Health, Female Sample . . . . .	86
C.106 Treatment Effects on Hypertension, Female Sample . . . . .	86
C.107 Treatment Effects on Cholesterol, Female Sample . . . . .	86
C.108 Treatment Effects on Diabetes, Female Sample . . . . .	86
C.109 Treatment Effects on Vitamin D Deficiency, Female Sample . . . . .	86
C.110 Treatment Effects on Obesity, Female Sample . . . . .	87
C.111 Treatment Effects on Mental Health, Female Sample . . . . .	87
E.1 Treatment - Control Mean Difference by Cohort, ABC . . . . .	110
E.2 Treatment - Control Mean Difference, IHDP . . . . .	112
E.3 Assessing Non-compliance in ABC, Exercise 1 . . . . .	116
E.4 Assessing Non-compliance in ABC, Exercise 2 . . . . .	117
H.1 Variables Estimated without IPW . . . . .	130
H.2 Model Selection for Attrition IPW Pooled . . . . .	131
H.3 Model Selection for Attrition IPW Females . . . . .	132
H.4 Model Selection for Attrition IPW Males . . . . .	133
H.5 Components of Benefits and Costs . . . . .	136
J.1 Yearly Individual Education Costs . . . . .	144
K.1 Crime Categories . . . . .	148
K.2 NCDPS Regressions of Ages 35–40 on Ages 16–35, Females . . . . .	155
K.3 NCDPS Regressions of Ages 35–40 on Ages 16–35, Males . . . . .	156
K.4 NCDPS Regressions of Ages 40–45 on Ages 16–35, Females . . . . .	157
K.5 NCDPS Regressions of Ages 40–45 on Ages 16–35, Males . . . . .	158
K.6 NCDPS Regressions of Ages 45–50 on Ages 16–35, Females . . . . .	159
K.7 NCDPS Regressions from Ages 45–50 on Ages 16–35, Males . . . . .	160
K.8 Unitary Costs of Crime for Victims . . . . .	170

N.1	Health State Transitions, Age $a$ as Predictor of Age $a + 1$ . . . . .	200
N.2	Prevalence of Disease by Age 65 (%) . . . . .	209
N.3	Present Discounted Value of Health Net of Medical Costs, by Treatment Status	215
O.1	Combining Functions, Pooled Sample . . . . .	216
O.2	Combining Functions, Male Sample . . . . .	217
O.3	Combining Functions, Female Sample . . . . .	217
O.4	Combining Functions by Category, Pooled Sample . . . . .	218
O.5	Combining Functions by Category   10% Significance, Pooled Sample . . . . .	219
O.6	Combining Functions by Category, Male Sample . . . . .	220
O.7	Combining Functions by Category   10% Significance, Male Sample . . . . .	221
O.8	Combining Functions by Category, Female Sample . . . . .	222
O.9	Combining Functions by Category   10% Significance, Female Sample . . . . .	223

# A Background<sup>1</sup>

## A.1 Overview

The Carolina Abecedarian Project (ABC) and the Carolina Approach to Responsive Education (CARE) were high-quality early childhood education programs each with two phases of randomized controlled design. They were both implemented at the Frank Porter Graham Center (FPGC) of the University of North Carolina in Chapel Hill. ABC served four cohorts of children born between 1972 and 1977, and CARE served two cohorts of children born between 1977 and 1980. In the main paper, we explain the eligibility requirements, the randomization protocol, and the programmatic contents of both programs. In this section of the appendix, we expand on additional important details of all these aspects.

## A.2 Eligibility Criteria and Populations Served

The mothers of the ABC and CARE subjects were typically recruited during the last trimester of pregnancy. Potential families were referred by local social service agencies and hospitals. Eligibility was determined by a score of 11 or more on a weighted 13-factor High-risk Index (HRI).

The HRI for ABC was based on 13 weighted variables, which are listed here with weights in parentheses: (i) maternal education level measured by years of education— 6 (8), 7 (7), 8 (6), 9 (3), 10 (2), 11 (1), 12 (0); (ii) paternal education level with weights identical to those for maternal education; (iii) yearly family income measured in current dollars—\$1,000 or less (8), \$1,001-\$2000 (7), \$2,001-\$3,000 (6), \$3,001-\$4,000 (5), \$4,001-\$5,000 (4), \$5,001-\$6,000 (0); (iv) father’s absence from the household for reasons other than health or death (3); (v) lack of maternal relatives in the area (3); (vi) siblings of school age one or more

---

<sup>1</sup>This section of the appendix is based on joint work with Sylvi Kuperman.

grades behind age-appropriate level, or with equivalently low scores on school-administered achievement tests (3); (vii) received payments from welfare agencies within the past 3 years (3); (viii) record of father’s work indicates instability or unskilled and semi-skilled labor (3); (ix) record of maternal or paternal IQ score of 90 or below (3); (x) record of a sibling with an IQ score of 90 or below (3); (xi) relevant social agencies in the community indicate the family is in need of assistance (3); (xii) one or more family members has sought counseling or professional help in the past 3 years (1); and (xiii) special circumstances not included in any of the above that are likely contributors to cultural or social disadvantage (1).<sup>2</sup> The weighting scale aimed to establish the relative importance of each item in the index.<sup>3</sup> Race was not considered for eligibility; however, 98% of the families who agreed to participate were African-American.<sup>4</sup>

The HRI for CARE was similar to that of ABC—it also contained 13 weighted variables and a score of 11 or above was required to be considered eligible. The items for maternal and paternal education levels have the same categories and weights as the ABC HRI. The other identical items are having an absent father, school-age siblings performing lower than the norm based on grade-level or achievement tests, a record of father’s unstable job history or unskilled labor, social agencies indicating a high level of need, and other circumstances related to cultural or social disadvantage. The specification of the following items were changed between the ABC and CARE HRI. The weight associated with household income depended on the number of individuals in the family for CARE and the income categories range from less than \$2,000 to \$13,500 or more. In the CARE HRI, it is asked if payments were received from welfare agencies in the past 5 years instead of the past 3 years. Similarly, it asks if any family member has sought counseling in the past 5 years instead of the past 3 years. The threshold for maternal or paternal IQ is 85 in the CARE HRI instead of 90 as in

---

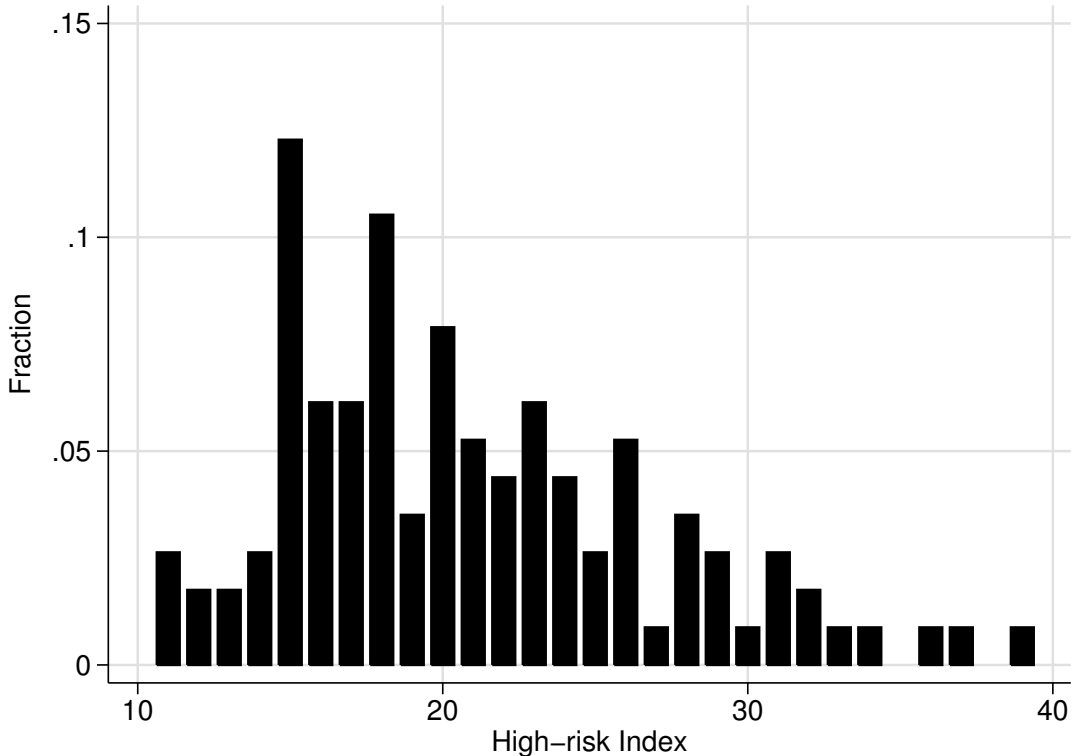
<sup>2</sup>Ramey and Smith (1977); Ramey and Campbell (1984, 1991); Ramey et al. (2000).

<sup>3</sup>Ramey and Smith (1977).

<sup>4</sup>Ramey and Smith (1977); Ramey and Campbell (1979).

the ABC HRI. It does not have an item related to the absence of maternal relatives in the area, but replaces that item with asking if any member of the mother or father’s immediate family has received services for the mentally disabled (the weight for this item is 3).<sup>5</sup>

Figure A.1: High-risk Index Distribution, ABC

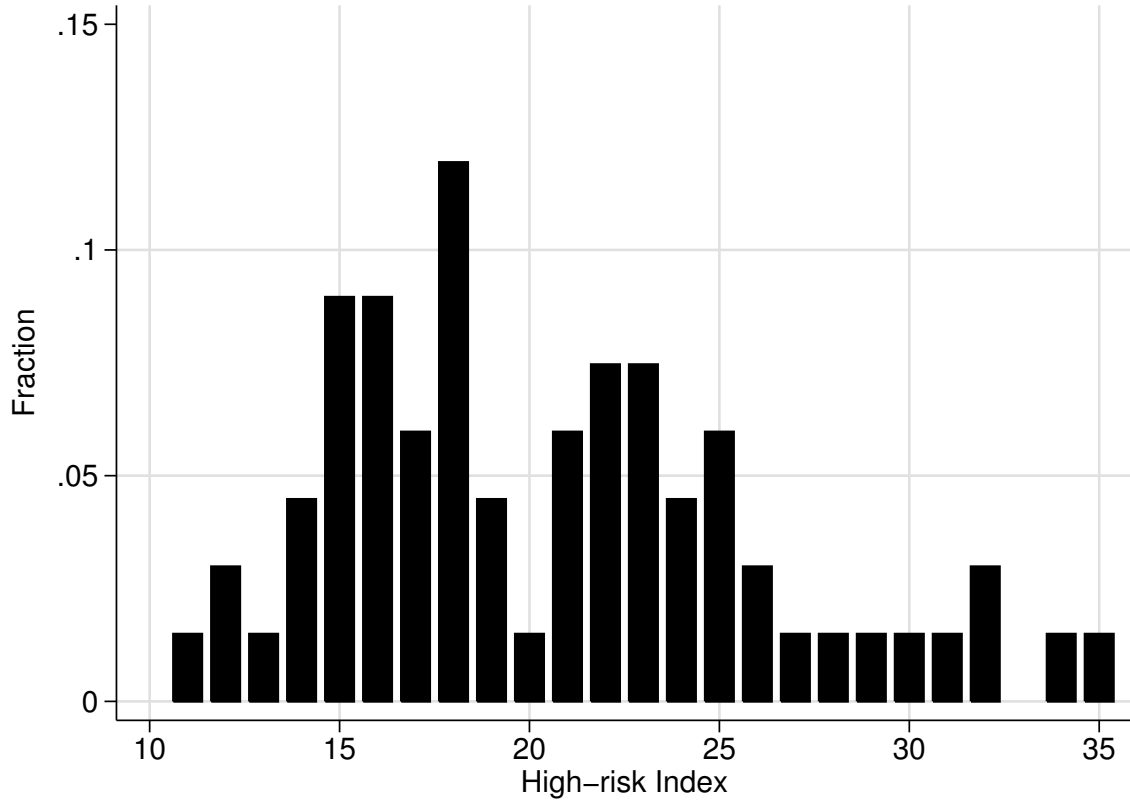


Note: This plot shows the distribution of the High-risk Index (HRI) for ABC, which determined eligibility. Subjects were eligible if they had a score of 11 or more.

---

<sup>5</sup>Ramey et al. (1985).

Figure A.2: High-risk Index Distribution, CARE



Note: This plot shows the distribution of the High-risk Index (HRI) for CARE, which determined eligibility. Subjects were eligible if they had a score of 11 or more.

Figure A.1 and Figure A.2 display the distribution of the HRI score among all subjects in ABC and CARE, respectively. All subjects were substantially disadvantaged. Maternal age when the subject was born was, on average, 19.9 years in ABC and 21.1 years in CARE. Approximately half of the mothers of both treatment-group and control-group subjects in ABC were 19 years or younger and one third were 17 years or younger. In CARE, approximately half of the mothers of both treatment-group and control-group subjects were 20 years or younger and one third were 17.2 years or younger. Mean maternal IQ score in ABC was approximately 85, one standard deviation below the national mean. In CARE, the mean maternal IQ score was approximately 87. Only 25% of the ABC subjects lived with both

biological parents, and more than 50% lived with extended families in multi-generational households (61% of treatment-group subjects and 56% of control-group subjects).<sup>6</sup> About 79% of subjects did not have a father in the home in both ABC and CARE.

### A.3 Randomization Protocol and Compromises

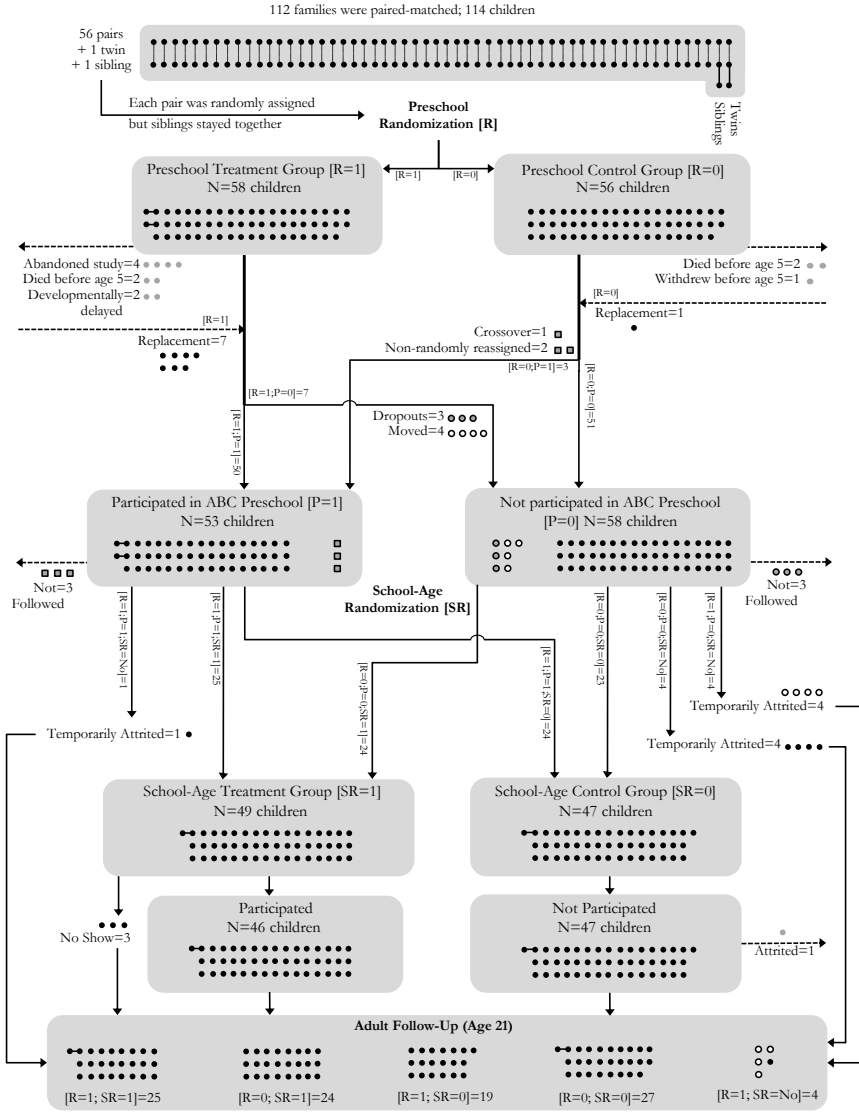
Randomization compromises throughout ABC's and CARE's implementations pose a challenge when evaluating the programs' effects. We discuss each case of compromise in detail. Figure A.3 and Figure A.4 are flow charts that depict the sample from the first-phase randomization through the last data follow-up accounting for all cases of attrition and non-compliance. In Section 3, we propose a methodology for adequately evaluating the programs while accounting for these compromises.

Although most randomization compromises occurred at early stages, this methodology also accounts for the fact that a few subjects were not in the sample either for the second-phase randomization or for the adult follow-ups. In Appendix A.6, we describe the sample reductions that attrition at different stages of the study generates and test potential differences between the subjects who completed data follow-ups and the subjects who did not.

---

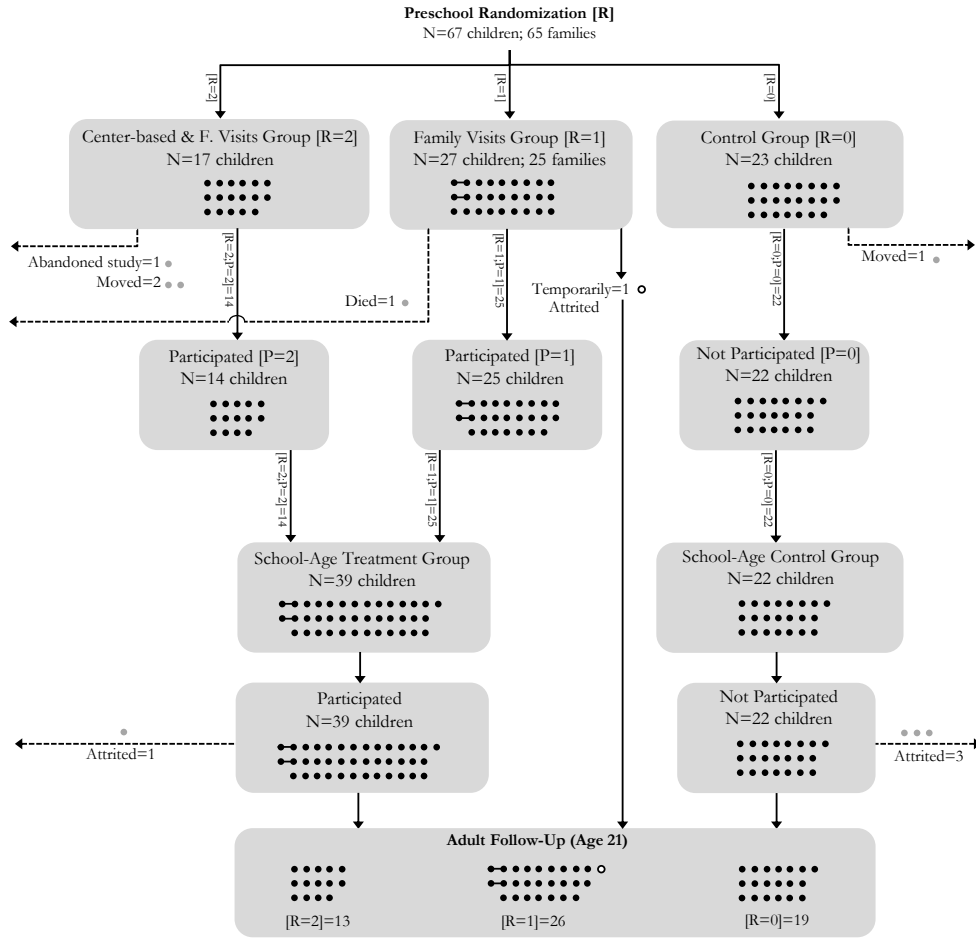
<sup>6</sup>Ramey and Campbell (1991); Campbell and Ramey (1994).

Figure A.3: Randomization Protocol and Treatment Compliance, ABC



Sources: Ramey et al. (1976); Ramey and Smith (1977); Ramey and Campbell (1979, 1984), internal documentation of the program, and own calculations. Note: The variable  $R$  represents randomization into treatment,  $[R = 1]$ , or control,  $[R = 0]$ , groups. After the original randomization, some subjects died or withdrew from the program early in life and were replaced.  $R$  also includes those replacements. Arrows pointing outside of the diagram indicate subjects who left the study permanently. The variable  $P$  represents participation in the preschool-age program. The variable  $SR$  represents randomization into the school-age program,  $[SR = 1]$ , or out of it,  $[SR = 0]$ . Some subjects were not randomized at school age,  $[SR = No]$ . We use the term “temporarily attrited” for subjects who did not participate in the study at school age, but were later interviewed in the age-21 followup.

Figure A.4: Randomization Protocol and Treatment Compliance, CARE



Sources: Wasik et al. (1990), internal documentation of the program, and own calculations. Note: The variable  $R$  represents randomization into center-based childcare and family education,  $[R = 2]$ , family education,  $[R = 1]$ , or control,  $[R = 0]$ . Arrows pointing outside of the diagram indicate subjects who left the study permanently. The variable  $P$  represents participation in the corresponding group of the preschool-age program. The variable  $SR$  represents those who participated in the school-age program,  $[SR = 1]$ , or did not,  $[SR = 0]$ . Unlike in ABC, there was no second-phase randomization in CARE. Rather, those in the center-based childcare and family education group and those in the family education group were automatically assigned to receive the school-age treatment. We use the term “temporarily attrited” for subjects who did not participate in the study at school age, but were later interviewed in the age-21 followup.

### A.3.1 ABC

Both the first and second phases of randomization were conducted at the family level, so pairs of siblings and twins were jointly randomized into either treatment or control groups.<sup>7</sup> Although we know that pairing was based on HRI, maternal IQ, maternal education, maternal age, and gender of the subject, we do not know the original pairs. The study collected an initial sample of 120 families. Twenty-two subjects did not complete the first-phase of treatment as initially assigned by the randomization (see Table A.1).<sup>8</sup>

Of these cases, there were four subjects assigned to treatment who left the study before any data on them was collected. In our main methodology, we assume that they are missing at random. Exercises assuming that these subjects had outcomes identical to the subject with the lowest outcome in the treatment group suggest that, even in this extreme scenario, there is little sensitivity (see Appendix E).

Second, four subjects died before age 5—two of them initially assigned to treatment and two of them initially assigned to control. For all of them, we observe baseline characteristics and any other data collected before their death. For methodological purposes, they represent cases of program attrition when we do not observe their outcomes.

Third, three subjects in the treatment group did not comply to treatment status. They are different from the four subjects who left the study before any data collection because we observe data collected for them from birth to age 8. Afterward, the program staff chose not

---

<sup>7</sup>Sibling pairs occurred when the two siblings were close enough in age such that both of them were eligible for the program.

<sup>8</sup>In Appendix B, we compare the observed baseline characteristics of the subjects in Table A.1 to the observed baseline characteristics of the subjects who complied to the initial treatment assignment. We find little evidence of differences.

to follow them anymore.<sup>9</sup> Therefore, these subjects remain in treatment sample until age 8 or before. After, they represent cases of program attrition, given that we do not observe them anymore. In Appendix E we find little sensitivity to adjusting for non-compliance of these children in the results we show for age 8 or before.

---

<sup>9</sup>Informal conversations with the program's staff do not indicate a clear reason for this.

Table A.1: Randomization Compromises, ABC

Initial Assignment	Compromise Description	Data Availability	Methodology Assumption
Treatment	Left the study	None	Missing at random
Treatment	Left the study	None	Missing at random
Treatment	Left the study	None	Missing at random
Treatment	Left the study	None	Missing at random
Control	Died (age 0), heart disease	Baseline; before death	Attrition after death
Control	Died (age 0), heart disease	Baseline; before death	Attrition after death
Treatment	Died (age 0), SIDS	Baseline; before death	Attrition after death
Treatment	Died (age 4), pedestrian accident	Baseline; before death	Attrition after death
Treatment	Non-compliance	Baseline; before age 8	Attrition after age 8
Treatment	Non-compliance	Baseline; before age 8	Attrition after age 8
Treatment	Non-compliance	Baseline; before age 8	Attrition after age 8
Control	Crossover from control to treatment	Baseline; before age 8	Attrition after age 8
Treatment	3 months of treatment	Baseline; after age 2	Same as treatment group
Treatment	10 months of treatment	Baseline; after age 2	Same as treatment group
Treatment	6 months of treatment	Baseline; after age 2	Same as treatment group
Treatment	9 months of treatment	Baseline; after age 2	Same as treatment group
Control	Left study at 54 months	Baseline; before 54 months	Attrition after 54 months
Treatment	Developmentally delayed at 6 months	No data after diagnosis	Dropped (non-eligible)
Treatment	Developmentally delayed at 36 months	No data after diagnosis	Dropped (non-eligible)
Control	Crossover from control to treatment	Baseline, before age 8	Attrition after age 8
Control	Crossover from control to treatment	Baseline, before age 8	Attrition after age 8

Note: This table describes the various randomization compromises in ABC. For each subject, we display: the nature of the compromise, the available data, and the methodological assumption when accounting for non-compliance and program attrition.

Fourth, one subject initially assigned to control was enrolled into treatment. Her mother wanted to work and the program staff decided to admit her child into center-based care.<sup>10</sup> Both in terms of data collection and in terms of methodological purposes, this subject is analogous to the subjects in the third case.<sup>11</sup>

Fifth, four subjects in the treatment group did not complete treatment in its entirety. They were treated for 3, 10, 6, and 9 months, respectively. Except for follow-ups during childhood, which our main results do not use, we observe most of the data for these subjects. We avoid taking a stance on how beneficial the program was at each age, because we do not have a way to document this. Therefore, we assume that they were treated as other subjects in the treatment group.<sup>12</sup>

Sixth, the family of one subject in the control group moved at age 54 months. We observe any data collection before the family moved, so we consider the subject as part of the control group in any estimation before this event. Afterwards, we do not observe any data on the subject, so we consider her a case of program attrition.

Seventh, two subjects initially assigned to treatment status were diagnosed as developmentally delayed after 6 and 36 months of treatment. No data for them are available after the diagnosis. We drop them from the sample because they were not eligible to be part of the program.

Finally, two subjects initially assigned to the control group were admitted into treatment. Local authorities requested this because the children were considered highly at risk. Data on them are available from birth to age 8 and we treat them as the subjects in the third and

---

<sup>10</sup>Correspondence with the program officers stating this permission is available under request from the authors.

<sup>11</sup>The sensitivity analysis finding little evidence when adjusting for non-compliance includes this case.

<sup>12</sup>If anything, this downward biases the effects of the program we estimate.

fourth cases. They crossed over from the control group to the treatment group, although the reasons were different.

Analysis of each of these cases leads to the following conclusions. For four subjects, we do not have data to methodologically assess them as cases of program attrition, though sensitivity analyses suggest that the treatment effects of the program persist after assigning them the same outcome as the subjects who did the worst in the treatment group. For the subjects who did not comply to treatment, adjusting our estimates for non-compliance when data are available makes little difference. The remaining 14 subjects who did not complete treatment as initially assigned represent various cases of program attrition, for which we propose a correction methodology in Section 3.

To increase the number of subjects in the sample, the program officers recruited additional subjects who were added to the program before the subjects were 6 months old. Our calculations indicate that there were eight replacements. We cannot distinguish in the data the subjects who were initially randomized from the replacement children and there is no documentation on how these subjects were recruited.<sup>13</sup> After the various compromises, the sample consisted of 111 subjects: 53 in the treatment group and 58 in the control group. The observed characteristics for each subjects indicate that they were eligible for the program; all subjects in the sample have an HRI of 11 or above.

Prior to the second phase of randomization, 3 subjects in the first-phase control group and 3 subjects in the first-phase treatment group could not be located for follow-up. One subject in the control group and eight subjects in the treatment group of the first phase did not participate in the second phase but later agreed to participate in the data collections during

---

<sup>13</sup>Three replacements are reported in [Ramey and Campbell \(1979\)](#). Three are documented in correspondence with the program officers, which is available from the authors of the present document upon request. The other two replacements are implied by the number of subjects who participated of the randomization protocol in each cohort.

adulthood. This yielded a sample of 96 subjects in the second phase: 49 in treatment and 47 in control. After the second-phase randomization, three subjects in the treatment group chose not to participate in the program, while all subjects in the control group adhered to their randomization status.

### A.3.2 CARE

The randomization protocol in CARE had no major compromises.<sup>14</sup> Of the 65 initial families, 23 were randomized to a control group, 25 to the family education treatment group, and 17 to the family education and center-based childcare treatment group. Two families in the family education treatment group had twins who were jointly randomized, as in ABC. We document four cases of program attrition (see Table A.2).<sup>15</sup> For methodological purposes, we consider these subjects analogous to their corresponding cases in ABC. We do not present exercises to evaluate the sensitivity to non-compliance because there was none in CARE. Figure A.4 illustrates CARE’s randomization protocol and the presence of subjects throughout the data follow-ups.

---

<sup>14</sup>Wasik et al. (1990); Burchinal et al. (1997).

<sup>15</sup>In Appendix B, we compare the observed baseline characteristics of the subjects in Table A.2 to the observed baseline characteristics of the subjects who complied to the initial treatment assignment. We find little evidence of differences.

Table A.2: Randomization Compromises, CARE

Initial Assignment	Compromise Description	Data Availability	Methodology Assumption
Family education	Died (age 0), unknown causes	Baseline	Attrition after dead
Center-based Childcare and Family Education	Left study at age 5	Baseline; before age 5	Attrition after age 5
Control	Move at 11 months old	Baseline; before 11 months	Attrition after 11 months
Center-based Childcare and Family Education	Move at 5 months old	Baseline; before 5 months	Attrition after 5 months
Center-based Childcare and Family Education	Move at age 5	Baseline; before age 5	Attrition after 5

Note: This table describes the various randomization compromises in CARE. For each subjects, we display: the nature of the compromise, the available data, and the methodological assumption when accounting for non-compliance and program attrition.

## A.4 Program Description and Content

### A.4.1 Goals

The original goals of treatment were to prevent mental retardation<sup>16</sup> by enhancing overall development from birth, in turn fostering school-readiness for an at-risk population. Additional curriculum goals were to (i) support language, motor, and cognitive development; (ii) minimize high-risk behaviors; and (iii) develop socio-emotional competencies considered crucial for school success including task-orientation, communicative competence, independence, and prosocial behavior.<sup>17</sup> Implementation of ABC's and CARE's educational treatments evolved each successive year as program staff evaluated ongoing outcome data.<sup>18</sup>

### A.4.2 Daily Schedule

For both ABC and CARE, FPGC was open to families from 7:45 a.m. to 5:30 p.m., 5 days per week and 50 weeks per year.<sup>19</sup> Subjects were offered free transportation to and from the center. A driver and second adult staffed each vehicle (one van and two station wagons) equipped with child safety seats.<sup>20</sup> Data provided by FPGC indicate that approximately 65% of treated ABC families utilized the free transportation.<sup>21</sup> Vehicles typically arrived by 9:00 a.m. to the center and departed around 3:45 p.m.<sup>22</sup> At FPGC, ABC and CARE treatment-group subjects received breakfast, lunch, and a snack planned by a nutritionist.<sup>23</sup> Meals were catered by off-site kitchens. Infants received iron-fortified formula until doctors advised adding solid food. The control-group subjects also received an unlimited amount of

---

<sup>16</sup>Note that the clinical understanding of mental retardation was once associated with disadvantages that hindered early-life development (Noll and Trent, 2004).

<sup>17</sup>Ramey et al. (1976, 1985); Sparling (1974); Wasik et al. (1990); Ramey et al. (2012).

<sup>18</sup>Ramey et al. (1975); Finkelstein (1982); McGinness (1982); Haskins (1985).

<sup>19</sup>Ramey et al. (1976, 1985).

<sup>20</sup>Ramey and Campbell (1979); Kuperman (2015).

<sup>21</sup>Barnett and Masse (2002).

<sup>22</sup>Ramey et al. (1977).

<sup>23</sup>Haskins (1985); Bryant et al. (1987); Ramey et al. (1977).

iron-fortified formula until approximately 15 months of age.<sup>24</sup>

### A.4.3 Program Staff and Physical Space

To promote trust in FPGC within the sample families, staff were recruited from the local community.<sup>25</sup> Infant and toddler caregivers and preschool teachers demonstrated varied educational backgrounds ranging from high school graduation to master's degrees. Their average professional working experience with young children was 7 years.<sup>26</sup> All classroom staff participated in extensive training and were closely observed by FPGC's academic staff, as part of a broad variety of ongoing clinical and social research related to early childhood education, psychology, and health. In ABC, child-caregiver ratios varied by age: 3:1 for infants up to 13 to 15 months of age; 4:1 for toddlers up to 36 months; and 5:1 or 6:1 for children aged 3 to 5 years, depending on cohort size.<sup>27</sup> Child-caregiver ratios were similar in CARE.<sup>28</sup>

The ABC and CARE staff included a program director, a secretary, 12 to 14 teachers and assistant teachers, 3 administrative staff members, and a transportation supervisor.<sup>29</sup> Lead caregivers and teachers had bachelor's or master's degrees. Teacher aides, recruited from the local community, held high school diplomas (at minimum) and were comparatively well-compensated in the childcare field. They remained a stable treatment component throughout the study. After 1980, following revisions to FIDCR regarding minimum requirements for early childhood education staff, several teacher aides pursued and received undergraduate degrees and became lead teachers. All classroom staff were supervised daily, received weekly mentoring, and professional development from outside consultants..<sup>30</sup>

---

<sup>24</sup>Campbell et al. (2014); Kuperman (2015).

<sup>25</sup>Ramey et al. (1977); Bryant et al. (1987); Feagans (1996); Kuperman (2015).

<sup>26</sup>Ramey et al. (1982, 1985); Wasik et al. (1990).

<sup>27</sup>Ramey et al. (1977); Ramey and Campbell (1979); Ramey et al. (1982).

<sup>28</sup>Burchinal et al. (1997); Ramey et al. (1985).

<sup>29</sup>Ramey et al. (1977, 1982); Bryant et al. (1987).

<sup>30</sup>O'Brien and Sanders (1974); Ramey et al. (1985); Sanders and Stokes (1979); Klein and Sanders (1982); Kuperman (2015).

Infant nurseries, toddler rooms, and preschool classrooms were housed on different floors of FPGC. Early reports indicate that FPGC allocated two floors to ABC, but later reports indicate the use of three floors.<sup>31</sup> Two infant nurseries were staffed by five adults in a suite of four adjoining rooms: two sleeping rooms contained seven cribs each, while the other two rooms were designated for activities.<sup>32</sup> The four rooms opened into a large, shared space with feeding tables, an area for food preparation, and a couch.<sup>33</sup> Offices for the medical staff, along with two examining rooms and facilities for laboratory tests were located around the corner from the infant nurseries.<sup>34</sup> Two multi-age toddler rooms were located one floor below the infant nurseries. One room served children who were 1 to 2 years old and the other served children 2 to 3 years old.<sup>35</sup> 3-year-olds were housed in a closed classroom near the toddler rooms. On the lowest floor, 4-year-olds shared an open classroom with a public kindergarten program; the two classes were separated by a long, low bookcase. In CARE, two floors of FPGC were allocated to nurseries and classrooms. A mixed-age classroom design was implemented combining children ages 1 and 3, and children ages 2 and 4. Teacher-child ratios for these ages remained 1:5. FPGC offered two outdoor play areas for both ABC and CARE: one for children up to age 3, and the other for older children.<sup>36</sup>

#### A.4.4 Approach to Child Development

Curriculum delivery enabled a highly customized learning experience for treated subjects in both ABC and CARE. Infant caregivers recorded child observations on progress charts and collaborated with FPGC's curriculum developers and academic researchers to rotate

---

<sup>31</sup>Ramey and Smith (1977); Ramey and Campbell (1979); Ramey and Haskins (1981).

<sup>32</sup>Ramey et al. (1977).

<sup>33</sup>Ramey and Campbell (1979).

<sup>34</sup>Kuperman (2015).

<sup>35</sup>Ramey and Smith (1977); Ramey and Campbell (1979).

<sup>36</sup>Ramey and Campbell (1979); Ramey et al. (1982).

learning activities every 2 to 3 weeks for each treated subject.<sup>37</sup> Preschool rooms featured intentionally organized environments to promote pre-literacy and access to a rich set of learning tools. The full-day curriculum emphasized active learning experiences, dramatic play, and pre-academics. Frequent 1:1 or 2:1 child-adult interactions prioritized language development for social competence. For ages 3 through 5, as the cohorts approached public school entry, classroom experiences were increasingly structured towards the development of pre-academic skills and “socio-linguistic and communicative competence.”<sup>38</sup> FPGC offered a summer program before the start of kindergarten designed to target specific skills to ensure success in a kindergarten classroom (e.g., lining up when exiting the classroom). This program was available to subjects in both the center-based childcare and family education group and the family education group.<sup>39</sup>

ABC’s and CARE’s learning programs were influenced by key developmental theorists, including Bowlby, Piaget, Vygotsky, and Tough.<sup>40</sup> All four ABC cohorts and two CARE cohorts participated in curriculum developers Sparling and Lewis’ “Learninggames for the First Three Years.”<sup>41</sup> The “Learninggames” were implemented daily by infant and toddler caregivers in 1:1 child-adult interactions. Each “Learninggames” activity stated a developmentally-appropriate objective, the necessary materials, directions for teacher behavior, and expected child outcome. The activities were designed for use both indoors and outdoors, while dressing, eating, bathing, or during play.<sup>42</sup>

Supplemental curricula for preschool rooms varied throughout the study, and included “Cook and Learn,” “Peabody Early Experiences Kit,” “GOAL Math Program,” and “My Friends

---

<sup>37</sup>Ramey et al. (1976); Campbell and Ramey (1994).

<sup>38</sup>Ramey et al. (1977); Haskins (1985); Ramey and Haskins (1981); Ramey and Campbell (1979); Ramey and Smith (1977); Ramey et al. (1982); Sparling and Lewis (1979, 1984).

<sup>39</sup>Ramey et al. (1985).

<sup>40</sup>Sparling (1974); McGinness and Ramey (1981); Kuperman (2015).

<sup>41</sup>Sparling and Lewis (1979).

<sup>42</sup>Ramey and Campbell (1979); Ramey and Haskins (1981); Sparling and Lewis (1979).

and Me.”<sup>43</sup>

CARE subjects randomized into the center-based childcare and family education group or the family education group also received home visits designed to transmit information on child development and skills involved with parenting including strategies for parent-child interactions based on “Learninggames” activities and problem-solving techniques.<sup>44</sup> Home visitors were trained to ensure they were able to form a strong relationship with the parent and successfully implement the curriculum.<sup>45</sup> The visits lasted about an hour, and occurred weekly until the child was 3 years old. After age 3, the home visits were less frequent and depended on the preferences of the parents. They were usually about once a month after age 3.<sup>46</sup>

#### A.4.5 Medical Care and Nutrition

ABC and CARE provided comprehensive on-site medical care because it was conducted in conjunction with a longitudinal medical research study on infectious respiratory diseases in group environments.<sup>47</sup> Treatment group children were monitored daily for signs of illness. All treated children received medical care while attending center-based childcare; the first ABC cohort of control-group children also received medical care during the program’s first year of implementation.<sup>48,49</sup>

In ABC, primary pediatric care was provided by a family nurse practitioner and a licensed practical nurse, both under the supervision of one pediatrician who was on continuous duty

---

<sup>43</sup>Greenberg and Epstein (1973); Karnes (1973); Dunn et al. (1976); Davis (1977); Wallach and Wallach (1976).

<sup>44</sup>Bryant et al. (1987); Wasik et al. (1990); Burchinal et al. (1997).

<sup>45</sup>Bryant et al. (1987).

<sup>46</sup>Bryant et al. (1987); Wasik et al. (1990); Burchinal et al. (1997).

<sup>47</sup>Henderson et al. (1982).

<sup>48</sup>Ramey et al. (1976); Bryant et al. (1987); Ramey and Campbell (1991); Campbell and Ramey (1994).

<sup>49</sup>Subjects in both the treatment and control groups of the first cohort received free medical care provided by ABC. The control group of the first cohort only received medical care in the first year of the program; the treatment group of the first cohort received medical care for all years of the program. In the subsequent cohorts, only subjects in the treatment group received free medical care provided by ABC. Both CARE cohorts of treated subjects received medical care.

at the center.<sup>50</sup> In CARE, the medical staff included two pediatricians, a family nurse practitioner, and a licensed practical nurse.<sup>51</sup> The medical staff provided regularly scheduled check-ups, immunizations, parental counseling, and initial assessment of illnesses.<sup>52</sup> The treatment group received standard check-ups when they were 2, 4, 6, 9, 12, 18, and 24 months old and annually thereafter. While in treatment, they also received the standard immunizations.<sup>53</sup> In ABC, a licensed practical nurse visited classrooms for up to two hours on a daily basis to monitor the subjects' health status.<sup>54</sup> Although this medical care was offered to the treatment-group families free of charge, it was the policy of the medical staff to refer families to a community hospital for serious treatment. While ABC and CARE provided aspirin, immunizations, and basic medicines, families were responsible for purchasing any prescription medication subjects required. There are no data currently available on treatment received for serious conditions or use of prescription medication.

Infants were supplied with iron-fortified formula. Children older than 15 months of age were provided breakfast, lunch, and an afternoon snack all planned by a nutritionist.<sup>55</sup> Control families received diapers for up to three years and unlimited iron-fortified bottled formula through 15 months.<sup>56</sup>

#### A.4.6 School-age Treatment

The ABC subjects participated in a second-phase randomization into a school-age treatment (95 subjects continued to this stage of treatment). The CARE subjects in the center-based childcare and family education group and the family education group received the school-age treatment without randomization. The school-age treatment lasted for the first three years

---

<sup>50</sup>Haskins et al. (1978).

<sup>51</sup>Bryant et al. (1987).

<sup>52</sup>Ramey et al. (1977); Bryant et al. (1987).

<sup>53</sup>Bryant et al. (1987); Campbell et al. (2014).

<sup>54</sup>Sanyal et al. (1980).

<sup>55</sup>Bryant et al. (1987); Campbell et al. (2014); Kuperman (2015).

<sup>56</sup>Ramey et al. (1976); Ramey and Campbell (1979); Ramey et al. (1985).

of elementary school and consisted of home visits conducted by a Home/School Resource Teacher.<sup>57</sup> These visits were structured to increase exposure to reading and mathematics and promote parental involvement in the academic process.

The curriculum was delivered through sets of activities (around 60 per year) that developed skills such as handwriting, phonics, and math facts.<sup>58</sup> Teachers worked to encourage parental involvement in the subjects' academics and provided incentives to families to comply with the treatment, such as giving gift certificates to restaurants and books for the subjects upon the completion of activity packets. Home activities were designed at the appropriate level to promote success.

Teachers had graduate-level education, training in special education, *or* were qualified to act as consultants for in-school teachers to address any problems that arose.<sup>59</sup> They met with parents at home and with teachers in the schools to deliver new activities for them to complete with their children and discuss the child's level of success with the previous set of activities. In addition, they helped parents with issues such as adult literacy, housing, and medical care. Thus, the teacher had a dual role as a parent educator and an advocate for the subject in their educational institution.

## A.5 Control Substitution

In ABC, the families of almost 70% of the control-group subjects enrolled their children in high-quality childcare. In CARE, 74% of families in the control group and 62% of families in the family education group enrolled their children in high-quality childcare. We refer to this phenomenon as control substitution; accounting for it is fundamental when evaluating the program, as we argue in Section 3. In this Appendix, we thoroughly describe the character-

---

<sup>57</sup>Burchinal et al. (1997).

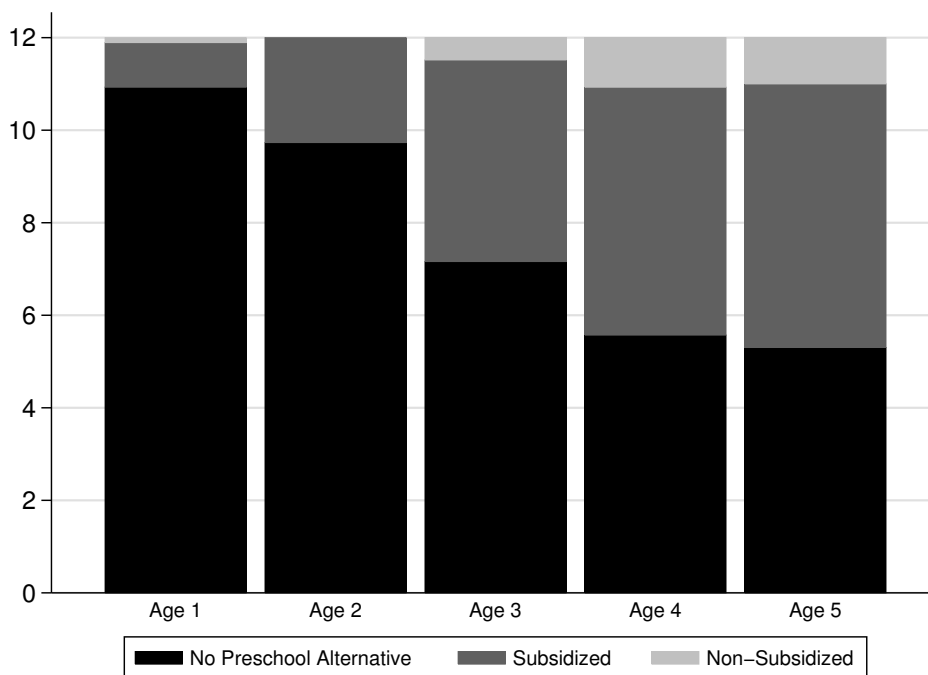
<sup>58</sup>Campbell and Ramey (1989).

<sup>59</sup>Ramey and Campbell (1991).

istics and costs of the childcare centers providing alternative treatment, in order to create a comparison with the treatments offered by ABC and CARE.

Most of the families in the ABC and CARE control groups enrolled their children in alternative preschool that received federal subsidies and, therefore, were regulated by the Federal Interagency of Day Care Requirements (FIDCR). Figure A.5 and Figure A.6 show the amount of enrollment into subsidized and non-subsidized care for ABC and CARE, respectively. Subsidized centers were required to have trained staff who were able to implement curricula designed to enhance cognitive, social, and linguistic competence in disadvantaged children.<sup>60</sup> Thus, we consider these centers to offer high-quality center-based childcare.

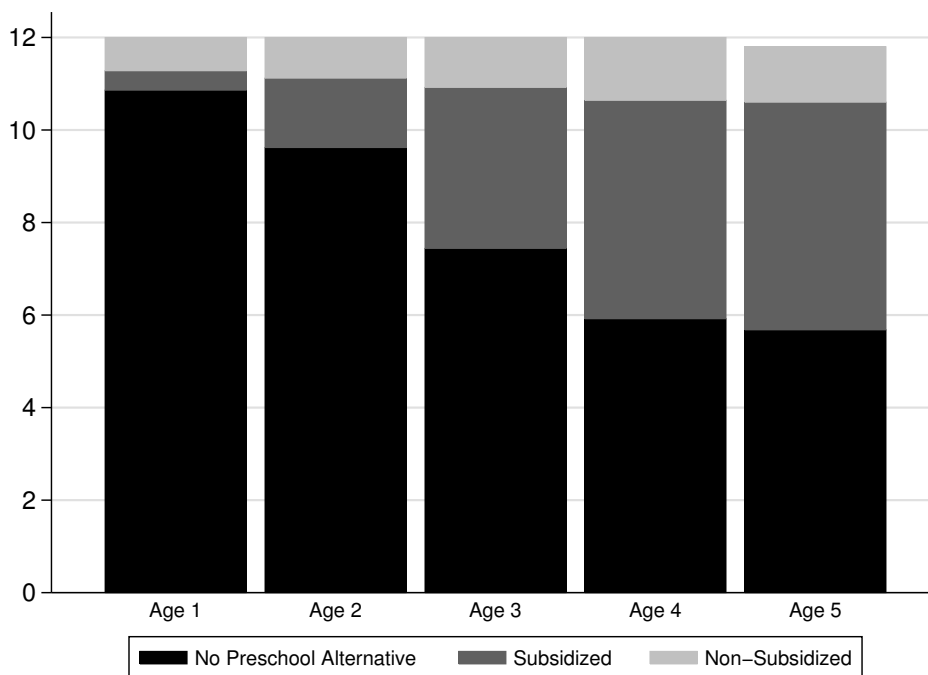
Figure A.5: Average Number of Months in Alternative Preschool, ABC Control Group



Note: This figure describes the take-up of alternative preschool by families in the ABC control group. The vertical axis represents the average number of months per year the subjects of the control group spent in alternative preschool. Subsidized centers were highly regulated and, therefore, relatively high-quality. Non-subsidized childcare services were center-based but not regulated. Other sources of childcare could have included care by parents, relatives, or non-relatives.

<sup>60</sup>Burchinal et al. (1989).

Figure A.6: Average Number of Months in Alternative Preschool, CARE Control and Family Education Groups



Note: This figure describes the take-up of alternative preschool by families in the CARE family education and control groups. The vertical axis represents the average number of months per year the subjects of the control group spent in alternative preschool. Subsidized centers were highly regulated and, therefore, relatively high-quality. Non-subsidized childcare services were center-based but not regulated. Other sources of childcare could have included care by parents, relatives, or non-relatives.

### A.5.1 Regulation

During the ABC and CARE programs, North Carolina had an active, high-quality system of public childcare for vulnerable families funded by several public programs. Examples include Title IV-A of the Social Service Administration (SSA), Aid to Families with Dependent Children (AFDC), and Title IV-B of Child Welfare Services. These funding efforts were amplified in 1975 by Title XX of the SSA, Social Services Block Grant, which was the main federal source of childcare financing in the US when ABC was active.<sup>61</sup>

Federally funded childcare services were regulated according to FIDCR standards, which

<sup>61</sup>Robins (1988).

defined stringent regulation for center-based programs for children between the ages of 3 and 6.<sup>62</sup> These requirements were enforced and childcare providers were aware of them.<sup>63</sup> Additionally, North Carolina had a mandatory licensing law for childcare facilities. This regulation only applied to centers serving children below the age of 3 while FIDCR applied to centers for older children (between the ages of 3 and 6). The relative weakness of this regulation is not very relevant for our study because treatment substitution occurred mostly after age 3.<sup>64</sup> Table A.3 compares a widely-used quality standard, the child-staff ratio, between the North Carolina and FIDCR standards and the actual ABC and CARE numbers.

Table A.3: Child-Staff Ratios for North Carolina, FIDCR, and Actual ABC and CARE Ratios

Age	NC Standards Level I	FIDCR Standards	ABC and CARE Ratios
0–1	6:1*		3:1
1–2	8:1*		4-5:1
2–3	12:1*		4-5:1
3–4	15:1	5:1*	4-5:1
4–5	20:1	7:1*	5-6:1
5–6	25:1	7:1*	5-6:1

Sources: Department of Health, Education, and Welfare (1968); North Carolina General Assembly (1971); Ramey et al. (1977); Ramey and Campbell (1979); Ramey et al. (1982); Burchinal et al. (1997).

Note: The starred ratios represent the ones we believe were the most relevant for ABC control-group subjects and CARE control-group and family-education-group subjects.

### A.5.2 Costs

Previous papers have used childcare cost rates that are not specific to North Carolina and do not account for the contemporaneous structure of the subsidies. We use the local subsidy rates that were in place when the ABC subjects were in preschool to impute different costs of the alternative preschools. These costs depend on the specific preschool attended and the

<sup>62</sup>Department of Health, Education, and Welfare (1968).

<sup>63</sup>Kuperman and Hojman (2015b).

<sup>64</sup>North Carolina General Assembly (1971).

eligibility of the families to receive the subsidies.

When ABC was in operation, center-based childcare was subsidized by several federal programs (the Department of Social Services categorized these programs as Child Welfare, AFDC, and Work Incentive Programs).<sup>65</sup> However, our calculations of the cost of alternative preschool is simplified by the fact that the subsidies were centralized and regulated by the County Department of Social Services. Those departments used a uniform subsidy rate, regardless of the origin of the funds.<sup>66</sup> We collected information about the subsidy rate at the time, which approximates the price of the centers, as centers pegged their fees and services to the maximum subsidy rate. Moreover, we know which centers each ABC control subject attended. We interviewed North Carolina childcare staff and academics that study childcare to document which of those centers were subsidized and regulated at the time.<sup>67</sup> For subsidized centers, we impute the maximum Department of Social Services fee established at the time: \$633/month in 2014 USD.<sup>68</sup> For non-subsidized centers, we impute the mean of costs for Level-1 centers (minimum accepted quality level) according to a 1982 North Carolina study of the cost of childcare: \$298/month in 2014 USD.<sup>69</sup> Although the information in this survey is not ideal to assess the cost of subsidized preschools, as the subsidies greatly changed after the end of FIDCR (1981), it provides an approximation for assessing the cost of the non-subsidized centers.

Finally, we determine if the families paid the costs themselves or if they were subsidized, in which case we also add deadweight costs. We consider if a subject was eligible for subsidies if the family lived in poverty according to the federal guidelines and all parents living at home worked. If a family is deemed eligible, then we assume the child's preschool was fully

---

<sup>65</sup>North Carolina State Department of Social Services (1972).

<sup>66</sup>Ad Hoc Committee of Professionals in Child Care Services, North Carolina (1974).

<sup>67</sup>Kuperman and Hojman (2015b,a).

<sup>68</sup>Ad Hoc Committee of Professionals in Child Care Services, North Carolina (1974); Community Planning Services (1973).

<sup>69</sup>Administrative Branch, Office of Day Care Services (1982).

subsidized using the rates described above without additional subsidies.

## **A.6 Data**

In Table A.4 through Table A.9, we summarize the data availability for both ABC and CARE. The data collection processes in both programs were analogous by design. For both programs, the treatment and control groups were followed into adulthood with relatively low attrition. For ABC, subjects were followed annually through elementary school and at ages 12, 15, 21, and 30. Health and administrative crime data were collected when the subjects reached their mid-30s. For CARE, the exact same follow-ups are available with the exception of the age 15 follow-up.

Table A.4: Early Childhood Data (Part I)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure	
Demographics	Gender	Gender of subject	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54	Demographic Interview	
	Race	Race/Cultural identity of subject	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54	Demographic Interview	
	Birth Date	Date of birth of subject	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54	Demographic Interview	
Cognitive Assessments	Language Ability	Auditory association, Verbal expression, etc.	36, 42, 48, 54	30, 42, 54	ITPA <sup>ABC</sup> , GPB <sup>ABC</sup> , PLP <sup>ABC</sup> , MSCD	
	Intelligence Levels	SBIS	24, 36, 48, 60	24, 36, 48, 60	SBIS	
		WPPSI	60	60	WPPSI	
		BSID	3, 6, 9, 12, 18, 24	6, 12, 18, 24	BSID	
		UOSPD	15	-	UOSPD <sup>ABC</sup>	
	Memory	RPM	60	-	RPM <sup>ABC</sup>	
		Quantitative	BSID	3, 6, 9, 12, 18, 24	6, 12, 18, 24	BSID
	Motor Development	MSCD	MSCD	30, 42, 54	30, 42, 54	MSCD
		Memory	BSID	3, 6, 9, 12, 18, 24	6, 12, 18, 24	BSID
	Critical Thinking	MSCD	MSCD	30, 42, 54	30, 42, 54	MSCD
Curiosity		Curiosity	30, 36, 42, 48, 54, 60, 66, 72	-	Infant Behavior Inventory <sup>ABC</sup>	
Non-Cognitive Assessments	Social Skills	Positive social response	30, 36, 42, 48, 54, 60, 66, 72	6, 12, 18, 24	Infant Behavior Inventory <sup>ABC</sup> , Bayley Infant Inventory <sup>CARE</sup>	
		Creativity	30, 36, 42, 48, 54, 60, 66, 72	-	Infant Behavior Inventory <sup>ABC</sup>	
	Self-Control	Locus of control	3, 18	6, 18	RIES	
Emotional Health	KRT	Distractibility, Attentiveness	30, 36, 42, 48, 54, 60, 66, 72	6, 12, 18, 24	Infant Behavior Inventory <sup>ABC</sup> , Bayley Infant Inventory <sup>CARE</sup>	
		KRT	24, 36, 48, 60	24, 30, 36, 42, 48, 60	KRT	
Self-Consciousness	Self-consciousness	Self-consciousness	30, 36, 42, 48, 54, 60, 66, 72	-	Infant Behavior Inventory <sup>ABC</sup>	

Sources: Authors' description.  
 Note: This table describes the main categories of variables that were measured for ABC and CARE subjects up to age 6. ABC and CARE ages are measured in months. This is not an exhaustive list of variables, nor does it include variables from auxiliary data. Instruments or questionnaires available for only one of the studies are indicated with the superscript *ABC* or *CARE*. **Abbreviations are as follows.** ITPA: Illinois Test of Psycholinguistic Ability. GPB: Gordon Psycholinguistic Battery. PLP: Preschool Language Performance. MSCD: McCarthy Scales of Children's Development. BSID: Bayley Scales of Infant Development and Infant Behavior. UOSPD: Uzgiris-Hunt Ordinal Scales of Psychological Development. RPM: Raven's Progressive Matrices. RIES: Rotter's Internality-Externality Scale. KRT: Kohn and Rosman Test Behavior Inventory.

Table A.5: Early Childhood Data (Part II)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Family Environment	Family Members	Number of primary caretakers	Birth, 18, 30, 42, 54, 60	18, 30, 42, 54, 60	Demographic Interview
		Relationship with family members, including father, mother, siblings, etc.	Birth, 18, 30, 42, 54	18, 30, 42, 54, 60	Demographic Interview
		Number of siblings	Birth, 18, 30, 42, 54, 60	Birth, 18, 30, 42, 54, 60	Demographic Interview
		Marital status of parents	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54, 60	Demographic Interview
		Marital conflicts between parents	6, 18	Birth, 6, 18, 36	Demographic Interview <i>CARE</i> , Parental Attitudes Research Inventory
		Father at home	18, 30, 42, 54	18, 30, 42, 54, 60	Demographic Interview
		Family Economic Environment	Parents' occupation	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54, 60
	Mother works		18, 30, 42, 54	18, 30, 42, 54, 60	Demographic Interview
	Source of child support		Birth, 18, 30, 42, 54	18, 30, 42, 54, 60	Demographic Interview
	Parents and Home Environment	Family income	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54, 60	Demographic Interview
Parents' authority, warmth, family conflict, etc.		6, 18, 30, 42, 54	6, 12, 18, 30, 42, 54	Parent Interview	
Family Social Status	Parents' education background	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54, 60	Demographic Interview	
	Risk taking of family members	Birth	-	Parent Interview <i>ABC</i>	
Family Members' Physical Health	Health issues of parents	Birth	Birth	Parent Interview	
	Pregnancy history	Birth	Birth	Parent Interview	
Childcare	Day-care Experience	Time and location of childcare, Age when begun	Birth, 18, 30, 42, 54	18, 30, 42, 54	Demographic Interview
		Home visits	-	6, 18, 30, 42, 54, 60	Home Visit Data <i>CARE</i>
	Parental Care	Maternal warmth, Maternal involvement with child	6, 18, 30, 42, 54	6, 12, 18, 30, 42, 54	Home Stimulation
		Provision of appropriate play materials	6, 18, 30, 42, 54	6, 12, 18, 30, 42, 54	Home Stimulation
		Avoidance of restriction and punishment	6, 18, 30, 42, 54	6, 12, 18, 30	Home Stimulation
		Authoritarian control	6, 18, 30, 42, 54	6, 12, 18, 30, 36, 42, 102	Home Stimulation, Parental Attitudes Research Inventory
		Democratic attitudes	6, 18	6, 18, 36	Parental Attitudes Research Inventory
		Hostility and rejection	6, 18	6, 18, 36	Parental Attitudes Research Inventory
	Physical Health	Parents' knowledge of childcare	Birth	-	Parent Interview <i>ABC</i>
		Growth Data	Height, Weight, Head circumference, etc.	3, 6, 9, 12, 18, 24, 36, 48, 60	Birth, 6, 12, 18, 24, 36, 48, 60

Sources: Authors' description.

Note: This table describes the main categories of variables that were measured for ABC and CARE subjects up to age 6. ABC and CARE ages are measured in months. This is not an exhaustive list of variables, nor does it include variables from auxiliary data. Instruments or questionnaires available for only one of the studies are indicated with the superscript *ABC* or *CARE*.

Table A.6: Childhood and Adolescence Data (Part I)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure	
Cognitive Assessment	Language Ability	Adaptive Language Inventory	6, 7, 8	6, 7, 8	Adaptive Language Inventory	
		Language Questionnaire	12	-	Language Questionnaire <sup>ABC</sup>	
	Intelligence Tests	MSCD	MSCD	7	-	MSCD <sup>ABC</sup>
			SBIS	6	7	SBIS
		WIS	6, 7, 8, 12, 15	6, 8	WIS	
		Kaufman <sup>CARE</sup>	-	6	Kaufman <sup>CARE</sup>	
		Quantitative Skills	MSCD <sup>ABC</sup>	7	-	MSCD <sup>ABC</sup>
			MSCD <sup>ABC</sup>	7	-	MSCD <sup>ABC</sup>
		Memory	MSCD <sup>ABC</sup>	7	-	MSCD <sup>ABC</sup>
			MSCD <sup>ABC</sup>	7	-	MSCD <sup>ABC</sup>
Non-Cognitive Assessment	Interpersonal Skills	Gets along with people	6, 8, 12, 15	8, 12	PEI, CAS, PMI <sup>ABC</sup> , SAI <sup>ABC</sup> , Subject Interview <sup>ABC</sup> , Quality Rank <sup>CARE</sup>	
		Relationship with the other sex	15	-	SAI <sup>ABC</sup> , Subject What I Am Like (Harter) <sup>ABC</sup>	
	Critical Thinking	Thinks for self, questions things	6, 8	8, 12	PEI, Harter Child <sup>CARE</sup> , CBI	
		Concept Attainment Kit	6, 7, 8	-	Concept Attainment Kit <sup>ABC</sup>	
	Self-Control	Distracted in class	6, 7, 8, 12, 15	12	SCAN <sup>ABC</sup> , CBI, WPB <sup>ABC</sup> , PMI <sup>ABC</sup> , SAI <sup>ABC</sup> , Self-Evaluation Inventory <sup>ABC</sup>	
		Locus of control	15	-	Nowicki-Strickland Data, Pearlin Mastery Scale <sup>ABC</sup>	
	Work Ethic	Task Orientation	6, 7, 8, 12, 15	6, 7, 8, 9, 12	SCAN <sup>ABC</sup> , CBI, PMI <sup>ABC</sup>	
		Harms self, suicidal thoughts	8, 12, 15	8, 12	Achenbach Parent, Subject Risk Taking Survey <sup>ABC</sup>	
	Emotional Health	Depression, anxiety, fear, etc.	6, 7, 8, 12, 15	7, 8, 9, 12	KRT, CAS, ETS, Achenbach Parent	
		Athletic activities	8, 12, 15	8, 12	Achenbach Parent, SAI <sup>ABC</sup> , Subject What I Am Like (Harter) <sup>ABC</sup> , PEI <sup>CARE</sup>	
Social Activities	Participant of organizations, e.g. religions	8, 12, 15	8, 12	Achenbach Parent, SAI <sup>ABC</sup> , Subject Interview <sup>ABC</sup>		
	Reading list	12, 15	12	CAS, SAI <sup>ABC</sup>		
Self-Consciousness	TV/music	12, 15	12	CAS, SAI <sup>ABC</sup> , Television Checklist <sup>ABC</sup>		
	Self-conscious emotions	8, 12, 15	8, 12	Achenbach Parent, Subject What I Am Like (Harter)		

Sources: Authors' descriptions.

Note: This table describes the main categories of variables that were measured for ABC and CARE subjects at ages 6 to 18. ABC and CARE ages are measured in years. This is not an exhaustive list of variables, nor does it include variables from auxiliary data. Instruments or questionnaires available for only one of the studies are indicated with the superscript <sup>ABC</sup> or <sup>CARE</sup>. **Abbreviations are as follows.** MSCD: McCarthy Scales of Children's Development. SBIS: Stanford-Binet Intelligence Scale. WIS: Wechsler Intelligence Scale for Children. KRT: Kohn and Rosman Test Behavior Inventory. WJCA: Woodcock-Johnson Test of Cognitive Abilities. PEI: Parents as Educator Interview. CAS: Child Assessment Schedule. PMI: Psychosocial Maturity Inventory. SAI: Social Adjustment Inventory for Children and Adolescents. SCAN: Schedule of Classroom Activity Norms. CBI: Classroom Behavior Inventory. WPB: Walker Problem Behavior Checklist. ETS: Emotional/Activity/Sociability/Impulsivity Temperament Survey. FES: Family Environment Scale. PIAT: Peabody Individual Achievement Test. CAT: California Achievement Test. MARS: Mid-Adolescence Rating Scale Data.

Table A.7: Childhood and Adolescence Data (Part II)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure	
Family Environment	Family Members	Number of adults in house	6, 8, 12, 15	8, 12	PEI, Parent Interview, Subject Person In Household <sup>ABC</sup>	
		Relationship with family members, including father, mother, siblings, etc.	6, 8, 12, 15	8, 12	PEI, FES, SAI, Subject Interview <sup>ABC</sup> , Adult, Self Report <sup>ABC</sup> , Parent Interview, Achenbach Parent PEI <sup>ABC</sup> , Parent Interview	
		Number of siblings	6, 8, 12, 15	7, 8, 12	PEI <sup>ABC</sup> , Parent Interview	
		Marital status of parents	6, 8, 12, 15	7, 8, 12	PEI <sup>ABC</sup> , Parent Interview	
	Parents' Education Style	Father at home		18, 30, 42, 54	54, 60	Demographic Interview
		Role of parents in education		6, 8	8, 12	PEI, Parent Interview <sup>CARE</sup>
		Parents' education beliefs & methods		6, 8	8, 12	PEI, Parent Interview <sup>CARE</sup>
		Parents' aspiration & attitudes towards child		6, 8, 12, 15	8, 12	PEI, Parent Interview
		Parents' occupation		6, 8, 12, 15	7, 8, 12	PEI <sup>ABC</sup> , Parent Interview
		Mother works		9	5, 7, 8	Demographic Interview
Family Economic Environment	Source of child support		6, 8, 12, 15	7, 8, 12	PEI <sup>ABC</sup> , Parent Interview	
	Family income		6, 8, 12, 15	7, 8, 12	PEI <sup>ABC</sup> , Parent Interview	
	Parents' authority, warmth, family conflict, etc.		8	8	Parent Interview	
	Parents' education background		6, 8, 12, 15	7, 8, 12	PEI <sup>ABC</sup> , Parent Interview	
Family Social Status	Criminal history and risk taking of family members		8, 12, 15	-	Subject Taylor Life Events <sup>ABC</sup> , Parent Interview <sup>ABC</sup>	
	Health issues of adults in house		8, 12, 15	12	Parent Interview, Subject Taylor Life Events <sup>ABC</sup>	
Academic Achievements	Standardized Tests	Reading, mathematics, and language abilities	6, 7, 8, 12	6, 8, 9, 12	CAT <sup>ABC</sup> , PIAT <sup>ABC</sup> , WJCA	
		Drop in grades	12, 15	12	CAS	
	Performance in Schoolwork	Lack of interest in school	12, 15	12	CAS	
		Total years in special education	17	11	Retention and Special Services Data	
		Total years retained in school	17	11	Retention and Special Services Data	
	Health Issues	Health issues of subject	8, 12, 15	8, 12	Achenbach Parent, Subject Interview <sup>ABC</sup> , Adult Self Report <sup>ABC</sup> , PEI <sup>CARE</sup> , Parent Interview <sup>CARE</sup>	
		Vision, weight, height	8	8	Growth Data	
	Teenage Pregnancy	Teenage Pregnancy	15	-	Subject Interview <sup>ABC</sup>	
		Law Breaking	15	-	MARS <sup>ABC</sup> , Subject Interview <sup>ABC</sup>	

Sources: Authors' descriptions.

Note: This table describes the major categories of variables that were measured for ABC and CARE subjects at ages 6 to 18. ABC and CARE age are measured in years. This is not an exhaustive list of variables, nor does it include variables from auxiliary data. Instruments or questionnaires available for only one of the studies are indicated with the superscript *ABC* or *CARE*. **Abbreviations are as follows.** MSCD: McCarthy Scales of Children's Development. SBIS: Stanford-Binet Intelligence Scale. WIS: Wechsler Intelligence Scale for Children. KRT: Kohn and Rosman Test Behavior Inventory. WJCA: Woodcock-Johnson Test of Cognitive Abilities. PEI: Parents as Educator Interview. CAS: Child Assessment Schedule. PMI: Psychosocial Maturity Inventory. SAI: Social Adjustment Inventory for Children and Adolescents. SCAN: Schedule of Classroom Activity Norms. CBI: Classroom Behavior Inventory. WPB: Walker Problem Behavior Checklist. ETS: Emotional/Activity/Sociability/Impulsivity Temperament Survey. FES: Family Environment Scale. PIAT: Peabody Individual Achievement Test. CAT: California Achievement Test. MARS: Mid-Adolescence Rating Scale Data.

Table A.8: Adulthood Data (Part I)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Cognitive Assessments	Intelligence Tests	WIS	21	-	WIS
	Interpersonal Skills	Gets along with people	21, 30	-	Subject Interview
Non-Cognitive Assessment	Self-Control	Locus of control	21, 30	-	Nowicki-Strickland Data <sup>ABC</sup> , Pearlin Mastery Scale <sup>ABC</sup>
	Emotional Health	Proud of working, interest in working	21, 30	21, 30	Job Satisfaction Survey <sup>ABC</sup> , Subject Interview
Social Activities	Emotional Health	Harms self, suicidal thoughts, depression, anxiety, fear, etc.	21, 30	21, 30	Achenbach, Subject Risk Taking Survey KRT, Achenbach Parent, CAS, Brief Symptom Inventory, ETS
	Social Activities	Athletic activities Participant of organizations, e.g. religions	21, 30	21, 30	Achenbach, Subject Interview
Family Environment	Family Members	Number of adults in house Relationship with family members, including father, mother, siblings, etc.	21, 30	-	Parent Interview <sup>ABC</sup> , Subject Interview
	Family Economic Environment	Number of siblings Marital status of parents Number of children, childcare basics	21, 30	30	Parent Interview, Adult Self Report Parent Interview <sup>ABC</sup> , Subject Interview Parent Interview <sup>ABC</sup> , Subject Interview Subject Interview, Childcare Questionnaire
Family Members and Children	Family Economic Environment	Parents' occupation Source of child support Family income	21, 30	-	Parent Interview <sup>ABC</sup> , Subject Interview Parent Interview <sup>ABC</sup> , Subject Interview Parent Interview <sup>ABC</sup> , Subject Interview
	Family Members and Children	Relationship quality, family health issues, attitude toward child learning	30	30	Parent Interview, Taylor Life Events <sup>ABC</sup> , Child Health Questionnaire, PEI
Marital Status	Marital Status	Marital status, spouse income Spouse details, marriage history Relationship with spouse	21, 30	21, 30	Subject Interview Subject Interview Subject Interview, Adult Self Report
	Education Level	Years in school, plans for future education College type, certificate earned WJCA	21, 30	21, 30	Subject Interview, Adult Self Report Subject Interview, Adult Self Report WJCA
Achievement	Achievement Test	WJCA	21, 30	-	Subject Interview, Adult Self Report

Sources: Authors' description.

Note: This table describes the major categories of variables that were measured for ABC and CARE subjects at ages 21 and 30. ABC and CARE age are measured in years. This is not an exhaustive list of variables, nor does it include variables from auxiliary data. Instruments or questionnaires available for only one of the studies are indicated with the superscript *ABC* or *CARE*. **Abbreviations are as follows.** KRT: Kohn and Rosman Test Behavior Inventory. CAS: Child Assessment Schedule. ETS: Emotional/Activity/Sociability/Impulsivity Temperament Survey. WIS: Wechsler Adult Intelligence Scale. WJCA: Woodcock-Johnson Test of Cognitive Abilities. PEI: Parents as Educator Interview.

Table A.9: Adulthood Data (Part II)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Physical Health	Health Insurance	Covered by health insurance	21, 30	21, 30	Subject Interview
	Health Issues	Health conditions, diseases, regular checkups and tests, mental health	21, 30	21	Brief Symptom Inventory, Subject Interview, Adult Self Report
Social Conduct	Risk Taking	Smoking, drinking, carry gun, fight, drug use	21, 30	21, 30	Subject Risk Taking Survey, Adult Self Report
Economic Status	Law Breaking	Felony, Time spent incarcerated	21	21, 30	Subject Interview
	Living Circumstances	Number of rooms	21, 30	21, 30	Subject Interview
		Own or rent apartment	21, 30	21	Subject Interview
		Number living in same domicile	21, 30	21	Subject Interview
Working Condition	Currently employed	21, 30	21, 30	Subject Interview	
	Job title	21, 30	21, 30	Subject Interview, Adult Self Report	
	Job category	21, 30	21, 30	Subject Interview, Adult Self Report	
	Hours	21, 30	21, 30	Subject Interview, Adult Self Report	
Transportation	Satisfied with current job	21, 30	21, 30	Subject Interview, Subject What I Am Like (Harter), Adult Self Report	
	Own reliable transportation	21, 30	21	Subject Interview, Adult Self Report	
		Public transportation	21, 30	21	Subject Interview, Adult Self Report
	Income	Income from job	21, 30	21, 30	Subject Interview, Adult Self Report
Income from welfare programs		21, 30	30	Subject Interview, Adult Self Report	
Income from investment		21, 30	-	Subject Interview, Adult Self Report	

Sources: Authors' description.

Note: This table describes the major categories of variables that were measured for ABC and CARE subjects at ages 21 and 30. ABC and CARE age are measured in years. This is not an exhaustive list of variables, nor does it include variables from auxiliary data. Instruments or questionnaires available for only one of the studies are indicated with the superscript *ABC* or *CARE*.

Attrition was low in ABC. Information is available on 100 subjects in the age 30 follow-up, which we call the adult follow-up. In addition, 80 subjects—40 from the control group and 40 from the treatment group—consented to the release of their criminal records. Further, 70 participants consented to the release of information regarding a full-range biomedical panel—31 from the control group and 39 from the treatment group.

Attrition was also low for CARE subjects. Information is available on 58 subjects (more than 85% of the initial sample) in the age-30 follow-up. Additionally, 40 participants (11 from the control group, 18 from the family education group, and 11 from the center-based childcare and family education group) released information on the full-range biomedical sweep. Administrative crime data are not available for CARE. We do not evaluate the second-phase of treatment in CARE because it was not randomized. Rather, those in the center-based childcare and family education group and the family education group were offered school-age treatment, and those in the control group were not.

In the following set of tables (Table A.10 through Table A.18), we compare the observed, baseline characteristics between the first-phase control and treatment groups in ABC, which are the main groups we analyze, at different stages of the data collection follow-ups. For each observed characteristic, we present the bootstrapped  $p$ -value associated with the standard  $t$ -test. We also present the bootstrapped, step-down  $p$ -value on jointly testing the difference in observed characteristics across the two blocks of variables separated by the horizontal line.<sup>70</sup>

First, we compare the first-phase treatment and control groups on baseline characteristics.

---

<sup>70</sup>Lehmann and Romano (2005).

Table A.10: First-phase Treatment vs. Control Groups, ABC

Variable	Age	Control	Treated	Control	Treated	$p$ -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	57	59	0.438	0.489	(0.580)	(0.700)
Birth Weight	0	56	58	7.191	6.829	(0.130)	(0.205)
No. Siblings in Household	0	57	59	0.750	0.516	(0.245)	(0.425)
Birth Year	0	57	59	1974	1974	(0.785)	(0.865)
Mother's Education	0	57	59	9.864	10.505	<b>(0.050)</b>	(0.105)
Mother's Age	0	57	59	20.103	19.564	(0.555)	(0.695)
Mother Employed	0	57	59	0.216	0.317	(0.190)	(0.370)
Parental Income	0	57	58	6,211	7,019	(0.645)	(0.755)
Mother's IQ	0	57	59	83.419	85.393	(0.360)	(0.555)
Father at Home	0	57	59	0.346	0.223	(0.135)	(0.310)

Note: This table shows the balance in observed characteristics between the treatment and control groups in ABC at baseline. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Second, we present the same exercise for each of the four cohorts ABC served.

Table A.11: First-phase Treatment vs. Control Groups, ABC Cohort 1

Variable	Age	Control	Treated	Control	Treated	$p$ -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	14	14	0.348	0.286	(0.730)	(0.738)
Birth Weight	0	14	13	6.755	6.491	(0.550)	(0.655)
No. Siblings in Household	0	14	14	1.741	0.606	<b>(0.035)</b>	<b>(0.085)</b>
Birth Year	0	14	14	1972	1972	(0.240)	(0.350)
Mother's Education	0	14	14	9.885	10.561	(0.265)	(0.480)
Mother's Age	0	14	14	23.869	19.552	<b>(0.050)</b>	(0.135)
Mother Employed	0	14	14	0.152	0.205	(0.695)	(0.895)
Parental Income	0	14	13	7,164	8,298	(0.755)	(0.910)
Mother's IQ	0	14	14	76.042	81.108	(0.270)	(0.485)
Father at Home	0	14	14	0.559	0.368	(0.340)	(0.493)

Note: This table shows the balance in observed characteristics between the treatment and control groups in ABC at baseline for cohort 1. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.12: First-phase Treatment vs. Control Groups, ABC Cohort 2

Variable	Control		Treated	Control		Treated	$p$ -value	
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$	
Male	0	13	16	0.457	0.503	(0.805)	(0.875)	
Birth Weight	0	13	16	7.256	6.534	(0.160)	(0.270)	
No. Siblings in Household	0	13	16	0.388	0.316	(0.755)	(0.835)	
Birth Year	0	13	16	1973	1973	(0.850)	(0.925)	
Mother's Education	0	13	16	10.225	10.307	(0.885)	(0.940)	
Mother's Age	0	13	16	18.446	17.637	(0.380)	(0.630)	
Mother Employed	0	13	16	0.307	0.248	(0.690)	(0.850)	
Parental Income	0	13	16	5,398	4,427	(0.790)	(0.880)	
Mother's IQ	0	13	16	86.873	85.597	(0.730)	(0.855)	
Father at Home	0	13	16	0.220	0.183	(0.790)	(0.895)	

Note: This table shows the balance in observed characteristics between the treatment and control groups in ABC at baseline for cohort 2. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.13: First-phase Treatment vs. Control Groups, ABC Cohort 3

Variable	Control		Treated	Control		Treated	$p$ -value	
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$	
Male	0	14	15	0.376	0.596	(0.265)	(0.320)	
Birth Weight	0	14	15	7.424	7.138	(0.470)	(0.730)	
No. Siblings in Household	0	14	15	0.423	0.203	(0.385)	(0.645)	
Birth Year	0	14	15	1975	1975	(0.510)	(0.520)	
Mother's Education	0	14	15	10.133	10.704	(0.405)	(0.595)	
Mother's Age	0	14	15	18.602	19.558	(0.355)	(0.570)	
Mother Employed	0	14	15	0.162	0.467	<b>(0.070)</b>	(0.155)	
Parental Income	0	14	15	7,034	4,981	(0.430)	(0.675)	
Mother's IQ	0	14	15	85.590	88.715	(0.435)	(0.610)	
Father at Home	0	14	15	0.424	0.209	(0.265)	(0.425)	

Note: This table shows the balance in observed characteristics between the treatment and control groups in ABC at baseline for cohort 3. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.14: First-phase Treatment vs. Control Groups, ABC Cohort 4

Variable	Age	Control	Treated	Control	Treated	<i>p</i> -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	15	14	0.599	0.567	(0.870)	(0.905)
Birth Weight	0	15	14	7.321	7.150	(0.725)	(0.840)
No. Siblings in Household	0	15	14	0.490	0.977	(0.220)	(0.380)
Birth Year	0	15	14	1977	1977	(0.615)	(0.728)
Mother's Education	0	15	14	9.530	10.424	(0.240)	(0.410)
Mother's Age	0	15	14	19.941	21.712	(0.320)	(0.570)
Mother Employed	0	15	14	0.260	0.347	(0.650)	(0.840)
Parental Income	0	15	14	5,827	10,781	<b>(0.065)</b>	(0.135)
Mother's IQ	0	15	14	85.561	86.004	(0.920)	(0.960)
Father at Home	0	15	14	0.208	0.138	(0.570)	(0.777)

Note: This table shows the balance in observed characteristics between the treatment and control groups in ABC at baseline for cohort 4. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Third, we compare the second-phase treatment and control groups on baseline characteristics.

Table A.15: Second-phase Treatment vs. Control Groups, ABC

Variable	Age	Control	Treated	Control	Treated	<i>p</i> -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	47	48	0.551	0.460	(0.420)	(0.552)
Birth Weight	0	47	48	7.084	6.929	(0.610)	(0.700)
No. Siblings in Household	0	47	48	0.748	0.504	(0.285)	(0.445)
Birth Year	0	47	48	1974	1974	(0.835)	(0.915)
Mother's Education	0	47	48	10.150	10.388	(0.480)	(0.725)
Mother's Age	0	47	48	21.122	18.884	<b>(0.035)</b>	<b>(0.075)</b>
Mother Employed	0	47	48	0.314	0.256	(0.530)	(0.725)
Parental Income	0	47	48	7,589	6,714	(0.625)	(0.825)
Mother's IQ	0	47	48	83.000	85.831	(0.185)	(0.365)
Father at Home	0	47	48	0.279	0.287	(0.920)	(0.965)

Note: This table shows the balance in observed characteristics between the school-age treatment and control groups in ABC at baseline. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Fourth, we compare the observed, baseline characteristics of compliant and non-compliant

subjects in the first-phase treatment assignment.

Table A.16: Observed vs. Attritted Children, ABC

Variable	Age	Obs.	Att.	Observed	Attritted	$p$ -value	
				Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	103	13	0.488	0.248	<b>(0.085)</b>	(0.140)
Birth Weight	0	103	11	7.014	6.948	(0.825)	(0.875)
No. Siblings in Household	0	103	13	0.609	0.829	(0.600)	(0.705)
Birth Year	0	103	13	1974	1973	<b>(0.045)</b>	<b>(0.095)</b>
Mother's Education	0	103	13	10.302	9.192	<b>(0.100)</b>	(0.165)
Mother's Age	0	103	13	20.016	18.178	<b>(0.080)</b>	(0.160)
Mother Employed	0	103	13	0.268	0.255	(0.925)	(0.955)
Parental Income	0	103	12	6,622	6,442	(0.950)	(0.960)
Mother's IQ	0	103	13	85.050	78.834	<b>(0.070)</b>	(0.135)
Father at Home	0	103	13	0.278	0.329	(0.735)	(0.835)

Note: This table shows the balance in observed characteristics between ABC subjects who were followed up to at least age 21 and ABC subjects who attritted before age 21. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Fifth, we compare the observed, baseline characteristics between the subjects in the treatment and the control groups, excluding those who did not comply to treatment.

Table A.17: First-phase Treatment vs. Control Groups, Dropping Attritted Children, ABC

Variable	Age	Control		Treated		$p$ -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	51	52	0.452	0.524	(0.430)	(0.600)
Birth Weight	0	51	52	7.210	6.822	(0.115)	(0.220)
No. Siblings in Household	0	51	52	0.767	0.455	(0.150)	(0.230)
Birth Year	0	51	52	1974	1974	(0.635)	(0.785)
Mother's Education	0	51	52	10.000	10.598	<b>(0.085)</b>	(0.185)
Mother's Age	0	51	52	20.412	19.635	(0.405)	(0.615)
Mother Employed	0	51	52	0.221	0.314	(0.245)	(0.455)
Parental Income	0	51	52	6,409	6,846	(0.765)	(0.870)
Mother's IQ	0	51	52	84.472	85.635	(0.560)	(0.755)
Father at Home	0	51	52	0.349	0.208	(0.115)	(0.255)

Note: This table shows the balance in observed characteristics between the treatment and control groups of ABC subjects who were followed up to at least age 21. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Finally, we compare the observed, baseline characteristics between the children in the first-phase treatment, restricting the sample to the children for whom we have information on the age-34 medical data collection.

Table A.18: First-phase Treatment vs. Control Groups, Subjects Completing the Health Follow-up, ABC

Variable	Control		Treated		<i>p</i> -value		
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	31	39	0.293	0.533	<b>(0.050)</b>	<b>(0.055)</b>
Birth Weight	0	31	39	7.233	6.826	(0.190)	(0.295)
No. Siblings in Household	0	31	39	0.613	0.493	(0.580)	(0.750)
Birth Year	0	31	39	1975	1974	(0.360)	(0.510)
Mother's Education	0	31	39	10.039	10.597	(0.190)	(0.385)
Mother's Age	0	31	39	19.389	19.595	(0.825)	(0.945)
Mother Employed	0	31	39	0.195	0.349	(0.185)	(0.315)
Parental Income	0	31	39	5,509	7,520	(0.280)	(0.535)
Mother's IQ	0	31	39	83.822	84.922	(0.655)	(0.860)
Father at Home	0	31	39	0.355	0.231	(0.205)	(0.450)

Note: This table shows the balance in observed characteristics between the treatment and control groups in ABC at baseline for subjects who completed the health follow-up at age 34. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Despite some exceptions, these tables indicate balance between the treatment and control groups from the first-phase randomization, which is the primary comparison we analyze in the main paper. The balance in observed characteristics holds for the different samples we consider, which differs from the initial sample due to various instances of item non-response. For the second-phase randomization, there is also balance in observed characteristics.

Table A.19 through Table A.26 are the analogous tables for CARE. We compare the two treatment groups (center-based childcare and family education, and only family education) separately across the full sample and by cohort. The inference statistics are constructed using the same methods as for Table A.10 through Table A.18.

Table A.19: CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare

Variable	Age	Control	Treated	Control	Treated	<i>p</i> -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	23	17	0.611	0.524	(0.565)	(0.740)
Birth Weight	0	23	15	7.102	7.508	(0.335)	(0.515)
No. Siblings in Household	0	23	17	0.619	0.653	(0.895)	(0.945)
Birth Year	0	23	17	1979	1979	(0.890)	(0.920)
Mother's Education	0	23	17	11.195	10.693	(0.390)	(0.500)
Mother's Age	0	23	17	21.636	21.896	(0.870)	(0.915)
Mother's IQ	0	23	17	87.584	86.624	(0.725)	(0.825)
Father at Home	0	23	17	0.127	0.351	<b>(0.095)</b>	(0.175)

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.20: CARE Baseline Characteristics, Control vs. Family Education

Variable	Age	Control	Treated	Control	Treated	<i>p</i> -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	23	27	0.611	0.632	(0.895)	(0.880)
Birth Weight	0	23	26	7.102	6.963	(0.755)	(0.830)
No. Siblings in Household	0	23	27	0.619	0.758	(0.680)	(0.715)
Birth Year	0	23	27	1979	1979	(0.330)	(0.480)
Mother's Education	0	23	27	11.195	10.689	(0.245)	(0.380)
Mother's Age	0	23	27	21.636	20.257	(0.305)	(0.400)
Mother's IQ	0	23	27	87.584	87.167	(0.855)	(0.915)
Father at Home	0	23	27	0.127	0.190	(0.455)	(0.585)

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.21: CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare, Cohort 5

Variable	Age	Control	Treated	Control	Treated	<i>p</i> -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	7	6	0.560	0.655	(0.810)	(0.860)
Birth Weight	0	7	4	7.223	7.502	(0.570)	(0.730)
No. Siblings in Household	0	7	6	0.428	0.541	(0.800)	(0.870)
Birth Year	0	7	6	1978	1978	(0.425)	(0.355)
Mother's Education	0	7	6	11.035	11.164	(0.865)	(0.875)
Mother's Age	0	7	6	18.808	21.652	(0.140)	(0.220)
Mother's IQ	0	7	6	89.202	92.345	(0.620)	(0.680)
Father at Home	0	7	6	0.289	0.322	(0.935)	(0.938)

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline for cohort 5. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.22: CARE Baseline Characteristics, Control vs. Family Education, Cohort 5

Variable	Age	Control	Treated	Control	Treated	<i>p</i> -value	
		Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$
Male	0	7	14	0.560	0.504	(0.885)	(0.885)
Birth Weight	0	7	14	7.223	6.742	(0.580)	(0.710)
No. Siblings in Household	0	7	14	0.428	1.046	(0.295)	(0.365)
Birth Year	0	7	14	1978	1978	(0.175)	(0.190)
Mother's Education	0	7	14	11.035	10.699	(0.610)	(0.735)
Mother's Age	0	7	14	18.808	20.824	(0.210)	(0.285)
Mother's IQ	0	7	14	89.202	90.710	(0.695)	(0.775)
Father at Home	0	7	14	0.289	0.219	(0.755)	(0.790)

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline for cohort 5. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.23: CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare, Cohort 6

Variable	Control		Treated	Control		Treated	$p$ -value	
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$	
Male	0	16	11	0.636	0.453	(0.395)	(0.625)	
Birth Weight	0	16	11	7.041	7.509	(0.410)	(0.645)	
No. Siblings in Household	0	16	11	0.703	0.720	(0.975)	(0.980)	
Birth Year	0	16	11	1979	1979	(0.565)	(0.498)	
Mother's Education	0	16	11	11.268	10.441	(0.250)	(0.385)	
Mother's Age	0	16	11	22.884	22.039	(0.690)	(0.750)	
Mother's IQ	0	16	11	86.841	83.388	(0.245)	(0.370)	
Father at Home	0	16	11	0.057	0.358	<b>(0.045)</b>	<b>(0.095)</b>	

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline for cohort 6. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.24: CARE Baseline Characteristics, Control vs. Family Education, Cohort 6

Variable	Control		Treated	Control		Treated	$p$ -value	
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$	
Male	0	16	12	0.636	0.747	(0.575)	(0.655)	
Birth Weight	0	16	12	7.041	7.208	(0.675)	(0.745)	
No. Siblings in Household	0	16	12	0.703	0.490	(0.515)	(0.600)	
Birth Year	0	16	12	1979	1979	(0.420)	(0.540)	
Mother's Education	0	16	12	11.268	10.668	(0.355)	(0.493)	
Mother's Age	0	16	12	22.884	19.905	<b>(0.075)</b>	(0.125)	
Mother's IQ	0	16	12	86.841	82.920	(0.185)	(0.295)	
Father at Home	0	16	12	0.057	0.177	(0.380)	(0.420)	

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline for cohort 6. For each characteristic, we present the  $p$ -value from a single hypothesis test. We also present the  $p$ -values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both  $p$ -values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.25: CARE Baseline Characteristics, Control vs. Family Education and Center-based Childcare  
Subjects Completing the Health Follow-up

Variable	Control		Treated	Control		Treated		<i>p</i> -value	
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$		
Male	0	11	11	0.467	0.550	(0.690)	(0.815)		
Birth Weight	0	11	11	6.783	7.633	(0.110)	(0.200)		
No. Siblings in Household	0	11	11	0.372	0.546	(0.665)	(0.740)		
Birth Year	0	11	11	1979	1979	(0.115)	(0.193)		
Mother's Education	0	11	11	11.391	11.027	(0.615)	(0.703)		
Mother's Age	0	11	11	22.142	21.607	(0.865)	(0.860)		
Mother's IQ	0	11	11	86.317	87.505	(0.745)	(0.825)		
Father at Home	0	11	11	0.085	0.362	(0.110)	(0.185)		

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline for subjects who completed the health follow-up at age 34. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Table A.26: CARE Baseline Characteristics, Control vs. Family Education  
Subjects Completing the Health Follow-up

Variable	Control		Treated	Control		Treated		<i>p</i> -value	
	Age	Obs.	Obs.	Mean	Mean	Single $H_0$	Multiple $H_0$		
Male	0	11	18	0.467	0.446	(0.940)	(0.935)		
Birth Weight	0	11	18	6.783	6.262	(0.325)	(0.430)		
No. Siblings in Household	0	11	18	0.372	0.383	(0.960)	(0.955)		
Birth Year	0	11	18	1979	1979	(0.120)	(0.240)		
Mother's Education	0	11	18	11.391	11.236	(0.795)	(0.845)		
Mother's Age	0	11	18	22.142	19.941	(0.230)	(0.290)		
Mother's IQ	0	11	18	86.317	87.611	(0.700)	(0.745)		
Father at Home	0	11	18	0.085	0.237	(0.240)	(0.320)		

Note: This table shows the balance in observed characteristics between the treatment and control groups at baseline for subjects who completed the health follow-up at age 34. For each characteristic, we present the *p*-value from a single hypothesis test. We also present the *p*-values from multiple hypothesis testing, where we collectively test the baseline characteristics within the blocks separated by the horizontal line. Both *p*-values are two-sided and non-parametric. We construct them based on 200 re-draws of the full sample. The estimates we display are the means of the empirical bootstrap distribution.

Overall, these tables indicate a balance between the treatment and control groups, both when considering center-based childcare and family education and only family education as treatment.

## B Background Variables and Tested Outcomes

### B.1 Controlling for Background Variables

We select the background variables that we control for as follows. First, we select the three variables, from 17 potential variables, that best predict enrollment into alternative preschool for subjects who were randomly assigned not to receive center-based childcare—the control group in ABC and the control and family education treatment groups in CARE. To choose the most predictive control set, we estimate the  $\binom{17}{3}$  linear models, using the variables in Panels (1) and (2) of Table B.1. The most predictive model is the one generating the lowest Bayesian Information Criterion (BIC).

We proceed in a similar way to choose the second set of background variables. Instead of predicting enrollment into alternative preschool, we predict the relevant outcome of interest. For this exercise we use the variables in Panel (2), so we estimate  $\binom{14}{3}$  models.

Our set of control variables consists of these two sets of background variables.

Table B.1: Background Variables

1. Prediction of Preschool Take-up		
Number of relatives index	Grandma in county	Born in the fall
2. Prediction of Outcomes		
Maternal IQ	Maternal education	Mother's age at birth
High Risk Index	Parent income	Premature birth
1 minute Apgar score	5 minute Apgar score	Mother married
Teen pregnancy	Father at home	Number of siblings
Cohort	Mother is employed	

Note: This table lists the variables we permute over when selecting the background variables we control for in our estimations. We use two sets of background variables as controls. The first set is the one that best predicts selection into preschool. We choose the three variables that minimize the Bayesian Information Criterion in a regression of a preschool enrollment indicator on three out of the 17 variables listed in Panels (1) and (2) in the table. The second set of background variables is the one that best predicts each outcome, using a similar criterion, based on the 14 variables listed in Panel (2). We pool the six controls resulting from these two selection processes.

## B.2 Tested Outcomes

The following table lists the outcomes we test treatment effects for when calculating our statistics. We reverse the outcomes for which we consider a negative treatment effect socially positive.

Table B.2: Outcome Variables

Category	Variable	Age	ABC	CARE	Re-versed
IQ Scores	Std. IQ Test	2	✓	✓	
		2.5		✓	
		3	✓	✓	
		3.5	✓	✓	
		4	✓	✓	
		4.5	✓	✓	
		5	✓	✓	
		6.6	✓	✓	

Category	Variable	Age	ABC	CARE	Re-versed	
		7	✓	✓		
		8	✓	✓		
		12	✓	✓		
		15	✓			
		21	✓			
	IQ Factor	2 to 5	✓	✓		
		6 to 12	✓	✓		
		15 to 21	✓			
Achievement Scores	Std. Achv. Test	5.5	✓	✓		
		6	✓	✓		
		6.5	✓			
		7	✓			
		7.5	✓	✓		
		8	✓	✓		
		8.5	✓	✓		
		12		✓		
		15	✓			
	21	✓				
		PIAT Math Std. Score	7	✓	✓	
		Achievement Factor	5.5 to 12	✓	✓	
			15 to 21	✓		
HOME Scores	HOME Score	0.5	✓	✓		
		1.5	✓	✓		
		2.5	✓	✓		
		3.5	✓	✓		
		4.5	✓	✓		
		8	✓	✓		
		HOME Factor	0.5 to 8	✓	✓	
	Parent Income	Parental income	1.5	✓	✓	
2.5			✓	✓		
3.5			✓	✓		
4.5			✓	✓		

Category	Variable	Age	ABC	CARE	Re-versed
		8	✓		
		12	✓		
		15	✓		
	Parental Income Factor	1.5 to 15	✓	✓	
Mother's Employment	Mother Works	2	✓	✓	
		3	✓	✓	
		4	✓	✓	
		5	✓	✓	
		21	✓		
	Mother Works Factor	2 to 21	✓	✓	
Mother's Education	Mother's Years of Edu.	2	✓		
		3	✓		
		4	✓		
		5	✓		
		9	✓		
	Mother's Edu. Factor	2 to 9	✓		
Father at Home	Father at Home	2	✓	✓	
		3	✓	✓	
		4	✓	✓	
		5	✓	✓	
		8	✓	✓	
	Father at Home Factor	2 to 8	✓	✓	
Adoption	Ever Adopted		✓		
Education	Graduated High School	30	✓	✓	
	Attended Voc./Tech./Com. College	30	✓	✓	
	Graduated 4-year College	30	✓	✓	
	Years of Edu.	30	✓	✓	
	Education Factor	30	✓	✓	

Category	Variable	Age	ABC	CARE	Re-versed
Employment and Income	Employed	30	✓	✓	
	Labor Income	21	✓	✓	
		30	✓	✓	
	Public-Transfer Income	21	✓	✓	✓
		30	✓	✓	✓
Employment Factor	21 to 30	✓	✓		
Crime	Total Felony Arrests	Mid-30s	✓	✓	✓
	Total Misdemeanor Arrests	Mid-30s	✓	✓	✓
	Total Years Incarcerated	30	✓	✓	✓
	Crime Factor	30 to Mid-30s	✓	✓	✓
Tobacco, Drugs, Alcohol	Cig. Smoked per day last month	30	✓	✓	✓
	Days drank alcohol last month	30	✓	✓	✓
	Days binge drank alcohol last month	30	✓	✓	✓
	Self-reported drug user	Mid-30s	✓	✓	✓
	Substance Use Factor	30 to Mid-30s	✓	✓	✓
Self-Reported Health	Self-reported Health	30	✓	✓	✓
		Mid-30s	✓	✓	✓
	Self-reported Health Factor	30 to Mid-30s	✓	✓	✓
Hypertension	Systolic Blood Pressure (mm Hg)	Mid-30s	✓	✓	✓
	Diastolic Blood Pressure (mm Hg)	Mid-30s	✓	✓	✓

Category	Variable	Age	ABC	CARE	Reversed
	Prehypertension	Mid-30s	✓	✓	✓
	Hypertension	Mid-30s	✓	✓	✓
	Hypertension Factor	Mid-30s	✓	✓	✓
Cholesterol	High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	✓	✓	
	Dyslipidemia	Mid-30s	✓	✓	✓
	Cholesterol Factor	Mid-30s	✓	✓	✓
Diabetes	Hemoglobin Level (%)	Mid-30s	✓	✓	✓
	Prediabetes	Mid-30s	✓	✓	✓
	Diabetes	Mid-30s	✓	✓	✓
	Diabetes Factor	Mid-30s	✓	✓	✓
Vitamin D Deficiency	Vitamin D Deficiency	Mid-30s	✓	✓	✓
Obesity	Measured BMI	Mid-30s	✓	✓	✓
	Obesity	Mid-30s	✓	✓	✓
	Severe Obesity	Mid-30s	✓	✓	✓
	Waist-hip Ratio	Mid-30s	✓	✓	✓
	Abdominal Obesity	Mid-30s	✓	✓	✓
	Framingham Risk Score	Mid-30s	✓	✓	✓
	Obesity Factor	Mid-30s	✓	✓	✓

Category	Variable	Age	ABC	CARE	Re-versed
Mental Health (BSI)	Somatization	21	✓	✓	✓
		34	✓	✓	✓
	Depression	21	✓	✓	✓
		34	✓	✓	✓
	Anxiety	21	✓	✓	✓
		34	✓	✓	✓
	Hostility	21	✓	✓	✓
		34	✓	✓	✓
	Global Severity Index	21	✓	✓	✓
		34	✓	✓	✓
	21				
	Mental Health Factor	and 34	✓	✓	✓
Child Behavior (CAS)	Participates in Activity	12	✓		
	Time Spent Reading	12	✓		
	Good Description of Self	12	✓		
	Views Self as Dumb	12	✓		✓
	Views Self as Clumsy	12	✓		✓
	Views Self as Not Liked	12	✓		✓
	Proud About Self	12	✓		
	Family Proud of You	12	✓		
	Feels Inadequate, Inferior	12	✓		✓
	Withdraws Excessively	12	✓		✓
	Ignores Situation	12	✓		✓
	Not Cope with Prob.	12	✓		✓
	Often Mad or Angry	12	✓		✓
	Impulsivity	12	✓		✓
	Significant Fears	12	✓		✓
	Denies Any Worries	12	✓		✓

Note: This table lists the outcomes that we test treatment effects for. We reverse the outcomes for which we consider a negative treatment effect socially positive.

## C Results

In this appendix we present the treatment effects estimates of the center-based treatment in ABC and CARE. For each set of estimates, we first present a summary of the effect of the

program(s) using a combining function counting the number of socially positive treatment effects. We then present tables of treatment effect estimates for each outcome. Finally, we present the treatment effect estimates using the step-down procedure to test multiple hypotheses.

We analyze 95 outcomes when combining the ABC and CARE samples. Table C.1 shows that across all methods of estimation, pooling males and females, over 70% of the treatment effect estimates are socially positive. When conditioning on 10% statistical significance, almost 40% of all estimates are socially positive. These statistics allow us to reject the hypothesis that the treatment and control groups have the same marginal distributions in outcomes.

For both males and females, we find positive effects in IQ test scores, achievement test scores, as well as educational attainment. Males also enjoy additional benefits in the areas of employment, labor earnings, and hypertension.

In each of the tables for combining functions and treatment effect estimates, we present 8 different estimates. Column (1) corresponds to the mean difference between the groups randomly assigned to receive center-based childcare and the groups randomly assigned not to. Column (2) adjusts the estimates in (1) for attrition and controls for a set of covariates. Column (3) corresponds to the mean difference between the groups randomly assigned to receive center-based childcare and the groups randomly assigned not to, restricting the latter to subjects who did not receive preschool alternatives. Column (4) adjusts the estimates in (3) for attrition and controls for a set of covariates. Column (5) corresponds to the mean difference between the groups randomly assigned to receive center-based childcare and the groups randomly assigned not to, placing a relatively high weight on the subjects who are likely not to be enrolled in alternative preschools. Column (6) corresponds to the mean

difference between the groups randomly assigned to receive center-based childcare and the groups randomly assigned not to, restricting the latter to subjects who received preschool alternatives. Column (7) adjusts the estimates in (6) for attrition and controls for a set of covariates. Column (8) corresponds to the mean difference between the groups randomly assigned to receive center-based childcare and the groups randomly assigned not to, placing a relatively high weight on the children who are likely to be enrolled in alternative preschools. The results in bold are significant at the 10% level in a single-sided, non-parametric, bootstrapped test.

## C.1 Combining Functions, Aggregated

Table C.1: Combining Functions, Pooled Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Pos. TE	80	78	81	80	78	80	74	81
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
% Pos. TE   10% Significance	49	41	40	37	41	37	36	42
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>

Table C.2: Combining Functions, Male Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Pos. TE	73	78	51	55	47	74	78	79
	<b>(0.000)</b>	<b>(0.000)</b>	(0.539)	(0.342)	(0.658)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
% Pos. TE   10% Significance	28	29	15	13	15	29	25	29
	<b>(0.026)</b>	<b>(0.000)</b>	(0.289)	(0.263)	(0.276)	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.013)</b>

Table C.3: Combining Functions, Female Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Pos. TE	84	78	82	82	84	78	71	79
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>
% Pos. TE   10% Significance	46	31	53	41	55	33	15	33
	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.026)</b>	(0.237)	<b>(0.013)</b>

## C.2 Combining Functions, by Category

Table C.4: Combining Functions by Category, Pooled Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
IQ Scores	100	100	100	100	100	100	100	100	12
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Achievement Scores	100	100	100	100	100	100	100	100	6
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
HOME Scores	100	100	100	100	100	86	29	86	7
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.118)	(0.618)	(0.184)	
Parent Income	100	100	80	80	100	100	100	100	5
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Mother's Employment	100	100	100	100	100	100	80	100	5
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.224)	<b>(0.000)</b>	
Father at Home	17	33	0	0	0	83	100	100	6
	(1.000)	(0.618)	(1.000)	(1.000)	(1.000)	(0.382)	<b>(0.000)</b>	<b>(0.000)</b>	
Education	80	80	100	100	80	80	80	80	5
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Employment and Income	100	83	100	100	100	83	83	83	6
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Crime	25	25	50	25	25	25	25	25	4
	(0.987)	(0.882)	(0.526)	(0.908)	(0.921)	(0.961)	(0.855)	(0.842)	
Tobacco, Drugs, Alcohol	20	20	80	80	60	20	20	20	5
	(0.961)	(0.987)	<b>(0.026)</b>	<b>(0.053)</b>	(0.250)	(0.868)	(0.908)	(0.908)	
Self-Reported Health	67	67	100	100	100	67	33	33	3
	(0.171)	(0.145)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.329)	(0.882)	(0.921)	
Hypertension	100	100	40	20	40	100	100	100	5
	<b>(0.000)</b>	<b>(0.000)</b>	(0.526)	(0.737)	(0.566)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Cholesterol	67	100	67	67	67	67	100	67	3
	(0.487)	<b>(0.000)</b>	(0.329)	(0.329)	(0.316)	(0.382)	<b>(0.000)</b>	(0.434)	
Diabetes	50	25	25	100	25	75	25	100	4
	(0.474)	(0.711)	(0.684)	<b>(0.000)</b>	(0.737)	(0.250)	(0.671)	<b>(0.000)</b>	
Vitamin D Deficiency	100	100	100	100	100	100	100	100	1
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Obesity	71	43	86	71	71	43	29	29	7
	<b>(0.079)</b>	(0.632)	<b>(0.000)</b>	<b>(0.092)</b>	<b>(0.039)</b>	(0.579)	(0.750)	(0.776)	
Mental Health	100	100	100	91	100	91	100	100	11
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	

Table C.5: Combining Functions by Category | 10% Significance, Pooled Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
IQ Scores	100 <b>(0.000)</b>	100 <b>(0.000)</b>	75 <b>(0.000)</b>	75 <b>(0.000)</b>	75 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	12
Achievement Scores	83 <b>(0.000)</b>	100 <b>(0.000)</b>	50 (0.132)	83 <b>(0.000)</b>	50 (0.118)	100 <b>(0.000)</b>	83 <b>(0.000)</b>	83 <b>(0.000)</b>	6
HOME Scores	71 <b>(0.013)</b>	0 (1.000)	71 <b>(0.000)</b>	57 <b>(0.026)</b>	86 <b>(0.000)</b>	0 (1.000)	0 (1.000)	0 (1.000)	7
Parent Income	0 (0.447)	20 (0.250)	0 (0.553)	60 <b>(0.092)</b>	40 (0.263)	0 (0.329)	0 (0.355)	20 (0.276)	5
Mother's Employment	100 <b>(0.000)</b>	40 (0.224)	100 <b>(0.000)</b>	80 <b>(0.000)</b>	100 <b>(0.000)</b>	0 (0.408)	0 (0.342)	0 (0.342)	5
Father at Home	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	17 (0.105)	33 (0.211)	67 (0.105)	6
Education	80 <b>(0.000)</b>	60 <b>(0.000)</b>	100 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.000)</b>	60 <b>(0.000)</b>	40 (0.250)	60 <b>(0.000)</b>	5
Employment and Income	50 <b>(0.066)</b>	50 <b>(0.013)</b>	67 <b>(0.013)</b>	0 (0.579)	83 <b>(0.000)</b>	50 <b>(0.026)</b>	33 <b>(0.026)</b>	50 <b>(0.013)</b>	6
Crime	25 <b>(0.066)</b>	0 (1.000)	25 (0.263)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	4
Tobacco, Drugs, Alcohol	20 (0.447)	20 <b>(0.053)</b>	0 (1.000)	20 <b>(0.039)</b>	20 (0.421)	0 (1.000)	0 (1.000)	0 (1.000)	5
Self-Reported Health	0 (1.000)	0 (1.000)	33 (0.105)	33 <b>(0.039)</b>	33 (0.105)	0 (1.000)	0 (1.000)	0 (1.000)	3
Hypertension	60 <b>(0.039)</b>	80 <b>(0.000)</b>	0 (1.000)	0 (1.000)	0 (1.000)	80 <b>(0.000)</b>	80 <b>(0.000)</b>	100 <b>(0.000)</b>	5
Cholesterol	33 (0.197)	33 (0.171)	33 <b>(0.079)</b>	0 (1.000)	33 (0.118)	0 (0.237)	33 (0.224)	33 (0.158)	3
Diabetes	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	4
Vitamin D Deficiency	100 <b>(0.000)</b>	0 (0.184)	100 <b>(0.000)</b>	0 (1.000)	100 <b>(0.000)</b>	0 (0.342)	0 (1.000)	0 (0.237)	1
Obesity	14 (0.316)	0 (1.000)	29 (0.237)	0 (0.526)	0 (0.579)	14 (0.461)	14 (0.421)	0 (1.000)	7
Mental Health	45 <b>(0.079)</b>	55 <b>(0.039)</b>	18 (0.263)	36 <b>(0.066)</b>	9 (0.461)	45 (0.105)	45 <b>(0.066)</b>	55 <b>(0.013)</b>	11

Table C.6: Combining Functions by Category, Male Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
IQ Scores	100 <b>(0.000)</b>	92 <b>(0.000)</b>	83 <b>(0.000)</b>	75 <b>(0.000)</b>	83 <b>(0.000)</b>	100 <b>(0.000)</b>	92 <b>(0.000)</b>	92 <b>(0.000)</b>	12
Achievement Scores	100 <b>(0.000)</b>	100 <b>(0.000)</b>	50 (0.487)	67 (0.408)	50 (0.474)	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	6
HOME Scores	86 <b>(0.000)</b>	86 (0.132)	86 <b>(0.000)</b>	43 (0.553)	86 <b>(0.000)</b>	71 (0.355)	86 (0.105)	71 (0.171)	7
Parent Income	100 <b>(0.000)</b>	40 (0.447)	60 (0.474)	40 (0.461)	40 (0.592)	60 (0.579)	40 (0.421)	80 (0.316)	5
Mother's Employment	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	80 <b>(0.000)</b>	100 <b>(0.000)</b>	80 <b>(0.000)</b>	5
Father at Home	17 (0.776)	100 <b>(0.000)</b>	0 (1.000)	67 (0.382)	0 (1.000)	80 (0.487)	100 <b>(0.000)</b>	100 <b>(0.000)</b>	6
Education	80 <b>(0.000)</b>	80 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.053)</b>	5
Employment and Income	67 (0.118)	67 <b>(0.053)</b>	33 (0.671)	17 (1.000)	50 (0.474)	83 <b>(0.000)</b>	67 <b>(0.092)</b>	67 (0.105)	6
Crime	25 (0.882)	25 (0.750)	25 (1.000)	0 (1.000)	25 (1.000)	25 (0.961)	25 (0.789)	25 (0.921)	4
Tobacco, Drugs, Alcohol	20 (1.000)	20 (1.000)	40 (0.605)	80 (0.237)	20 (0.921)	20 (1.000)	20 (1.000)	20 (0.974)	5
Self-Reported Health	67 (0.434)	33 (0.724)	67 <b>(0.053)</b>	67 <b>(0.039)</b>	67 <b>(0.053)</b>	33 (0.895)	33 (0.763)	67 (0.355)	3
Hypertension	100 <b>(0.000)</b>	100 <b>(0.000)</b>	40 (0.658)	20 (0.618)	40 (0.697)	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	5
Cholesterol	100 <b>(0.000)</b>	100 <b>(0.000)</b>	0 (0.908)	0 (0.908)	0 (0.908)	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	3
Diabetes	25 (0.750)	25 (0.763)	25 (0.711)	25 (0.474)	25 (0.789)	25 (0.763)	25 (0.763)	25 (0.829)	4
Vitamin D Deficiency	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	1
Obesity	67 (0.395)	100 <b>(0.000)</b>	57 (0.421)	86 <b>(0.000)</b>	14 (0.737)	86 <b>(0.000)</b>	86 <b>(0.000)</b>	100 <b>(0.000)</b>	6
Mental Health	73 (0.303)	91 <b>(0.000)</b>	9 (1.000)	36 (0.763)	18 (0.947)	73 (0.132)	100 <b>(0.000)</b>	91 <b>(0.000)</b>	11

Table C.7: Combining Functions by Category | 10% Significance, Male Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
IQ Scores	75 <b>(0.000)</b>	75 <b>(0.000)</b>	50 <b>(0.053)</b>	33 (0.184)	50 <b>(0.026)</b>	83 <b>(0.000)</b>	75 <b>(0.000)</b>	83 <b>(0.000)</b>	12
Achievement Scores	33 (0.224)	83 <b>(0.000)</b>	17 (0.197)	17 (0.263)	17 (0.237)	67 <b>(0.079)</b>	67 <b>(0.079)</b>	50 (0.224)	6
HOME Scores	0 (0.408)	0 (0.368)	0 (0.421)	0 (1.000)	0 (0.382)	0 (1.000)	0 (1.000)	0 (0.382)	7
Parent Income	0 (0.237)	0 (1.000)	0 (0.250)	0 (1.000)	0 (0.211)	0 (0.250)	0 (1.000)	0 (0.237)	5
Mother's Employment	60 <b>(0.092)</b>	20 (0.316)	40 (0.342)	0 (0.237)	40 (0.289)	0 (0.513)	0 (0.487)	0 (0.474)	5
Father at Home	0 (1.000)	17 (0.224)	0 (1.000)	0 (1.000)	0 (1.000)	0 (0.303)	17 (0.289)	50 (0.197)	6
Education	40 (0.237)	60 <b>(0.066)</b>	0 (0.421)	60 <b>(0.026)</b>	0 (0.487)	40 (0.184)	20 (0.276)	40 (0.184)	5
Employment and Income	33 <b>(0.053)</b>	33 (0.105)	0 (1.000)	0 (1.000)	0 (1.000)	33 <b>(0.066)</b>	33 (0.158)	33 <b>(0.079)</b>	6
Crime	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	4
Tobacco, Drugs, Alcohol	20 <b>(0.039)</b>	20 <b>(0.026)</b>	20 <b>(0.066)</b>	20 (0.158)	20 <b>(0.079)</b>	20 (0.395)	20 (0.500)	20 (0.474)	5
Self-Reported Health	0 (1.000)	0 (1.000)	67 <b>(0.000)</b>	67 <b>(0.000)</b>	67 <b>(0.000)</b>	0 (1.000)	0 (1.000)	0 (1.000)	3
Hypertension	60 (0.118)	80 <b>(0.000)</b>	20 (0.382)	20 (0.158)	20 (0.368)	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	5
Cholesterol	67 <b>(0.000)</b>	0 (0.289)	0 (0.908)	0 (0.908)	0 (0.908)	67 <b>(0.000)</b>	33 (0.145)	67 <b>(0.000)</b>	3
Diabetes	0 (1.000)	0 (1.000)	0 (0.368)	0 (0.908)	0 (0.395)	0 (1.000)	0 (1.000)	0 (1.000)	4
Vitamin D Deficiency	100 <b>(0.000)</b>	0 (0.211)	100 <b>(0.000)</b>	0 (0.895)	100 <b>(0.000)</b>	0 (0.329)	0 (1.000)	0 (0.316)	1
Obesity	17 (0.145)	0 (0.395)	0 (0.329)	0 (0.500)	0 (0.342)	14 (0.184)	0 (1.000)	0 (0.368)	6
Mental Health	0 (1.000)	18 (0.224)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (0.500)	0 (1.000)	11

Table C.8: Combining Functions by Category, Female Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
IQ Scores	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	92 <b>(0.000)</b>	100 <b>(0.000)</b>	12
Achievement Scores	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	6
HOME Scores	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	29 <b>(0.829)</b>	29 <b>(0.868)</b>	71 <b>(0.303)</b>	7
Parent Income	100 <b>(0.000)</b>	100 <b>(0.000)</b>	80 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	5
Mother's Employment	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	80 <b>(0.250)</b>	60 <b>(0.526)</b>	80 <b>(0.368)</b>	5
Father at Home	17 <b>(1.000)</b>	33 <b>(0.579)</b>	0 <b>(1.000)</b>	17 <b>(1.000)</b>	0 <b>(1.000)</b>	67 <b>(0.487)</b>	83 <b>(0.000)</b>	83 <b>(0.368)</b>	6
Education	80 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.000)</b>	80 <b>(0.026)</b>	80 <b>(0.000)</b>	5
Employment and Income	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	83 <b>(0.000)</b>	67 <b>(0.276)</b>	83 <b>(0.000)</b>	6
Crime	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	75 <b>(0.408)</b>	4
Tobacco, Drugs, Alcohol	80 <b>(0.211)</b>	20 <b>(0.882)</b>	80 <b>(0.000)</b>	40 <b>(0.658)</b>	80 <b>(0.000)</b>	100 <b>(0.000)</b>	20 <b>(0.789)</b>	40 <b>(0.737)</b>	5
Self-Reported Health	33 <b>(0.526)</b>	67 <b>(0.276)</b>	67 <b>(0.184)</b>	67 <b>(0.066)</b>	67 <b>(0.132)</b>	0 <b>(1.000)</b>	100 <b>(0.000)</b>	67 <b>(0.066)</b>	3
Hypertension	80 <b>(0.105)</b>	60 <b>(0.395)</b>	20 <b>(0.776)</b>	20 <b>(0.816)</b>	40 <b>(0.447)</b>	80 <b>(0.158)</b>	60 <b>(0.474)</b>	80 <b>(0.316)</b>	5
Cholesterol	67 <b>(0.171)</b>	67 <b>(0.289)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	33 <b>(0.711)</b>	67 <b>(0.487)</b>	67 <b>(0.079)</b>	3
Diabetes	75 <b>(0.000)</b>	75 <b>(0.158)</b>	67 <b>(0.395)</b>	67 <b>(0.592)</b>	67 <b>(0.395)</b>	75 <b>(0.276)</b>	75 <b>(0.158)</b>	75 <b>(0.000)</b>	4
Vitamin D Deficiency	100 <b>(0.000)</b>	0 <b>(1.000)</b>	100 <b>(0.000)</b>	100 <b>(0.474)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	0 <b>(1.000)</b>	0 <b>(1.000)</b>	1
Obesity	71 <b>(0.329)</b>	29 <b>(0.711)</b>	71 <b>(0.263)</b>	71 <b>(0.132)</b>	71 <b>(0.237)</b>	57 <b>(0.303)</b>	14 <b>(1.000)</b>	43 <b>(0.592)</b>	7
Mental Health	91 <b>(0.000)</b>	91 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	91 <b>(0.000)</b>	91 <b>(0.000)</b>	91 <b>(0.000)</b>	11

Table C.9: Combining Functions by Category | 10% Significance, Female Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	
IQ Scores	100 <b>(0.000)</b>	83 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	100 <b>(0.000)</b>	83 <b>(0.000)</b>	25 (0.171)	83 <b>(0.000)</b>	12
Achievement Scores	100 <b>(0.000)</b>	100 <b>(0.000)</b>	83 <b>(0.000)</b>	100 <b>(0.000)</b>	83 <b>(0.000)</b>	100 <b>(0.000)</b>	50 (0.145)	100 <b>(0.000)</b>	6
HOME Scores	29 (0.224)	0 (1.000)	86 <b>(0.000)</b>	57 <b>(0.053)</b>	86 <b>(0.000)</b>	0 (1.000)	0 (1.000)	14 (0.237)	7
Parent Income	40 (0.184)	40 (0.105)	0 (0.434)	40 <b>(0.053)</b>	60 <b>(0.092)</b>	20 (0.276)	40 <b>(0.053)</b>	40 (0.184)	5
Mother's Employment	40 (0.197)	0 (0.276)	60 (0.145)	20 (0.197)	60 (0.132)	0 (0.197)	0 (1.000)	0 (1.000)	5
Father at Home	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (0.329)	0 (0.289)	6
Education	80 <b>(0.000)</b>	40 (0.118)	80 <b>(0.000)</b>	60 <b>(0.000)</b>	80 <b>(0.000)</b>	40 (0.132)	20 (0.250)	40 (0.211)	5
Employment and Income	67 <b>(0.026)</b>	17 (0.224)	83 <b>(0.000)</b>	67 <b>(0.053)</b>	100 <b>(0.000)</b>	17 (0.224)	0 (1.000)	17 (0.211)	6
Crime	100 <b>(0.000)</b>	25 (0.224)	33 (0.368)	0 (0.184)	0 (0.447)	75 <b>(0.013)</b>	25 (0.303)	25 (0.118)	4
Tobacco, Drugs, Alcohol	0 (0.355)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	0 (1.000)	5
Self-Reported Health	0 (1.000)	0 (1.000)	33 (0.105)	33 <b>(0.000)</b>	33 (0.118)	0 (1.000)	0 (1.000)	0 (1.000)	3
Hypertension	20 (0.105)	20 <b>(0.092)</b>	0 (1.000)	0 (1.000)	0 (1.000)	20 (0.132)	20 <b>(0.066)</b>	20 (0.197)	5
Cholesterol	0 (0.224)	0 (0.303)	67 <b>(0.000)</b>	0 (1.000)	67 <b>(0.000)</b>	0 (1.000)	0 (1.000)	0 (1.000)	3
Diabetes	50 <b>(0.066)</b>	0 (1.000)	0 (0.368)	0 (1.000)	0 (0.342)	25 (0.237)	0 (1.000)	50 <b>(0.066)</b>	4
Vitamin D Deficiency	0 (0.987)	0 (1.000)	0 (0.355)	0 (1.000)	0 (0.316)	0 (0.105)	0 (1.000)	0 (1.000)	1
Obesity	0 (0.579)	0 (1.000)	57 <b>(0.013)</b>	0 (1.000)	57 <b>(0.026)</b>	0 (1.000)	0 (1.000)	0 (1.000)	7
Mental Health	45 <b>(0.053)</b>	55 <b>(0.000)</b>	55 <b>(0.026)</b>	45 <b>(0.000)</b>	45 <b>(0.026)</b>	55 <b>(0.026)</b>	27 (0.250)	45 (0.105)	11

### C.3 Treatment Effects for Pooled Sample

Table C.10: Treatment Effects on IQ Scores, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. IQ Test	2	10.116	10.461	10.609	12.312	11.140	9.863	10.116	10.214
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3	13.450	14.071	19.242	21.795	20.287	11.314	12.053	11.777
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3.5	8.387	7.886	11.255	12.220	11.438	7.276	6.711	7.008
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>
	4	9.166	8.768	11.985	12.570	12.562	8.149	7.853	8.525
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	4.5	8.380	8.004	13.287	13.937	13.523	6.717	6.246	6.827
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.013)</b>
	5	6.362	5.621	8.310	9.003	8.749	5.760	4.776	5.595
		<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.039)</b>	<b>(0.013)</b>
	6.6	5.956	5.497	4.088	4.646	4.845	5.850	5.369	6.138
		<b>(0.013)</b>	<b>(0.013)</b>	(0.132)	<b>(0.092)</b>	<b>(0.079)</b>	<b>(0.026)</b>	<b>(0.039)</b>	<b>(0.013)</b>
7	5.373	4.912	6.575	7.626	6.822	5.066	3.970	4.945	
	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.026)</b>	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.026)</b>	<b>(0.079)</b>	<b>(0.039)</b>	
8	4.932	3.848	2.570	2.766	2.709	4.948	3.952	4.799	
	<b>(0.026)</b>	<b>(0.066)</b>	(0.276)	(0.289)	(0.303)	<b>(0.026)</b>	<b>(0.079)</b>	<b>(0.026)</b>	
12	4.524	3.612	3.251	2.881	2.731	4.766	3.628	3.579	
	<b>(0.000)</b>	<b>(0.066)</b>	(0.224)	(0.276)	(0.250)	<b>(0.013)</b>	<b>(0.053)</b>	<b>(0.039)</b>	
IQ Factor	2 to 5	0.744	0.723	1.017	1.113	1.055	0.662	0.625	0.670
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
	6 to 12	0.403	0.377	0.373	0.434	0.400	0.411	0.355	0.411
	<b>(0.026)</b>	<b>(0.039)</b>	(0.105)	<b>(0.079)</b>	(0.105)	<b>(0.026)</b>	<b>(0.066)</b>	<b>(0.039)</b>	

Table C.11: Treatment Effects on Achievement Scores, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. Achv. Test	5.5	8.029	8.366	14.284	16.754	15.265	6.223	5.634	7.268
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.066)</b>	<b>(0.000)</b>
	6	4.543	5.026	6.178	7.971	7.124	4.075	4.167	4.722
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	7.5	1.937	2.136	0.667	3.618	0.558	2.308	1.762	2.129
		(0.118)	<b>(0.066)</b>	(0.447)	(0.132)	(0.447)	(0.118)	(0.132)	(0.132)
8	4.207	4.715	1.630	5.607	2.393	4.958	4.637	5.508	
	<b>(0.000)</b>	<b>(0.000)</b>	(0.342)	<b>(0.039)</b>	(0.224)	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.013)</b>	
8.5	5.938	6.119	5.046	7.926	5.390	5.507	5.008	5.828	
	<b>(0.000)</b>	<b>(0.000)</b>	(0.118)	<b>(0.066)</b>	(0.105)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
Achievement Factor	5.5 to 12	0.483	0.540	0.570	0.819	0.620	0.456	0.458	0.496
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.053)</b>	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	

Table C.12: Treatment Effects on HOME Scores, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOME Score	0.5	1.005 (0.132)	0.168 (0.487)	1.332 (0.132)	0.523 (0.382)	0.773 (0.303)	0.566 (0.276)	0.071 (0.487)	0.195 (0.368)
	1.5	1.126 <b>(0.092)</b>	0.128 (0.474)	2.706 <b>(0.039)</b>	1.474 (0.158)	2.577 <b>(0.053)</b>	0.368 (0.329)	-0.098 (0.487)	0.435 (0.342)
	2.5	0.441 (0.355)	0.236 (0.382)	3.089 <b>(0.026)</b>	3.726 <b>(0.013)</b>	3.493 <b>(0.013)</b>	-0.588 (0.724)	-0.554 (0.697)	-0.051 (0.539)
	3.5	2.112 <b>(0.066)</b>	0.942 (0.276)	8.288 <b>(0.000)</b>	7.799 <b>(0.013)</b>	8.165 <b>(0.013)</b>	0.306 (0.342)	-0.492 (0.566)	0.325 (0.382)
	4.5	1.927 <b>(0.039)</b>	0.451 (0.408)	8.156 <b>(0.013)</b>	6.570 <b>(0.026)</b>	8.047 <b>(0.013)</b>	0.146 (0.421)	-0.674 (0.697)	0.339 (0.395)
	8	1.004 (0.263)	0.502 (0.329)	3.102 (0.105)	0.991 (0.329)	3.624 <b>(0.079)</b>	0.492 (0.382)	0.359 (0.355)	1.303 (0.224)
	HOME Factor	0.5 to 8	0.263 (0.105)	0.087 (0.395)	0.780 <b>(0.013)</b>	0.545 <b>(0.092)</b>	0.783 <b>(0.013)</b>	0.134 (0.197)	-0.018 (0.539)

Table C.13: Treatment Effects on Parent Income, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parental Income	1.5	2,248 (0.132)	3,277 <b>(0.066)</b>	2,553 (0.211)	4,721 (0.118)	5,028 <b>(0.066)</b>	1,870 (0.184)	2,901 (0.105)	3,718 <b>(0.026)</b>
	2.5	516 (0.355)	366 (0.500)	-2,455 (0.763)	-851 (0.500)	93.152 (0.461)	988 (0.329)	469 (0.447)	1,555 (0.237)
	3.5	1,821 (0.184)	1,901 (0.237)	3,984 <b>(0.066)</b>	4,990 <b>(0.092)</b>	5,265 <b>(0.053)</b>	961 (0.289)	961 (0.329)	2,108 (0.197)
	4.5	2,336 (0.105)	3,565 <b>(0.092)</b>	4,159 <b>(0.066)</b>	4,738 <b>(0.079)</b>	2,703 (0.197)	1,434 (0.263)	2,926 (0.132)	2,697 (0.158)
Parental Income Factor	1.5 to 15	0.156 (0.184)	0.120 (0.329)	0.158 (0.355)	0.303 (0.184)	0.207 (0.224)	0.127 (0.224)	0.063 (0.434)	0.183 (0.211)

Table C.14: Treatment Effects on Mother's Employment, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother Works	2	0.114 <b>(0.039)</b>	0.076 (0.158)	0.296 <b>(0.000)</b>	0.254 <b>(0.026)</b>	0.277 <b>(0.026)</b>	0.048 (0.237)	0.034 (0.316)	0.039 (0.289)
	3	0.119 <b>(0.039)</b>	0.087 (0.145)	0.219 <b>(0.053)</b>	0.172 (0.105)	0.199 <b>(0.079)</b>	0.092 <b>(0.092)</b>	0.065 (0.224)	0.087 (0.145)
	4	0.127 <b>(0.013)</b>	0.109 <b>(0.079)</b>	0.306 <b>(0.000)</b>	0.275 <b>(0.013)</b>	0.290 <b>(0.013)</b>	0.076 (0.105)	0.071 (0.145)	0.071 (0.132)
	5	0.089 <b>(0.066)</b>	0.059 (0.197)	0.342 <b>(0.000)</b>	0.321 <b>(0.013)</b>	0.346 <b>(0.000)</b>	0.005 (0.513)	-0.025 (0.605)	0.016 (0.434)
Mother Works Factor	2 to 21	0.256 (0.105)	0.179 (0.237)	0.717 <b>(0.000)</b>	0.641 <b>(0.026)</b>	0.679 <b>(0.026)</b>	0.133 (0.237)	0.065 (0.421)	0.126 (0.263)

Table C.15: Treatment Effects on Father at Home, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father at Home	2	-0.010 (0.553)	0.063 (0.197)	-0.187 (0.921)	-0.116 (0.776)	-0.148 (0.895)	0.047 (0.316)	0.142 <b>(0.026)</b>	0.130 <b>(0.066)</b>
	3	-0.076 (0.842)	-0.015 (0.605)	-0.291 (0.987)	-0.226 (0.961)	-0.258 (0.974)	0.002 (0.447)	0.066 (0.158)	0.079 (0.158)
	4	-0.071 (0.842)	-0.009 (0.566)	-0.331 (1.000)	-0.252 (0.974)	-0.293 (0.987)	0.021 (0.395)	0.074 (0.184)	0.101 <b>(0.092)</b>
	5	-0.093 (0.868)	-0.026 (0.671)	-0.369 (0.987)	-0.249 (0.921)	-0.343 (0.987)	-0.006 (0.487)	0.042 (0.263)	0.062 (0.184)
	8	0.052 (0.197)	0.073 (0.184)	-0.124 (0.803)	-0.036 (0.645)	-0.113 (0.803)	0.113 <b>(0.079)</b>	0.115 <b>(0.066)</b>	0.134 <b>(0.066)</b>
Father at Home Factor	2 to 8	-0.173 (0.829)	-0.028 (0.592)	-0.805 (1.000)	-0.548 (0.934)	-0.748 (1.000)	0.033 (0.421)	0.152 (0.171)	0.207 (0.145)

Table C.16: Treatment Effects on Education, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduated High School	30	0.164 <b>(0.013)</b>	0.116 <b>(0.079)</b>	0.390 <b>(0.000)</b>	0.340 <b>(0.013)</b>	0.353 <b>(0.000)</b>	0.103 <b>(0.066)</b>	0.049 (0.382)	0.059 (0.303)
Attended Voc./Tech./Com. College	30	-0.091 (0.855)	-0.101 (0.868)		0.116 (0.211)	-0.044 (0.658)	-0.100 (0.882)	-0.171 (0.961)	-0.152 (0.921)
Graduated 4-year College	30	0.161 <b>(0.000)</b>	0.141 <b>(0.026)</b>	0.188 <b>(0.013)</b>	0.168 <b>(0.039)</b>	0.181 <b>(0.026)</b>	0.148 <b>(0.000)</b>	0.120 <b>(0.026)</b>	0.136 <b>(0.013)</b>
Years of Edu.	30	1.367 <b>(0.000)</b>	1.070 <b>(0.000)</b>	2.513 <b>(0.000)</b>	2.108 <b>(0.000)</b>	2.430 <b>(0.000)</b>	0.986 <b>(0.013)</b>	0.759 <b>(0.013)</b>	0.886 <b>(0.013)</b>
Education Factor	30	0.513 <b>(0.000)</b>	0.390 <b>(0.000)</b>	0.843 <b>(0.000)</b>	0.725 <b>(0.000)</b>	0.787 <b>(0.000)</b>	0.389 <b>(0.000)</b>	0.294 <b>(0.039)</b>	0.316 <b>(0.026)</b>

Table C.17: Treatment Effects on Employment and Income, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	30	0.125 <b>(0.039)</b>	0.132 <b>(0.026)</b>	0.164 (0.105)	0.128 (0.158)	0.205 <b>(0.053)</b>	0.111 <b>(0.066)</b>	0.148 <b>(0.039)</b>	0.163 <b>(0.000)</b>
Labor Income	21	167 (0.382)	-579 (0.566)	1,577 (0.289)	1,469 (0.395)	1,749 (0.368)	-429 (0.566)	-1,211 (0.697)	-1,527 (0.776)
	30	12,377 <b>(0.026)</b>	10,752 (0.158)	17,677 <b>(0.013)</b>	13,888 (0.132)	21,198 <b>(0.026)</b>	10,847 <b>(0.092)</b>	9,790 (0.132)	11,232 (0.145)
Public-Transfer Income	21	-728 (0.211)	-923 (0.158)	-247 (0.382)	-1,161 (0.171)	-1,625 (0.171)	-1,054 (0.118)	-824 (0.184)	-813 (0.197)
	30	-1,832 <b>(0.026)</b>	-1,005 (0.171)	-1,613 <b>(0.079)</b>	-1,480 (0.105)	-1,600 (0.118)	-1,483 <b>(0.079)</b>	-992 (0.145)	-1,680 <b>(0.079)</b>
Employment Factor	21 to 30	0.344 <b>(0.000)</b>	0.267 <b>(0.013)</b>	0.393 <b>(0.053)</b>	0.315 (0.132)	0.423 <b>(0.026)</b>	0.307 <b>(0.039)</b>	0.291 <b>(0.026)</b>	0.313 <b>(0.026)</b>

Table C.18: Treatment Effects on Crime, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Felony Arrests	Mid-30s	0.045 (0.566)	0.237 (0.737)	-0.132 (0.342)	0.227 (0.566)	0.205 (0.658)	0.112 (0.645)	0.224 (0.671)	0.185 (0.671)
Total Misdemeanor Arrests	Mid-30s	-0.689 <b>(0.039)</b>	-0.481 (0.197)	-1.445 <b>(0.079)</b>	-1.198 (0.171)	-1.277 (0.158)	-0.546 (0.105)	-0.293 (0.276)	-0.310 (0.211)
Total Years Incarcerated	30	0.168 (0.895)	0.187 (0.895)	0.280 (1.000)	0.325 (1.000)	0.350 (1.000)	0.153 (0.868)	0.176 (0.855)	0.215 (0.908)
Crime Factor	30 to Mid-30s	0.054 (0.632)	0.072 (0.684)	0.012 (0.408)	0.045 (0.474)	0.051 (0.500)	0.076 (0.645)	0.091 (0.711)	0.131 (0.750)

Table C.19: Treatment Effects on Tobacco, Drugs, Alcohol, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cig. Smoked per day last month	30	0.033 (0.461)	0.321 (0.618)	-0.826 (0.250)	-0.150 (0.434)	-0.802 (0.237)	0.434 (0.645)	0.576 (0.684)	0.434 (0.658)
Days drank alcohol last month	30	0.244 (0.645)	0.419 (0.618)	-0.156 (0.447)	-0.700 (0.342)	0.113 (0.526)	0.207 (0.566)	0.618 (0.645)	0.630 (0.658)
Days binge drank alcohol last month	30	0.085 (0.632)	0.412 (0.803)	-0.267 (0.395)	-0.218 (0.408)	-0.126 (0.434)	0.151 (0.632)	0.674 (0.882)	0.393 (0.776)
Self-reported drug user	Mid-30s	-0.142 <b>(0.026)</b>	-0.183 <b>(0.053)</b>	-0.253 <b>(0.013)</b>	-0.303 <b>(0.039)</b>	-0.274 <b>(0.026)</b>	-0.090 (0.197)	-0.122 (0.158)	-0.114 (0.118)
Substance Use Factor	30 to Mid-30s	0.187 (0.842)	0.213 (0.921)	0.335 (0.987)	0.330 (0.908)	0.339 (0.987)	0.149 (0.724)	0.201 (0.855)	0.202 (0.855)

Table C.20: Treatment Effects on Self-Reported Health, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported Health	30	-0.066 (0.316)	0.007 (0.487)	-0.369 (0.105)	-0.264 (0.211)	-0.330 (0.184)	-0.009 (0.434)	0.004 (0.553)	0.036 (0.553)
	Mid-30s	0.079 (0.684)	-0.019 (0.447)	-0.429 <b>(0.039)</b>	-0.442 <b>(0.066)</b>	-0.498 <b>(0.039)</b>	0.179 (0.789)	0.113 (0.658)	0.066 (0.697)
Self-reported Health Factor	30 to Mid-30s	-0.012 (0.434)	-0.052 (0.342)	-0.083 (0.303)	-0.093 (0.329)	-0.129 (0.224)	-0.012 (0.447)	-0.021 (0.474)	-0.051 (0.329)

Table C.21: Treatment Effects on Hypertension, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Systolic Blood Pressure (mm Hg)	Mid-30s	-5.625 <b>(0.092)</b>	-10.849 <b>(0.026)</b>	5.375 (0.882)	6.809 (0.803)	4.383 (0.829)	-9.438 <b>(0.066)</b>	-16.448 <b>(0.000)</b>	-15.882 <b>(0.013)</b>
Diastolic Blood Pressure (mm Hg)	Mid-30s	-5.312 <b>(0.039)</b>	-7.745 <b>(0.026)</b>	-1.424 (0.316)	0.519 (0.579)	-1.637 (0.303)	-7.219 <b>(0.026)</b>	-9.589 <b>(0.053)</b>	-10.696 <b>(0.000)</b>
Prehypertension	Mid-30s	-0.176 <b>(0.000)</b>	-0.181 <b>(0.053)</b>	-0.049 (0.382)	-0.055 (0.355)	-0.060 (0.342)	-0.240 <b>(0.000)</b>	-0.268 <b>(0.000)</b>	-0.280 <b>(0.000)</b>
Hypertension	Mid-30s	-0.036 (0.316)	-0.097 (0.158)	0.083 (0.724)	0.143 (0.763)	0.045 (0.658)	-0.083 (0.171)	-0.151 (0.118)	-0.187 <b>(0.066)</b>
Hypertension Factor	Mid-30s	-0.293 <b>(0.053)</b>	-0.475 <b>(0.039)</b>	0.089 (0.632)	0.185 (0.684)	0.046 (0.539)	-0.447 <b>(0.026)</b>	-0.670 <b>(0.000)</b>	-0.703 <b>(0.000)</b>

Table C.22: Treatment Effects on Cholesterol, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	3.872 <b>(0.092)</b>	4.790 <b>(0.053)</b>	5.806 <b>(0.013)</b>	3.305 (0.211)	5.518 <b>(0.092)</b>	2.964 (0.184)	5.364 <b>(0.039)</b>	4.456 <b>(0.053)</b>
Dyslipidemia	Mid-30s	0.013 (0.579)	-0.003 (0.500)	0.035 (0.553)	0.082 (0.671)	0.031 (0.566)	0.032 (0.645)	-0.009 (0.474)	0.010 (0.553)
Cholesterol Factor	Mid-30s	-0.097 (0.289)	-0.141 (0.158)	-0.127 (0.237)	-0.000 (0.513)	-0.124 (0.237)	-0.049 (0.395)	-0.165 (0.132)	-0.117 (0.276)

Table C.23: Treatment Effects on Diabetes, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hemoglobin Level (%)	Mid-30s	0.003 (0.461)	0.075 (0.684)	0.032 (0.645)	-0.093 (0.355)	0.073 (0.684)	-0.029 (0.395)	0.086 (0.658)	-0.002 (0.434)
Prediabetes	Mid-30s	0.004 (0.487)	-0.039 (0.408)	-0.040 (0.421)	-0.110 (0.289)	-0.097 (0.289)	0.004 (0.447)	-0.045 (0.368)	-0.054 (0.329)
Diabetes	Mid-30s	-0.002 (0.487)	0.015 (0.671)	0.043 (0.855)	-0.003 (0.447)	0.049 (0.882)	-0.015 (0.382)	0.020 (0.711)	-0.007 (0.421)
Diabetes Factor	Mid-30s	-0.001 (0.461)	0.060 (0.632)	0.079 (0.697)	-0.065 (0.368)	0.104 (0.803)	-0.038 (0.408)	0.073 (0.684)	-0.017 (0.434)

Table C.24: Treatment Effects on Vitamin D Deficiency, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vitamin D Deficiency	Mid-30s	-0.153 <b>(0.039)</b>	-0.069 (0.237)	-0.264 <b>(0.000)</b>	-0.085 (0.342)	-0.233 <b>(0.000)</b>	-0.118 (0.132)	-0.031 (0.408)	-0.078 (0.263)

Table C.25: Treatment Effects on Obesity, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measured BMI	Mid-30s	0.999 (0.750)	3.094 (0.921)	-0.202 (0.500)	1.254 (0.671)	0.710 (0.605)	1.072 (0.711)	3.548 (0.961)	1.827 (0.816)
Obesity	Mid-30s	-0.050 (0.329)	0.063 (0.724)	-0.256 <b>(0.000)</b>	-0.144 (0.118)	-0.144 (0.197)	-0.013 (0.461)	0.113 (0.803)	0.011 (0.513)
Severe Obesity	Mid-30s	-0.126 <b>(0.092)</b>	-0.018 (0.382)	-0.093 (0.289)	-0.069 (0.355)	-0.065 (0.303)	-0.147 (0.105)	-0.011 (0.382)	-0.107 (0.171)
Waist-hip Ratio	Mid-30s	-0.006 (0.382)	0.000 (0.474)	-0.037 (0.171)	-0.047 (0.197)	-0.031 (0.224)	0.003 (0.618)	0.012 (0.776)	0.006 (0.711)
Abdominal Obesity	Mid-30s	-0.091 (0.158)	-0.021 (0.487)	-0.230 <b>(0.013)</b>	-0.139 (0.105)	-0.157 (0.118)	-0.041 (0.342)	0.031 (0.671)	-0.030 (0.474)
Framingham Risk Score	Mid-30s	0.348 (0.711)	-0.399 (0.289)	0.948 (0.921)	0.299 (0.605)	0.902 (0.908)	0.351 (0.724)	-0.704 (0.145)	0.088 (0.500)
Obesity Factor	Mid-30s	-0.022 (0.434)	0.162 (0.789)	-0.268 (0.211)	-0.138 (0.303)	-0.134 (0.342)	0.033 (0.579)	0.250 (0.829)	0.077 (0.618)

Table C.26: Treatment Effects on Mental Health, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Somatization	21	-0.015 (0.461)	-0.068 (0.184)	-0.118 (0.211)	-0.240 <b>(0.066)</b>	-0.182 (0.105)	0.012 (0.539)	-0.046 (0.263)	-0.071 (0.171)
	34	-0.128 (0.171)	-0.119 (0.184)	-0.191 (0.250)	-0.202 (0.250)	-0.205 (0.237)	-0.122 (0.184)	-0.114 (0.197)	-0.123 (0.145)
Depression	21	-0.256 <b>(0.000)</b>	-0.207 <b>(0.039)</b>	-0.329 (0.105)	-0.373 <b>(0.066)</b>	-0.365 <b>(0.066)</b>	-0.184 <b>(0.079)</b>	-0.164 (0.132)	-0.202 <b>(0.039)</b>
	34	-0.165 (0.184)	-0.203 (0.132)	-0.061 (0.395)	0.001 (0.421)	-0.117 (0.289)	-0.211 (0.158)	-0.292 <b>(0.039)</b>	-0.236 <b>(0.079)</b>
Anxiety	21	-0.119 (0.132)	-0.110 (0.132)	-0.192 (0.158)	-0.344 <b>(0.053)</b>	-0.208 (0.158)	-0.077 (0.224)	-0.083 (0.237)	-0.102 (0.158)
	34	-0.256 <b>(0.039)</b>	-0.259 <b>(0.066)</b>	-0.224 (0.171)	-0.217 (0.237)	-0.262 (0.145)	-0.280 <b>(0.039)</b>	-0.286 <b>(0.026)</b>	-0.300 <b>(0.013)</b>
Hostility	21	-0.296 <b>(0.013)</b>	-0.279 <b>(0.013)</b>	-0.504 <b>(0.013)</b>	-0.666 <b>(0.026)</b>	-0.564 <b>(0.013)</b>	-0.191 <b>(0.079)</b>	-0.199 <b>(0.092)</b>	-0.229 <b>(0.039)</b>
	34	-0.145 (0.158)	-0.126 (0.118)	-0.090 (0.342)	-0.057 (0.355)	-0.129 (0.184)	-0.174 (0.118)	-0.144 (0.118)	-0.150 (0.118)
Global Severity Index	21	-0.147 <b>(0.066)</b>	-0.140 (0.118)	-0.221 <b>(0.053)</b>	-0.325 <b>(0.039)</b>	-0.258 <b>(0.039)</b>	-0.107 (0.132)	-0.107 (0.184)	-0.144 <b>(0.092)</b>
	34	-3.294 (0.118)	-3.487 (0.105)	-2.858 (0.276)	-2.506 (0.303)	-3.502 (0.197)	-3.674 <b>(0.039)</b>	-4.149 <b>(0.039)</b>	-3.952 <b>(0.066)</b>
BSI Factor	21 and 34	-0.559 <b>(0.000)</b>	-0.330 <b>(0.066)</b>	-0.600 <b>(0.066)</b>	-0.405 (0.197)	-0.485 (0.105)	-0.547 <b>(0.000)</b>	-0.371 <b>(0.079)</b>	-0.350 <b>(0.039)</b>

## C.4 Treatment Effects for Male Sample

Table C.27: Treatment Effects on IQ Scores, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. IQ Test	2	9.528	11.036	6.875	10.704	7.944	10.286	11.497	11.080
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.026)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3	13.410	14.873	13.896	17.254	15.474	13.271	14.229	14.302
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3.5	8.756	7.994	6.354	5.904	6.784	9.443	8.690	9.043
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.066)</b>	(0.105)	<b>(0.053)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	4	12.089	11.867	8.950	7.899	9.668	12.986	12.604	13.487
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.118)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	4.5	8.508	8.914	10.411	11.349	10.639	7.964	8.518	7.804
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	5	7.697	7.130	4.643	3.469	4.992	8.679	8.241	8.181
		<b>(0.000)</b>	<b>(0.000)</b>	(0.158)	(0.316)	(0.158)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	6.6	5.803	5.157	0.831	-1.794	1.479	5.916	5.542	5.886
		<b>(0.026)</b>	<b>(0.000)</b>	(0.500)	(0.579)	(0.474)	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.013)</b>
7	4.390	5.812	5.323	8.155	5.750	4.156	5.481	4.418	
	<b>(0.053)</b>	<b>(0.026)</b>	(0.237)	(0.132)	(0.145)	<b>(0.092)</b>	<b>(0.039)</b>	(0.105)	
8	4.160	2.644	-2.514	-4.445	-2.651	4.754	3.391	4.207	
	<b>(0.079)</b>	(0.211)	(0.645)	(0.684)	(0.632)	<b>(0.053)</b>	<b>(0.092)</b>	<b>(0.066)</b>	
12	0.686	-0.856	-0.343	-4.076	-1.018	0.943	-0.116	-0.794	
	(0.382)	(0.592)	(0.566)	(0.658)	(0.605)	(0.355)	(0.461)	(0.592)	
IQ Factor	2 to 5	0.831	0.845	0.708	0.766	0.759	0.867	0.870	0.874
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	
	6 to 12	0.268	0.227	0.308	0.251	0.304	0.258	0.245	0.226
	(0.105)	(0.184)	(0.237)	(0.303)	(0.263)	(0.105)	(0.105)	(0.184)	

Table C.28: Treatment Effects on Achievement Scores, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. Achv. Test	5.5	5.108	5.375	10.088	11.148	10.628	3.863	4.183	4.171
		<b>(0.026)</b>	<b>(0.092)</b>	<b>(0.039)</b>	(0.118)	<b>(0.053)</b>	(0.105)	(0.224)	(0.145)
	6	3.091	3.803	2.271	3.803	3.354	3.312	3.662	3.629
		<b>(0.079)</b>	<b>(0.066)</b>	(0.368)	(0.316)	(0.276)	<b>(0.053)</b>	<b>(0.039)</b>	<b>(0.039)</b>
	7.5	0.019	1.103	-2.767	-2.380	-3.284	0.799	1.762	0.215
		(0.513)	(0.316)	(0.697)	(0.632)	(0.724)	(0.382)	(0.184)	(0.474)
8	2.309	4.117	-3.386	-0.625	-2.713	3.903	5.198	4.240	
	(0.145)	<b>(0.039)</b>	(0.750)	(0.487)	(0.697)	<b>(0.053)</b>	<b>(0.000)</b>	<b>(0.039)</b>	
8.5	3.910	5.586	-1.771	0.775	-1.498	4.199	5.497	4.117	
	<b>(0.092)</b>	<b>(0.013)</b>	(0.618)	(0.447)	(0.579)	<b>(0.066)</b>	<b>(0.013)</b>	<b>(0.079)</b>	
Achievement Factor	5.5 to 12	0.250	0.330	0.030	0.055	0.042	0.308	0.367	0.288
		(0.145)	(0.118)	(0.513)	(0.461)	(0.513)	<b>(0.092)</b>	(0.105)	(0.132)

Table C.29: Treatment Effects on HOME Scores, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOME Score	0.5	0.372 (0.316)	0.039 (0.487)	0.944 (0.316)	-0.416 (0.618)	0.431 (0.421)	0.143 (0.447)	0.069 (0.474)	-0.080 (0.487)
	1.5	-0.500 (0.645)	-0.874 (0.750)	0.431 (0.447)	-1.123 (0.684)	0.216 (0.447)	-0.766 (0.711)	-0.865 (0.737)	-0.872 (0.724)
	2.5	0.141 (0.474)	0.886 (0.263)	1.654 (0.197)	3.806 (0.105)	2.232 (0.145)	-0.292 (0.618)	0.543 (0.395)	0.145 (0.487)
	3.5	1.404 (0.197)	0.668 (0.368)	2.897 (0.197)	1.917 (0.263)	2.835 (0.197)	0.962 (0.329)	0.762 (0.355)	0.743 (0.368)
	4.5	1.146 (0.263)	1.534 (0.237)	3.312 (0.184)	3.894 ( <b>0.066</b> )	2.807 (0.184)	0.527 (0.395)	1.279 (0.276)	0.232 (0.421)
	8	1.548 (0.197)	1.524 (0.224)	-0.898 (0.526)	-4.874 (0.789)	-0.503 (0.513)	2.062 (0.145)	2.245 (0.224)	2.634 (0.105)
	HOME Factor	0.5 to 8	0.266 (0.118)	0.180 (0.224)	0.130 (0.382)	-0.093 (0.592)	0.080 (0.408)	0.295 (0.132)	0.218 (0.211)

Table C.30: Treatment Effects on Parent Income, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parental Income	1.5	330 (0.408)	-97.199 (0.461)	-1,046 (0.579)	-2,384 (0.645)	-1,168 (0.632)	-9.245 (0.434)	-26.663 (0.487)	872 (0.329)
	2.5	673 (0.382)	-941 (0.645)	-1,167 (0.618)	-3,542 (0.750)	-1,858 (0.671)	478 (0.408)	-839 (0.605)	232 (0.513)
	3.5	1,036 (0.355)	223 (0.500)	3,085 (0.211)	-1,152 (0.553)	1,448 (0.395)	112 (0.461)	47.496 (0.447)	701 (0.461)
	4.5	821 (0.355)	1,677 (0.303)	1,561 (0.395)	3,815 (0.132)	-2,662 (0.803)	-81.743 (0.513)	917 (0.434)	-429 (0.461)
Parental Income Factor	1.5 to 15	0.130 (0.289)	-0.018 (0.513)	0.079 (0.461)	0.039 (0.421)	0.092 (0.461)	0.088 (0.316)	-0.062 (0.579)	0.116 (0.316)

Table C.31: Treatment Effects on Mother's Employment, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother Works	2	0.056 (0.289)	0.038 (0.368)	0.264 ( <b>0.066</b> )	0.184 (0.145)	0.240 ( <b>0.079</b> )	-0.004 (0.447)	0.008 (0.487)	-0.018 (0.539)
	3	0.150 ( <b>0.066</b> )	0.125 (0.184)	0.261 ( <b>0.079</b> )	0.184 (0.145)	0.240 ( <b>0.079</b> )	0.117 (0.158)	0.117 (0.197)	0.117 (0.171)
	4	0.134 ( <b>0.079</b> )	0.147 (0.105)	0.287 ( <b>0.053</b> )	0.262 (0.132)	0.270 ( <b>0.066</b> )	0.090 (0.184)	0.124 (0.158)	0.089 (0.197)
	5	0.111 ( <b>0.092</b> )	0.095 (0.211)	0.311 ( <b>0.053</b> )	0.259 ( <b>0.079</b> )	0.289 ( <b>0.066</b> )	0.061 (0.276)	0.071 (0.303)	0.054 (0.316)
Mother Works Factor	2 to 21	0.309 (0.118)	0.280 (0.211)	0.865 ( <b>0.066</b> )	0.730 (0.184)	0.805 ( <b>0.079</b> )	0.160 (0.237)	0.196 (0.250)	0.144 (0.250)

Table C.32: Treatment Effects on Father at Home, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father at Home	2	-0.018 (0.553)	0.160 ( <b>0.066</b> )	-0.282 (0.882)	0.111 (0.368)	-0.223 (0.816)	0.057 (0.355)	0.187 ( <b>0.039</b> )	0.171 ( <b>0.066</b> )
	3	-0.076 (0.737)	0.007 (0.447)	-0.243 (0.882)	0.026 (0.447)	-0.197 (0.816)	-0.029 (0.592)	0.010 (0.487)	0.070 (0.276)
	4	-0.075 (0.724)	0.008 (0.513)	-0.339 (0.974)	-0.050 (0.658)	-0.287 (0.961)		0.036 (0.395)	0.103 (0.224)
	5	-0.057 (0.697)	0.051 (0.303)	-0.429 (0.987)	-0.081 (0.724)	-0.376 (0.987)	0.036 (0.395)	0.089 (0.211)	0.142 (0.132)
	8	0.037 (0.355)	0.070 (0.303)	-0.177 (0.868)	0.058 (0.408)	-0.160 (0.776)	0.123 (0.132)	0.084 (0.289)	0.168 ( <b>0.092</b> )
Father at Home Factor	2 to 8	-0.160 (0.750)	0.034 (0.421)	-0.728 (0.961)	0.029 (0.553)	-0.617 (0.895)	0.039 (0.474)	0.073 (0.342)	0.308 (0.132)

Table C.33: Treatment Effects on Education, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduated High School	30	0.073 (0.237)	0.109 (0.250)	0.114 (0.342)	0.113 (0.355)	0.085 (0.382)	0.077 (0.237)	0.102 (0.289)	0.064 (0.329)
Attended Voc./Tech./Com. College	30	-0.099 (0.816)	-0.124 (0.816)	0.086 (0.395)	0.258 (0.145)	0.024 (0.474)	-0.138 (0.816)	-0.231 (0.947)	-0.233 (0.961)
Graduated 4-year College	30	0.170 <b>(0.066)</b>	0.221 <b>(0.066)</b>	0.124 (0.303)	0.323 <b>(0.079)</b>	0.118 (0.289)	0.179 <b>(0.013)</b>	0.210 <b>(0.066)</b>	0.170 <b>(0.079)</b>
Years of Edu.	30	0.525 <b>(0.079)</b>	0.708 <b>(0.079)</b>	0.857 (0.118)	1.302 <b>(0.092)</b>	0.791 (0.171)	0.385 (0.171)	0.540 (0.132)	0.347 (0.276)
Education Factor	30	0.338 <b>(0.039)</b>	0.437 <b>(0.079)</b>	0.298 (0.237)	0.722 <b>(0.026)</b>	0.249 (0.289)	0.346 <b>(0.066)</b>	0.385 <b>(0.092)</b>	0.288 (0.132)

Table C.34: Treatment Effects on Employment and Income, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	30	0.119 <b>(0.079)</b>	0.179 <b>(0.039)</b>	-0.029 (0.487)	-0.050 (0.579)	0.041 (0.355)	0.176 <b>(0.053)</b>	0.245 <b>(0.013)</b>	0.262 <b>(0.000)</b>
Labor Income	21	-1,672 (0.711)	-4,977 (0.895)	-3,951 (0.724)	-13,951 (0.974)	-4,585 (0.737)	-1,527 (0.724)	-3,973 (0.842)	-3,746 (0.855)
	30	19,810 <b>(0.079)</b>	24,902 (0.171)	17,909 (0.132)	21,069 (0.263)	24,012 (0.105)	20,065 <b>(0.066)</b>	28,483 (0.132)	21,170 (0.158)
Public-Transfer Income	21	315 (0.684)	456 (0.658)	1,376 (0.868)	2,654 (0.763)	1,543 (0.803)	-58,901 (0.513)	144 (0.618)	97,591 (0.645)
	30	-530 (0.118)	-176 (0.355)	287 (0.539)	722 (0.539)	548 (0.579)	-279 (0.237)	-82,612 (0.395)	-155 (0.355)
Employment Factor	21 to 30	0.351 <b>(0.026)</b>	0.433 <b>(0.053)</b>	0.053 (0.408)	-0.119 (0.539)	0.086 (0.382)	0.450 <b>(0.000)</b>	0.595 <b>(0.013)</b>	0.486 <b>(0.000)</b>

Table C.35: Treatment Effects on Crime, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Felony Arrests	Mid-30s	0.196 (0.618)	0.683 (0.711)	0.946 (0.934)	2.132 (0.895)	1.337 (0.974)	0.017 (0.487)	0.466 (0.645)	0.183 (0.553)
Total Misdemeanor Arrests	Mid-30s	-0.501 (0.132)	-0.239 (0.316)	-0.251 (0.408)	0.085 (0.500)	-0.040 (0.395)	-0.665 (0.105)	-0.343 (0.224)	-0.512 (0.118)
Total Years Incarcerated	30	0.347 (0.934)	0.489 (0.934)	0.553 (1.000)	0.811 (0.974)	0.701 (1.000)	0.338 (0.882)	0.489 (0.934)	0.470 (0.947)
Crime Factor	30 to Mid-30s	0.169 (0.776)	0.326 (0.776)	0.469 (1.000)	0.835 (0.921)	0.545 (1.000)	0.109 (0.658)	0.272 (0.684)	0.182 (0.737)

Table C.36: Treatment Effects on Tobacco, Drugs, Alcohol, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cig. Smoked per day last month	30	0.826 (0.789)	0.245 (0.539)	0.757 (0.618)	-2.228 (0.184)	0.624 (0.579)	1.429 (0.908)	1.083 (0.789)	1.215 (0.789)
Days drank alcohol last month	30	0.805 (0.711)	1.083 (0.645)	-0.186 (0.408)	-0.648 (0.355)	0.060 (0.421)	0.944 (0.711)	1.348 (0.724)	1.341 (0.776)
Days binge drank alcohol last month	30	0.500 (0.816)	0.474 (0.816)	0.543 (0.724)	-0.360 (0.368)	0.692 (0.829)	0.491 (0.789)	0.625 (0.803)	0.702 (0.868)
Self-reported drug user	Mid-30s	-0.333 <b>(0.026)</b>	-0.432 <b>(0.000)</b>	-0.500 <b>(0.000)</b>	-0.788 <b>(0.000)</b>	-0.555 <b>(0.000)</b>	-0.233 (0.118)	-0.326 <b>(0.053)</b>	-0.330 <b>(0.039)</b>
Substance Use Factor	30 to Mid-30s	0.337 (0.908)	0.232 (0.737)	0.180 (0.553)	0.014 (0.395)	0.219 (0.566)	0.403 (0.934)	0.273 (0.724)	0.292 (0.855)

Table C.37: Treatment Effects on Self-Reported Health, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported Health	30	-0.004 (0.539)	0.226 (0.789)	0.457 (0.868)	0.755 (0.934)	0.568 (0.961)	-0.145 (0.303)	0.096 (0.632)	-0.000 (0.500)
	Mid-30s	0.196 (0.684)	0.010 (0.539)	-0.852 (0.000)	-1.130 (0.026)	-0.925 (0.000)	0.382 (0.934)	0.207 (0.697)	0.170 (0.711)
Self-reported Health Factor	30 to Mid-30s	-0.035 (0.368)	-0.136 (0.224)	-0.563 (0.039)	-0.758 (0.026)	-0.596 (0.039)	0.065 (0.684)	-0.035 (0.395)	-0.044 (0.355)

Table C.38: Treatment Effects on Hypertension, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Systolic Blood Pressure (mm Hg)	Mid-30s	-9.791 (0.079)	-20.743 (0.000)	15.280 (0.908)	16.605 (0.855)	14.253 (0.908)	-19.920 (0.013)	-30.437 (0.000)	-30.587 (0.000)
Diastolic Blood Pressure (mm Hg)	Mid-30s	-10.854 (0.000)	-19.895 (0.000)	-8.640 (0.013)	-12.199 (0.079)	-8.150 (0.026)	-14.240 (0.013)	-22.740 (0.000)	-21.851 (0.000)
Prehypertension	Mid-30s	-0.137 (0.145)	-0.164 (0.184)	0.053 (0.474)	0.192 (0.645)	0.108 (0.592)	-0.280 (0.000)	-0.295 (0.066)	-0.296 (0.000)
Hypertension	Mid-30s	-0.291 (0.039)	-0.383 (0.013)	-0.053 (0.395)	0.090 (0.461)	-0.064 (0.368)	-0.420 (0.000)	-0.511 (0.000)	-0.517 (0.000)
Hypertension Factor	Mid-30s	-0.599 (0.000)	-1.063 (0.000)	0.074 (0.539)	0.112 (0.526)	0.076 (0.539)	-0.961 (0.000)	-1.401 (0.000)	-1.386 (0.000)

Table C.39: Treatment Effects on Cholesterol, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	7.753 (0.013)	4.139 (0.197)	-0.267 (0.461)	-8.079 (0.763)	-4.138 (0.645)	9.015 (0.013)	5.332 (0.118)	6.830 (0.066)
Dyslipidemia	Mid-30s	-0.094 (0.289)	-0.029 (0.421)	0.200 (0.868)	0.365 (0.789)	0.207 (0.855)	-0.108 (0.289)	-0.056 (0.342)	-0.130 (0.237)
Cholesterol Factor	Mid-30s	-0.333 (0.092)	-0.153 (0.316)	0.240 (0.750)	0.657 (0.882)	0.359 (0.816)	-0.385 (0.105)	-0.219 (0.303)	-0.348 (0.145)

Table C.40: Treatment Effects on Diabetes, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hemoglobin Level (%)	Mid-30s	0.322 (0.829)	0.435 (0.737)	0.240 (0.658)	0.469 (0.618)	0.355 (0.724)	0.286 (0.803)	0.386 (0.763)	0.396 (0.829)
Prediabetes	Mid-30s	-0.129 (0.250)	-0.190 (0.184)	-0.267 (0.197)	-0.478 (0.118)	-0.338 (0.132)	-0.139 (0.237)	-0.182 (0.224)	-0.173 (0.224)
Diabetes	Mid-30s	0.080 (0.842)	0.081 (0.776)	0.080 (0.750)	0.027 (0.461)	0.098 (0.789)	0.080 (0.842)	0.083 (0.803)	0.098 (0.882)
Diabetes Factor	Mid-30s	0.283 (0.908)	0.338 (0.763)	0.225 (0.711)	0.244 (0.579)	0.307 (0.737)	0.262 (0.882)	0.317 (0.789)	0.345 (0.882)

Table C.41: Treatment Effects on Vitamin D Deficiency, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vitamin D Deficiency	Mid-30s	-0.245 (0.066)	-0.185 (0.132)	-0.480 (0.000)	-0.216 (0.158)	-0.485 (0.000)	-0.172 (0.158)	-0.145 (0.263)	-0.189 (0.158)

Table C.42: Treatment Effects on Obesity, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measured BMI	Mid-30s	-0.125 (0.500)	-0.617 (0.408)	-0.684 (0.382)	-3.505 (0.303)	0.867 (0.447)	-0.627 (0.447)	0.025 (0.539)	-0.487 (0.421)
Obesity	Mid-30s		-0.103 (0.303)	-0.128 (0.329)	-0.331 <b>(0.079)</b>	0.031 (0.487)	-0.017 (0.487)	-0.045 (0.395)	-0.060 (0.395)
Severe Obesity	Mid-30s	-0.160 (0.132)	-0.132 (0.211)	-0.185 (0.224)	-0.319 (0.145)	-0.126 (0.276)	-0.185 (0.171)	-0.108 (0.289)	-0.131 (0.250)
Waist-hip Ratio	Mid-30s	0.004 (0.553)	-0.014 (0.237)	0.018 (0.526)	-0.043 (0.171)	0.030 (0.553)	-0.002 (0.461)	-0.008 (0.329)	-0.004 (0.421)
Abdominal Obesity	Mid-30s	0.003 (0.461)	-0.137 (0.211)	0.029 (0.487)	-0.377 (0.171)	0.100 (0.566)	0.029 (0.605)	-0.062 (0.382)	-0.031 (0.421)
Framingham Risk Score	Mid-30s	-0.766 (0.224)	-0.472 (0.368)	1.491 (0.737)	2.437 (0.684)	1.802 (0.816)	-1.202 (0.184)	-1.036 (0.250)	-0.705 (0.316)
Obesity Factor	Mid-30s	-0.040 (0.461)	-0.197 (0.263)	-0.036 (0.408)	-0.524 (0.211)	0.211 (0.474)	-0.101 (0.395)	-0.091 (0.434)	-0.116 (0.355)

Table C.43: Treatment Effects on Mental Health, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Somatization	21	0.035 (0.671)	-0.071 (0.276)	0.076 (0.697)	-0.196 (0.132)	0.035 (0.605)	0.032 (0.645)	-0.091 (0.276)	-0.043 (0.395)
	34	-0.239 (0.184)	-0.333 <b>(0.092)</b>	-0.025 (0.329)	-0.110 (0.171)	-0.075 (0.197)	-0.341 (0.105)	-0.370 (0.105)	-0.331 (0.118)
Depression	21	-0.023 (0.474)	-0.085 (0.276)	0.239 (0.855)	0.171 (0.789)	0.174 (0.697)	-0.024 (0.461)	-0.113 (0.197)	-0.074 (0.289)
	34	-0.145 (0.316)	-0.247 (0.250)	0.216 (0.868)	0.370 (0.803)	0.151 (0.803)	-0.295 (0.263)	-0.373 (0.145)	-0.273 (0.250)
Anxiety	21	0.145 (0.868)	0.008 (0.579)	0.307 (0.961)	0.012 (0.513)	0.249 (0.947)	0.105 (0.816)	-0.042 (0.368)	0.019 (0.566)
	34	-0.191 (0.224)	-0.293 <b>(0.079)</b>	0.099 (0.789)	0.117 (0.803)	0.040 (0.697)	-0.312 (0.145)	-0.364 <b>(0.092)</b>	-0.301 (0.145)
Hostility	21	-0.099 (0.303)	-0.148 (0.237)	0.126 (0.724)	-0.129 (0.355)	0.081 (0.645)	-0.134 (0.289)	-0.151 (0.197)	-0.144 (0.237)
	34	-0.157 (0.250)	-0.264 (0.145)	0.067 (0.697)	0.019 (0.553)	-0.007 (0.408)	-0.260 (0.184)	-0.322 (0.145)	-0.273 (0.158)
Global Severity Index	21	0.045 (0.724)	-0.061 (0.329)	0.229 (0.934)	-0.038 (0.408)	0.169 (0.908)	0.016 (0.632)	-0.089 (0.263)	-0.058 (0.276)
	34	-3.450 (0.237)	-5.240 (0.158)	1.741 (0.776)	2.260 (0.750)	0.696 (0.658)	-5.693 (0.171)	-6.643 (0.105)	-5.427 (0.158)
BSI Factor	21 and 34	-0.186 (0.303)	-0.150 (0.408)	0.016 (0.539)	0.328 (0.671)	0.042 (0.592)	-0.261 (0.289)	-0.237 (0.303)	-0.109 (0.447)

## C.5 Treatment Effects for Female Sample

Table C.44: Treatment Effects on IQ Scores, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. IQ Test	2	10.700	9.200	13.950	16.963	13.910	9.431	6.680	9.341
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>
	3	13.333	12.563	23.729	29.326	24.246	9.211	7.275	9.177
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.013)</b>
	3.5	8.049	6.310	16.188	20.843	16.047	5.049	1.970	4.960
		<b>(0.000)</b>	<b>(0.053)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.092)</b>	<b>(0.303)</b>	<b>(0.066)</b>
	4	6.035	4.483	14.812	19.509	15.321	3.007	0.543	3.475
		<b>(0.013)</b>	<b>(0.092)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.158)	(0.474)	(0.132)
	4.5	8.162	6.607	16.058	20.187	16.348	5.318	1.706	5.810
		<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.079)</b>	<b>(0.355)</b>	<b>(0.066)</b>
	5	4.921	3.762	12.425	17.533	12.947	2.698	-0.158	2.994
		<b>(0.092)</b>	(0.197)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.211)	(0.526)	(0.184)
	6.6	6.127	5.322	7.339	12.597	8.174	5.773	3.462	6.408
		<b>(0.053)</b>	<b>(0.039)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.053)</b>	<b>(0.237)</b>	<b>(0.053)</b>
7	6.365	4.279	7.796	13.116	7.861	5.992	2.304	5.487	
	<b>(0.039)</b>	(0.105)	<b>(0.066)</b>	<b>(0.000)</b>	<b>(0.066)</b>	<b>(0.066)</b>	(0.289)	<b>(0.079)</b>	
8	5.906	4.111	7.857	12.487	8.186	5.360	2.282	5.550	
	<b>(0.026)</b>	<b>(0.092)</b>	<b>(0.039)</b>	<b>(0.000)</b>	<b>(0.053)</b>	<b>(0.066)</b>	(0.276)	<b>(0.066)</b>	
12	8.688	7.857	6.850	9.960	6.441	9.120	6.952	8.429	
	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.013)</b>	
IQ Factor	2 to 5	0.644	0.529	1.317	1.705	1.345	0.437	0.167	0.457
		<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.053)</b>	(0.303)	<b>(0.053)</b>
	6 to 12	0.545	0.475	0.448	0.976	0.506	0.570	0.362	0.601
		<b>(0.053)</b>	<b>(0.026)</b>	<b>(0.053)</b>	<b>(0.000)</b>	<b>(0.053)</b>	<b>(0.039)</b>	(0.132)	<b>(0.039)</b>

Table C.45: Treatment Effects on Achievement Scores, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. Achv. Test	5.5	12.314	10.088	19.650	19.146	20.936	9.869	4.876	11.584
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.026)</b>	(0.171)	<b>(0.000)</b>
	6	6.269	6.756	10.379	14.081	10.986	5.018	4.220	5.887
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.000)</b>
	7.5	4.133	3.185	4.300	9.725	4.578	4.082	1.201	4.326
		<b>(0.092)</b>	<b>(0.092)</b>	(0.171)	<b>(0.013)</b>	(0.171)	(0.105)	(0.329)	(0.105)
8	6.619	5.350	7.125	11.989	7.738	6.465	3.618	7.036	
	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.053)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.039)</b>	(0.171)	<b>(0.039)</b>	
8.5	8.407	7.607	12.299	15.146	12.625	7.223	5.279	7.918	
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.066)</b>	<b>(0.000)</b>	
Achievement Factor	5.5 to 12	0.839	0.838	1.183	1.277	1.264	0.706	0.648	0.809
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.000)</b>

Table C.46: Treatment Effects on HOME Scores, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOME Score	0.5	1.581 <b>(0.039)</b>	0.749 <b>(0.250)</b>	1.684 <b>(0.158)</b>	1.131 <b>(0.250)</b>	1.067 <b>(0.316)</b>	0.980 <b>(0.224)</b>	0.498 <b>(0.342)</b>	0.438 <b>(0.316)</b>
	1.5	2.668 <b>(0.013)</b>	1.723 <b>(0.158)</b>	4.729 <b>(0.013)</b>	3.279 <b>(0.092)</b>	4.661 <b>(0.013)</b>	1.544 <b>(0.132)</b>	0.992 <b>(0.263)</b>	1.746 <b>(0.105)</b>
	2.5	0.762 <b>(0.224)</b>	0.832 <b>(0.237)</b>	4.434 <b>(0.000)</b>	5.502 <b>(0.000)</b>	4.637 <b>(0.000)</b>	-0.899 <b>(0.737)</b>	-0.932 <b>(0.737)</b>	-0.256 <b>(0.566)</b>
	3.5	2.858 <b>(0.158)</b>	2.561 <b>(0.132)</b>	13.719 <b>(0.000)</b>	15.918 <b>(0.026)</b>	13.454 <b>(0.000)</b>	-0.309 <b>(0.579)</b>	-1.851 <b>(0.737)</b>	-0.058 <b>(0.500)</b>
	4.5	2.736 <b>(0.132)</b>	1.288 <b>(0.329)</b>	12.957 <b>(0.013)</b>	11.926 <b>(0.026)</b>	13.245 <b>(0.013)</b>	-0.273 <b>(0.553)</b>	-1.462 <b>(0.737)</b>	0.468 <b>(0.421)</b>
	8	0.659 <b>(0.434)</b>	0.721 <b>(0.329)</b>	5.909 <b>(0.039)</b>	6.478 <b>(0.026)</b>	6.484 <b>(0.039)</b>	-0.773 <b>(0.658)</b>	-1.027 <b>(0.671)</b>	0.328 <b>(0.434)</b>
	HOME Factor	0.5 to 8	0.262 <b>(0.158)</b>	0.220 <b>(0.184)</b>	1.215 <b>(0.013)</b>	1.001 <b>(0.118)</b>	1.272 <b>(0.013)</b>	-0.010 <b>(0.553)</b>	-0.053 <b>(0.618)</b>

Table C.47: Treatment Effects on Parent Income, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parental Income	1.5	4,516 <b>(0.066)</b>	7,539 <b>(0.013)</b>	5,105 <b>(0.158)</b>	8,682 <b>(0.079)</b>	9,681 <b>(0.039)</b>	4,309 <b>(0.105)</b>	7,091 <b>(0.066)</b>	7,337 <b>(0.000)</b>
	2.5	222 <b>(0.474)</b>	400 <b>(0.434)</b>	-3,712 <b>(0.697)</b>	894 <b>(0.408)</b>	1,769 <b>(0.408)</b>	1,598 <b>(0.355)</b>	894 <b>(0.382)</b>	3,227 <b>(0.224)</b>
	3.5	2,756 <b>(0.211)</b>	4,043 <b>(0.184)</b>	4,469 <b>(0.184)</b>	7,429 <b>(0.118)</b>	8,577 <b>(0.066)</b>	2,125 <b>(0.276)</b>	2,120 <b>(0.316)</b>	3,764 <b>(0.211)</b>
	4.5	4,039 <b>(0.092)</b>	8,141 <b>(0.026)</b>	6,443 <b>(0.105)</b>	9,836 <b>(0.079)</b>	9,337 <b>(0.013)</b>	3,202 <b>(0.118)</b>	7,852 <b>(0.039)</b>	6,022 <b>(0.013)</b>
Parental Income Factor	1.5 to 15	0.193 <b>(0.237)</b>	0.209 <b>(0.276)</b>	0.192 <b>(0.316)</b>	0.291 <b>(0.211)</b>	0.284 <b>(0.276)</b>	0.193 <b>(0.276)</b>	0.163 <b>(0.342)</b>	0.283 <b>(0.184)</b>

Table C.48: Treatment Effects on Mother's Employment, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother Works	2	0.168 <b>(0.026)</b>	0.124 <b>(0.079)</b>	0.323 <b>(0.079)</b>	0.292 <b>(0.145)</b>	0.309 <b>(0.079)</b>	0.101 <b>(0.145)</b>	0.082 <b>(0.211)</b>	0.096 <b>(0.171)</b>
	3	0.087 <b>(0.237)</b>	0.047 <b>(0.368)</b>	0.177 <b>(0.158)</b>	0.145 <b>(0.316)</b>	0.160 <b>(0.184)</b>	0.066 <b>(0.276)</b>	0.021 <b>(0.461)</b>	0.057 <b>(0.276)</b>
	4	0.118 <b>(0.039)</b>	0.080 <b>(0.197)</b>	0.319 <b>(0.079)</b>	0.275 <b>(0.171)</b>	0.305 <b>(0.079)</b>	0.060 <b>(0.276)</b>	0.038 <b>(0.368)</b>	0.053 <b>(0.316)</b>
	5	0.067 <b>(0.263)</b>	0.017 <b>(0.434)</b>	0.367 <b>(0.000)</b>	0.268 <b>(0.066)</b>	0.397 <b>(0.000)</b>	-0.056 <b>(0.658)</b>	-0.097 <b>(0.895)</b>	-0.024 <b>(0.605)</b>
Mother Works Factor	2 to 21	0.200 <b>(0.211)</b>	0.076 <b>(0.434)</b>	0.578 <b>(0.132)</b>	0.484 <b>(0.197)</b>	0.564 <b>(0.145)</b>	0.101 <b>(0.329)</b>	-0.037 <b>(0.553)</b>	0.108 <b>(0.316)</b>

Table C.49: Treatment Effects on Father at Home, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father at Home	2	-0.012 <b>(0.579)</b>	0.015 <b>(0.500)</b>	-0.115 <b>(0.737)</b>	-0.166 <b>(0.816)</b>	-0.094 <b>(0.737)</b>	0.034 <b>(0.474)</b>	0.095 <b>(0.263)</b>	0.087 <b>(0.224)</b>
	3	-0.079 <b>(0.789)</b>	-0.038 <b>(0.671)</b>	-0.337 <b>(0.961)</b>	-0.330 <b>(0.974)</b>	-0.315 <b>(0.947)</b>	0.034 <b>(0.474)</b>	0.095 <b>(0.263)</b>	0.087 <b>(0.224)</b>
	4	-0.071 <b>(0.789)</b>	-0.042 <b>(0.671)</b>	-0.330 <b>(0.947)</b>	-0.350 <b>(0.961)</b>	-0.306 <b>(0.947)</b>	0.041 <b>(0.447)</b>	0.091 <b>(0.263)</b>	0.096 <b>(0.224)</b>
	5	-0.139 <b>(0.908)</b>	-0.101 <b>(0.895)</b>	-0.333 <b>(0.934)</b>	-0.278 <b>(0.921)</b>	-0.328 <b>(0.934)</b>	-0.056 <b>(0.724)</b>	-0.025 <b>(0.618)</b>	-0.020 <b>(0.592)</b>
	8	0.056 <b>(0.303)</b>	0.078 <b>(0.211)</b>	-0.064 <b>(0.605)</b>	0.003 <b>(0.434)</b>	-0.055 <b>(0.605)</b>	0.092 <b>(0.197)</b>	0.125 <b>(0.145)</b>	0.095 <b>(0.184)</b>
Father at Home Factor	2 to 8	-0.211 <b>(0.829)</b>	-0.094 <b>(0.658)</b>	-0.899 <b>(0.974)</b>	-0.818 <b>(0.961)</b>	-0.884 <b>(0.974)</b>	-0.002 <b>(0.539)</b>	0.143 <b>(0.289)</b>	0.090 <b>(0.421)</b>

Table C.50: Treatment Effects on Education, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduated High School	30	0.253 <b>(0.000)</b>	0.148 <b>(0.066)</b>	0.642 <b>(0.000)</b>	0.519 <b>(0.013)</b>	0.595 <b>(0.000)</b>	0.137 (0.158)	0.004 (0.461)	0.066 (0.263)
Attended Voc./Tech./Com. College	30	-0.057 (0.684)	-0.115 (0.803)	-0.050 (0.618)	-0.098 (0.671)	-0.071 (0.697)	-0.041 (0.618)	-0.145 (0.895)	-0.051 (0.724)
Graduated 4-year College	30	0.134 <b>(0.066)</b>	0.102 (0.158)	0.217 <b>(0.013)</b>	0.097 (0.184)	0.210 <b>(0.000)</b>	0.106 (0.145)	0.073 (0.316)	0.095 (0.197)
Years of Edu.	30	2.143 <b>(0.000)</b>	1.695 <b>(0.000)</b>	4.025 <b>(0.000)</b>	2.984 <b>(0.000)</b>	3.918 <b>(0.000)</b>	1.567 <b>(0.013)</b>	1.155 <b>(0.066)</b>	1.409 <b>(0.026)</b>
Education Factor	30	0.661 <b>(0.000)</b>	0.461 <b>(0.039)</b>	1.277 <b>(0.000)</b>	0.799 <b>(0.039)</b>	1.212 <b>(0.000)</b>	0.447 <b>(0.039)</b>	0.261 (0.250)	0.361 (0.105)

Table C.51: Treatment Effects on Employment and Income, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	30	0.131 <b>(0.092)</b>	0.080 (0.171)	0.333 <b>(0.053)</b>	0.363 <b>(0.092)</b>	0.340 <b>(0.053)</b>	0.056 (0.276)	0.003 (0.447)	0.070 (0.263)
Labor Income	21	1,741 (0.224)	1,694 (0.263)	6,932 <b>(0.000)</b>	10,072 <b>(0.000)</b>	8,248 <b>(0.000)</b>	496 (0.434)	-779 (0.579)	440 (0.461)
	30	2,548 (0.368)	2,391 (0.355)	14,356 <b>(0.000)</b>	8,166 (0.171)	14,835 <b>(0.013)</b>	-425 (0.566)	-803 (0.579)	-840 (0.566)
Public-Transfer Income	21	-1,424 <b>(0.092)</b>	-1,929 <b>(0.013)</b>	-1,322 (0.158)	-1,786 (0.132)	-2,879 <b>(0.039)</b>	-1,751 <b>(0.066)</b>	-1,144 (0.158)	-1,475 (0.118)
	30	-2,672 <b>(0.026)</b>	-1,560 (0.118)	-3,053 <b>(0.013)</b>	-2,783 <b>(0.092)</b>	-3,213 <b>(0.013)</b>	-2,269 (0.105)	-1,169 (0.224)	-2,620 (0.132)
Employment Factor	21 to 30	0.281 <b>(0.079)</b>	0.132 (0.211)	0.658 <b>(0.013)</b>	0.557 <b>(0.079)</b>	0.675 <b>(0.000)</b>	0.167 (0.263)	0.000 (0.487)	0.149 (0.276)

Table C.52: Treatment Effects on Crime, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Felony Arrests	Mid-30s	-0.328 (0.105)	-0.390 (0.145)	-1.345 (0.105)	-1.011 (0.171)	-0.968 (0.118)	-0.077 (0.237)	-0.022 (0.395)	0.005 (0.500)
Total Misdemeanor Arrests	Mid-30s	-0.973 <b>(0.039)</b>	-0.883 (0.145)	-2.708 <b>(0.079)</b>	-2.177 (0.184)	-2.453 (0.105)	-0.588 <b>(0.092)</b>	-0.142 (0.329)	-0.200 (0.250)
Total Years Incarcerated	30	-0.024 <b>(0.053)</b>	-0.025 <b>(0.066)</b>				-0.037 <b>(0.053)</b>	-0.032 <b>(0.053)</b>	-0.038 <b>(0.066)</b>
Crime Factor	30 to Mid-30s	-0.155 <b>(0.079)</b>	-0.176 (0.118)	-0.508 (0.118)	-0.523 (0.184)	-0.504 (0.145)	-0.071 <b>(0.092)</b>	-0.026 (0.316)	-0.041 (0.237)

Table C.53: Treatment Effects on Tobacco, Drugs, Alcohol, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cig. Smoked per day last month	30	-0.765 (0.276)	0.000 (0.566)	-2.337 (0.171)	0.463 (0.632)	-2.131 (0.158)	-0.530 (0.342)	0.253 (0.526)	-0.302 (0.395)
Days drank alcohol last month	30	-0.742 (0.289)	0.085 (0.526)	-0.567 (0.368)	0.274 (0.434)	-0.261 (0.395)	-0.918 (0.303)	0.043 (0.474)	-0.464 (0.421)
Days binge drank alcohol last month	30	-0.358 (0.276)	0.216 (0.579)	-1.062 (0.289)	-0.070 (0.461)	-0.922 (0.303)	-0.232 (0.342)	0.455 (0.776)	0.034 (0.526)
Self-reported drug user	Mid-30s	-0.033 (0.395)	-0.079 (0.276)	-0.116 (0.263)	-0.196 (0.158)	-0.101 (0.276)	-0.010 (0.434)	-0.043 (0.382)	0.033 (0.592)
Substance Use Factor	30 to Mid-30s	0.006 (0.500)	0.317 (0.895)	0.308 (0.961)	0.497 (0.934)	0.313 (0.974)	-0.076 (0.408)	0.264 (0.816)	0.067 (0.671)

Table C.54: Treatment Effects on Self-Reported Health, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported Health	30	-0.184 (0.237)	-0.215 (0.197)	-1.150 <b>(0.013)</b>	-0.988 <b>(0.000)</b>	-1.188 <b>(0.013)</b>	0.044 (0.618)	-0.027 (0.461)	-0.016 (0.461)
	Mid-30s	0.006 (0.487)	-0.108 (0.342)	-0.232 (0.250)	-0.171 (0.303)	-0.252 (0.237)	0.071 (0.618)	-0.087 (0.368)	0.010 (0.526)
Self-reported Health Factor	30 to Mid-30s	0.045 (0.605)	0.001 (0.434)	0.203 (0.803)	0.252 (0.855)	0.180 (0.829)	0.002 (0.513)	-0.071 (0.329)	-0.011 (0.421)

Table C.55: Treatment Effects on Hypertension, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Systolic Blood Pressure (mm Hg)	Mid-30s	-2.899 (0.289)	-3.901 (0.263)	1.065 (0.592)	2.263 (0.645)	-0.577 (0.526)	-3.980 (0.316)	-4.164 (0.316)	-6.433 (0.171)
Diastolic Blood Pressure (mm Hg)	Mid-30s	-0.002 (0.461)	1.317 (0.500)	4.725 (0.816)	7.497 (0.829)	4.337 (0.750)	-1.290 (0.316)	0.423 (0.461)	-1.783 (0.355)
Prehypertension	Mid-30s	-0.189 <b>(0.013)</b>	-0.224 <b>(0.000)</b>	-0.094 (0.316)	-0.194 (0.171)	-0.160 (0.250)	-0.215 <b>(0.013)</b>	-0.248 <b>(0.000)</b>	-0.275 <b>(0.000)</b>
Hypertension	Mid-30s	0.172 (0.895)	0.159 (0.842)	0.232 (0.895)	0.156 (0.803)	0.181 (0.829)	0.156 (0.829)	0.193 (0.842)	0.083 (0.724)
Hypertension Factor	Mid-30s	-0.033 (0.395)	-0.035 (0.421)	0.204 (0.763)	0.233 (0.697)	0.125 (0.632)	-0.097 (0.368)	-0.053 (0.434)	-0.202 (0.303)

Table C.56: Treatment Effects on Cholesterol, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	2.884 (0.197)	5.555 (0.105)	10.514 <b>(0.000)</b>	6.966 (0.158)	13.319 <b>(0.000)</b>	0.802 (0.434)	4.446 (0.211)	4.243 (0.171)
Dyslipidemia	Mid-30s	0.051 (0.803)	0.043 (0.645)	-0.080 (0.211)	-0.072 (0.461)	-0.093 (0.237)	0.087 (0.895)	0.101 (0.855)	0.096 (0.882)
Cholesterol Factor	Mid-30s	-0.024 (0.421)	-0.111 (0.329)	-0.396 <b>(0.039)</b>	-0.284 (0.342)	-0.492 <b>(0.039)</b>	0.078 (0.658)	-0.011 (0.487)	-0.011 (0.487)

Table C.57: Treatment Effects on Diabetes, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hemoglobin Level (%)	Mid-30s	-0.277 (0.211)	-0.101 (0.316)	-0.176 <b>(0.092)</b>	-0.088 (0.355)	-0.190 (0.105)	-0.304 (0.211)	-0.037 (0.474)	-0.355 (0.197)
Prediabetes	Mid-30s	0.088 (0.737)	0.163 (0.789)	0.076 (0.632)	0.176 (0.763)	0.035 (0.526)	0.091 (0.697)	0.161 (0.750)	0.029 (0.579)
Diabetes	Mid-30s	-0.071 <b>(0.066)</b>	-0.032 (0.171)				-0.091 <b>(0.066)</b>	-0.039 (0.171)	-0.095 <b>(0.039)</b>
Diabetes Factor	Mid-30s	-0.249 (0.145)	-0.086 (0.250)	-0.086 (0.211)	-0.029 (0.382)	-0.098 (0.158)	-0.294 (0.145)	-0.064 (0.408)	-0.334 (0.145)

Table C.58: Treatment Effects on Vitamin D Deficiency, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vitamin D Deficiency	Mid-30s	-0.047 (0.303)	0.061 (0.658)	-0.094 (0.316)	-0.007 (0.368)	-0.034 (0.434)	-0.034 (0.355)	0.093 (0.711)	0.038 (0.553)

Table C.59: Treatment Effects on Obesity, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measured BMI	Mid-30s	3.545 (0.855)	5.649 (0.974)	1.937 (0.724)	3.117 (0.724)	1.969 (0.697)	3.983 (0.868)	6.692 (0.947)	4.705 (0.961)
Obesity	Mid-30s	-0.011 (0.474)	0.087 (0.711)	-0.261 <b>(0.013)</b>	-0.115 (0.171)	-0.199 <b>(0.026)</b>	0.057 (0.684)	0.178 (0.829)	0.109 (0.789)
Severe Obesity	Mid-30s	-0.045 (0.368)	0.065 (0.632)	0.015 (0.513)	0.076 (0.632)	0.019 (0.539)	-0.061 (0.395)	0.064 (0.579)	-0.040 (0.408)
Waist-hip Ratio	Mid-30s	-0.022 (0.224)	0.004 (0.539)	-0.076 <b>(0.079)</b>	-0.076 (0.171)	-0.072 <b>(0.079)</b>	-0.007 (0.408)	0.036 (0.763)	0.005 (0.553)
Abdominal Obesity	Mid-30s	-0.159 (0.171)	-0.016 (0.447)	-0.381 <b>(0.000)</b>	-0.207 (0.105)	-0.294 <b>(0.013)</b>	-0.095 (0.316)	0.054 (0.592)	-0.029 (0.434)
Framingham Risk Score	Mid-30s	-0.259 (0.132)	-0.228 (0.158)	-0.488 <b>(0.053)</b>	-0.551 (0.105)	-0.525 <b>(0.066)</b>	-0.197 (0.237)	-0.078 (0.355)	-0.220 (0.184)
Obesity Factor	Mid-30s	0.088 (0.645)	0.303 (0.803)	-0.289 (0.211)	-0.133 (0.395)	-0.267 (0.224)	0.196 (0.724)	0.474 (0.842)	0.255 (0.763)

Table C.60: Treatment Effects on Mental Health, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Somatization	21	-0.059 (0.395)	-0.170 (0.158)	-0.281 (0.171)	-0.431 <b>(0.039)</b>	-0.370 <b>(0.092)</b>	-0.006 (0.526)	-0.109 (0.250)	-0.100 (0.211)
	34	0.003 (0.526)	0.033 (0.513)	-0.199 (0.368)	-0.404 (0.237)	-0.198 (0.316)	0.058 (0.605)	0.110 (0.763)	0.075 (0.645)
Depression	21	-0.475 <b>(0.000)</b>	-0.483 <b>(0.013)</b>	-0.818 <b>(0.000)</b>	-0.905 <b>(0.013)</b>	-0.814 <b>(0.000)</b>	-0.346 <b>(0.013)</b>	-0.359 <b>(0.079)</b>	-0.317 <b>(0.039)</b>
	34	-0.125 (0.303)	-0.180 <b>(0.079)</b>	-0.149 (0.368)	-0.246 (0.276)	-0.226 (0.263)	-0.118 (0.303)	-0.217 (0.118)	-0.179 (0.158)
Anxiety	21	-0.387 <b>(0.026)</b>	-0.365 <b>(0.053)</b>	-0.643 <b>(0.039)</b>	-0.732 <b>(0.026)</b>	-0.606 <b>(0.039)</b>	-0.273 <b>(0.066)</b>	-0.240 (0.158)	-0.223 (0.132)
	34	-0.245 (0.118)	-0.218 <b>(0.066)</b>	-0.342 (0.224)	-0.474 (0.171)	-0.389 (0.184)	-0.218 (0.171)	-0.177 (0.132)	-0.265 (0.118)
Hostility	21	-0.475 <b>(0.013)</b>	-0.602 <b>(0.013)</b>	-1.036 <b>(0.000)</b>	-1.218 <b>(0.000)</b>	-1.110 <b>(0.000)</b>	-0.240 (0.132)	-0.403 <b>(0.053)</b>	-0.305 (0.105)
	34	-0.080 (0.329)	-0.058 (0.316)	-0.109 (0.368)	-0.217 (0.211)	-0.132 (0.263)	-0.072 (0.395)	-0.018 (0.487)	-0.017 (0.395)
Global Severity Index	21	-0.335 <b>(0.000)</b>	-0.351 <b>(0.013)</b>	-0.615 <b>(0.000)</b>	-0.665 <b>(0.000)</b>	-0.620 <b>(0.000)</b>	-0.234 <b>(0.053)</b>	-0.247 (0.118)	-0.225 <b>(0.066)</b>
	34	-2.197 (0.303)	-2.191 (0.145)	-4.138 (0.329)	-6.745 (0.224)	-4.884 (0.250)	-1.668 (0.316)	-1.704 (0.237)	-2.210 (0.197)
BSI Factor	21 and 34	-0.641 <b>(0.026)</b>	-0.347 (0.105)	-0.847 (0.105)	-0.799 (0.184)	-0.674 (0.184)	-0.583 <b>(0.039)</b>	-0.243 (0.250)	-0.396 <b>(0.079)</b>

## C.6 Treatment Effects for Pooled Sample, Step Down

Table C.61: Treatment Effects on IQ Scores, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. IQ Test	2	10.116	10.461	10.609	12.312	11.140	9.863	10.116	10.214
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3	13.450	14.071	19.242	21.795	20.287	11.314	12.053	11.777
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3.5	8.387	7.886	11.255	12.220	11.438	7.276	6.711	7.008
		<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.026)</b>
	4	9.166	8.768	11.985	12.570	12.562	8.149	7.853	8.525
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>
	4.5	8.380	8.004	13.287	13.937	13.523	6.717	6.246	6.827
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.026)</b>	<b>(0.039)</b>
	5	6.362	5.621	8.310	9.003	8.749	5.760	4.776	5.595
		<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.053)</b>	<b>(0.013)</b>	<b>(0.053)</b>	(0.105)	<b>(0.053)</b>
	6.6	5.956	5.497	4.088	4.646	4.845	5.850	5.369	6.138
		<b>(0.013)</b>	<b>(0.013)</b>	(0.303)	(0.250)	(0.237)	<b>(0.039)</b>	<b>(0.066)</b>	<b>(0.053)</b>
7	5.373	4.912	6.575	7.626	6.822	5.066	3.970	4.945	
	<b>(0.013)</b>	<b>(0.066)</b>	(0.132)	<b>(0.066)</b>	(0.118)	<b>(0.092)</b>	(0.132)	<b>(0.066)</b>	
8	4.932	3.848	2.570	2.766	2.709	4.948	3.952	4.799	
	<b>(0.092)</b>	(0.145)	(0.579)	(0.553)	(0.579)	<b>(0.079)</b>	(0.132)	<b>(0.053)</b>	
12	4.524	3.612	3.251	2.881	2.731	4.766	3.628	3.579	
	<b>(0.053)</b>	(0.105)	(0.408)	(0.421)	(0.487)	<b>(0.079)</b>	(0.118)	<b>(0.092)</b>	
IQ Factor	2 to 5	0.744	0.723	1.017	1.113	1.055	0.662	0.625	0.670
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	6 to 12	0.403	0.377	0.373	0.434	0.400	0.411	0.355	0.411
		<b>(0.039)</b>	<b>(0.053)</b>	(0.237)	(0.250)	(0.237)	<b>(0.066)</b>	(0.118)	<b>(0.092)</b>

Table C.62: Treatment Effects on Achievement Scores, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. Achv. Test	5.5	8.029	8.366	14.284	16.754	15.265	6.223	5.634	7.268
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.092)</b>	<b>(0.000)</b>
	6	4.543	5.026	6.178	7.971	7.124	4.075	4.167	4.722
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.000)</b>
	7.5	1.937	2.136	0.667	3.618	0.558	2.308	1.762	2.129
		(0.118)	<b>(0.066)</b>	(0.566)	(0.132)	(0.579)	(0.118)	(0.132)	(0.132)
8	4.207	4.715	1.630	5.607	2.393	4.958	4.637	5.508	
	<b>(0.039)</b>	<b>(0.000)</b>	(0.421)	<b>(0.092)</b>	(0.342)	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.026)</b>	
8.5	5.938	6.119	5.046	7.926	5.390	5.507	5.008	5.828	
	<b>(0.013)</b>	<b>(0.000)</b>	(0.171)	<b>(0.066)</b>	(0.132)	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.026)</b>	
Achievement Factor	5.5 to 12	0.483	0.540	0.570	0.819	0.620	0.456	0.458	0.496
		<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.092)</b>	<b>(0.066)</b>	<b>(0.053)</b>	<b>(0.013)</b>	<b>(0.026)</b>	<b>(0.013)</b>

Table C.63: Treatment Effects on HOME Scores, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOME Score	0.5	1.005 (0.434)	0.168 (0.803)	1.332 (0.342)	0.523 (0.763)	0.773 (0.592)	0.566 (0.632)	0.071 (0.882)	0.195 (0.803)
	1.5	1.126 (0.421)	0.128 (0.829)	2.706 (0.105)	1.474 (0.368)	2.577 (0.145)	0.368 (0.816)	-0.098 (0.908)	0.435 (0.776)
	2.5	0.441 (0.724)	0.236 (0.789)	3.089 <b>(0.053)</b>	3.726 <b>(0.053)</b>	3.493 <b>(0.039)</b>	-0.588 (0.987)	-0.554 (0.974)	-0.051 (0.908)
	3.5	2.112 (0.329)	0.942 (0.684)	8.288 <b>(0.013)</b>	7.799 <b>(0.026)</b>	8.165 <b>(0.013)</b>	0.306 (0.855)	-0.492 (0.961)	0.325 (0.803)
	4.5	1.927 (0.368)	0.451 (0.750)	8.156 <b>(0.013)</b>	6.570 <b>(0.053)</b>	8.047 <b>(0.013)</b>	0.146 (0.868)	-0.674 (0.974)	0.339 (0.803)
	8	1.004 (0.592)	0.502 (0.737)	3.102 (0.250)	0.991 (0.750)	3.624 (0.171)	0.492 (0.842)	0.359 (0.816)	1.303 (0.579)
	HOME Factor	0.5 to 8	0.263 (0.368)	0.087 (0.697)	0.780 <b>(0.066)</b>	0.545 (0.211)	0.783 <b>(0.066)</b>	0.134 (0.592)	-0.018 (0.908)

Table C.64: Treatment Effects on Parent Income, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parental Income	1.5	2,248 (0.237)	3,277 (0.184)	2,553 (0.408)	4,721 (0.276)	5,028 (0.184)	1,870 (0.329)	2,901 (0.237)	3,718 (0.145)
	2.5	516 (0.632)	366 (0.711)	-2,455 (0.921)	-851 (0.842)	93,152 (0.724)	988 (0.553)	469 (0.684)	1,555 (0.434)
	3.5	1,821 (0.421)	1,901 (0.447)	3,984 (0.237)	4,990 (0.263)	5,265 (0.171)	961 (0.592)	961 (0.618)	2,108 (0.342)
	4.5	2,336 (0.171)	3,565 (0.171)	4,159 (0.158)	4,738 (0.105)	2,703 (0.395)	1,434 (0.382)	2,926 (0.303)	2,697 (0.276)
Parental Income Factor	1.5 to 15	0.156 (0.368)	0.120 (0.513)	0.158 (0.500)	0.303 (0.382)	0.207 (0.474)	0.127 (0.434)	0.063 (0.671)	0.183 (0.342)

Table C.65: Treatment Effects on Mother's Employment, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother Works	2	0.114 <b>(0.092)</b>	0.076 (0.329)	0.296 <b>(0.000)</b>	0.254 <b>(0.092)</b>	0.277 <b>(0.053)</b>	0.048 (0.474)	0.034 (0.645)	0.039 (0.513)
	3	0.119 (0.118)	0.087 (0.289)	0.219 <b>(0.053)</b>	0.172 (0.105)	0.199 <b>(0.079)</b>	0.092 (0.224)	0.065 (0.447)	0.087 (0.276)
	4	0.127 <b>(0.026)</b>	0.109 (0.171)	0.306 <b>(0.000)</b>	0.275 <b>(0.039)</b>	0.290 <b>(0.013)</b>	0.076 (0.211)	0.071 (0.316)	0.071 (0.276)
	5	0.089 (0.224)	0.059 (0.382)	0.342 <b>(0.000)</b>	0.321 <b>(0.026)</b>	0.346 <b>(0.000)</b>	0.005 (0.763)	-0.025 (0.934)	0.016 (0.658)
Mother Works Factor	2 to 21	0.256 (0.184)	0.179 (0.395)	0.717 <b>(0.053)</b>	0.641 <b>(0.079)</b>	0.679 <b>(0.092)</b>	0.133 (0.474)	0.065 (0.697)	0.126 (0.513)

Table C.66: Treatment Effects on Father at Home, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father at Home	2	-0.010 (0.789)	0.063 (0.408)	-0.187 (1.000)	-0.116 (0.961)	-0.148 (0.974)	0.047 (0.487)	0.142 <b>(0.079)</b>	0.130 (0.105)
	3	-0.076 (0.961)	-0.015 (0.882)	-0.291 (1.000)	-0.226 (0.987)	-0.258 (1.000)	0.002 (0.697)	0.066 (0.355)	0.079 (0.289)
	4	-0.071 (0.934)	-0.009 (0.842)	-0.331 (1.000)	-0.252 (1.000)	-0.293 (1.000)	0.021 (0.645)	0.074 (0.368)	0.101 (0.237)
	5	-0.093 (0.974)	-0.026 (0.908)	-0.369 (1.000)	-0.249 (0.987)	-0.343 (1.000)	-0.006 (0.776)	0.042 (0.474)	0.062 (0.395)
	8	0.052 (0.461)	0.073 (0.395)	-0.124 (0.961)	-0.036 (0.908)	-0.113 (0.947)	0.113 (0.184)	0.115 (0.171)	0.134 (0.132)
Father at Home Factor	2 to 8	-0.173 (0.947)	-0.028 (0.868)	-0.805 (1.000)	-0.548 (0.987)	-0.748 (1.000)	0.033 (0.671)	0.152 (0.355)	0.207 (0.250)

Table C.67: Treatment Effects on Education, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduated High School	30	0.164 <b>(0.026)</b>	0.116 <b>(0.118)</b>	0.390 <b>(0.000)</b>	0.340 <b>(0.026)</b>	0.353 <b>(0.000)</b>	0.103 <b>(0.118)</b>	0.049 <b>(0.579)</b>	0.059 <b>(0.461)</b>
Attended Voc./Tech./Com. College	30	-0.091 <b>(0.867)</b>	-0.101 <b>(0.987)</b>		0.116 <b>(0.211)</b>	-0.044 <b>(0.658)</b>	-0.100 <b>(1.000)</b>	-0.171 <b>(1.000)</b>	-0.152 <b>(0.987)</b>
Graduated 4-year College	30	0.161 <b>(0.000)</b>	0.141 <b>(0.079)</b>	0.188 <b>(0.026)</b>	0.168 <b>(0.066)</b>	0.181 <b>(0.053)</b>	0.148 <b>(0.026)</b>	0.120 <b>(0.092)</b>	0.136 <b>(0.066)</b>
Years of Edu.	30	1.367 <b>(0.000)</b>	1.070 <b>(0.000)</b>	2.513 <b>(0.000)</b>	2.108 <b>(0.000)</b>	2.430 <b>(0.000)</b>	0.986 <b>(0.013)</b>	0.759 <b>(0.066)</b>	0.886 <b>(0.013)</b>
Education Factor	30	0.513 <b>(0.000)</b>	0.390 <b>(0.000)</b>	0.843 <b>(0.000)</b>	0.725 <b>(0.000)</b>	0.787 <b>(0.000)</b>	0.389 <b>(0.013)</b>	0.294 <b>(0.092)</b>	0.316 <b>(0.053)</b>

Table C.68: Treatment Effects on Employment and Income, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	30	0.125 <b>(0.118)</b>	0.132 <b>(0.145)</b>	0.164 <b>(0.355)</b>	0.128 <b>(0.513)</b>	0.205 <b>(0.224)</b>	0.111 <b>(0.316)</b>	0.148 <b>(0.197)</b>	0.163 <b>(0.079)</b>
Labor Income	21	167 <b>(0.908)</b>	-579 <b>(0.961)</b>	1,577 <b>(0.724)</b>	1,469 <b>(0.842)</b>	1,749 <b>(0.829)</b>	-429 <b>(0.987)</b>	-1,211 <b>(0.987)</b>	-1,527 <b>(1.000)</b>
	30	12,377 <b>(0.145)</b>	10,752 <b>(0.658)</b>	17,677 <b>(0.092)</b>	13,888 <b>(0.592)</b>	21,198 <b>(0.118)</b>	10,847 <b>(0.329)</b>	9,790 <b>(0.737)</b>	11,232 <b>(0.513)</b>
Public-Transfer Income	21	-728 <b>(0.684)</b>	-923 <b>(0.579)</b>	-247 <b>(0.789)</b>	-1,161 <b>(0.566)</b>	-1,625 <b>(0.461)</b>	-1,054 <b>(0.605)</b>	-824 <b>(0.658)</b>	-813 <b>(0.645)</b>
	30	-1,832 <b>(0.118)</b>	-1,005 <b>(0.487)</b>	-1,613 <b>(0.237)</b>	-1,480 <b>(0.368)</b>	-1,600 <b>(0.382)</b>	-1,483 <b>(0.316)</b>	-992 <b>(0.605)</b>	-1,680 <b>(0.342)</b>
Employment Factor	21 to 30	0.344 <b>(0.013)</b>	0.267 <b>(0.132)</b>	0.393 <b>(0.158)</b>	0.315 <b>(0.461)</b>	0.423 <b>(0.118)</b>	0.307 <b>(0.105)</b>	0.291 <b>(0.171)</b>	0.313 <b>(0.079)</b>

Table C.69: Treatment Effects on Crime, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Felony Arrests	Mid-30s	0.045 <b>(0.842)</b>	0.237 <b>(0.934)</b>	-0.132 <b>(0.697)</b>	0.227 <b>(0.908)</b>	0.205 <b>(0.921)</b>	0.112 <b>(0.882)</b>	0.224 <b>(0.961)</b>	0.185 <b>(0.961)</b>
Total Misdemeanor Arrests	Mid-30s	-0.689 <b>(0.118)</b>	-0.481 <b>(0.395)</b>	-1.445 <b>(0.237)</b>	-1.198 <b>(0.461)</b>	-1.277 <b>(0.421)</b>	-0.546 <b>(0.211)</b>	-0.293 <b>(0.500)</b>	-0.310 <b>(0.461)</b>
Total Years Incarcerated	30	0.168 <b>(1.000)</b>	0.187 <b>(1.000)</b>	0.280 <b>(1.000)</b>	0.325 <b>(1.000)</b>	0.350 <b>(1.000)</b>	0.153 <b>(0.987)</b>	0.176 <b>(1.000)</b>	0.215 <b>(0.987)</b>
Crime Factor	30 to Mid-30s	0.054 <b>(0.895)</b>	0.072 <b>(0.895)</b>	0.012 <b>(0.855)</b>	0.045 <b>(0.816)</b>	0.051 <b>(0.882)</b>	0.076 <b>(0.934)</b>	0.091 <b>(0.961)</b>	0.131 <b>(0.987)</b>

Table C.70: Treatment Effects on Tobacco, Drugs, Alcohol, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cig. Smoked per day last month	30	0.033 <b>(0.842)</b>	0.321 <b>(0.921)</b>	-0.826 <b>(0.697)</b>	-0.150 <b>(0.882)</b>	-0.802 <b>(0.711)</b>	0.434 <b>(0.947)</b>	0.576 <b>(0.947)</b>	0.434 <b>(0.961)</b>
Days drank alcohol last month	30	0.244 <b>(0.921)</b>	0.419 <b>(0.921)</b>	-0.156 <b>(0.882)</b>	-0.700 <b>(0.776)</b>	0.113 <b>(0.868)</b>	0.207 <b>(0.908)</b>	0.618 <b>(0.947)</b>	0.630 <b>(0.961)</b>
Days binge drank alcohol last month	30	0.085 <b>(0.921)</b>	0.412 <b>(0.987)</b>	-0.267 <b>(0.803)</b>	-0.218 <b>(0.868)</b>	-0.126 <b>(0.855)</b>	0.151 <b>(0.934)</b>	0.674 <b>(0.987)</b>	0.393 <b>(0.974)</b>
Self-reported drug user	Mid-30s	-0.142 <b>(0.145)</b>	-0.183 <b>(0.158)</b>	-0.253 <b>(0.132)</b>	-0.303 <b>(0.105)</b>	-0.274 <b>(0.171)</b>	-0.090 <b>(0.513)</b>	-0.122 <b>(0.461)</b>	-0.114 <b>(0.368)</b>
Substance Use Factor	30 to Mid-30s	0.187 <b>(1.000)</b>	0.213 <b>(0.987)</b>	0.335 <b>(1.000)</b>	0.330 <b>(0.987)</b>	0.339 <b>(1.000)</b>	0.149 <b>(0.961)</b>	0.201 <b>(0.987)</b>	0.202 <b>(1.000)</b>

Table C.71: Treatment Effects on Self-Reported Health, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported Health	30	-0.066 (0.684)	0.007 (0.921)	-0.369 (0.303)	-0.264 (0.500)	-0.330 (0.395)	-0.009 (0.829)	0.004 (0.908)	0.036 (0.947)
	Mid-30s	0.079 (0.987)	-0.019 (0.855)	-0.429 (0.132)	-0.442 (0.289)	-0.498 (0.132)	0.179 (1.000)	0.113 (1.000)	0.066 (0.987)
Self-reported Health Factor	30 to Mid-30s	-0.012 (0.816)	-0.052 (0.750)	-0.083 (0.658)	-0.093 (0.684)	-0.129 (0.566)	-0.012 (0.829)	-0.021 (0.829)	-0.051 (0.697)

Table C.72: Treatment Effects on Hypertension, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Systolic Blood Pressure (mm Hg)	Mid-30s	-5.625 (0.118)	-10.849 <b>(0.079)</b>	5.375 (1.000)	6.809 (0.961)	4.383 (0.987)	-9.438 (0.105)	-16.448 <b>(0.000)</b>	-15.882 <b>(0.013)</b>
Diastolic Blood Pressure (mm Hg)	Mid-30s	-5.312 <b>(0.092)</b>	-7.745 <b>(0.079)</b>	-1.424 (0.605)	0.519 (0.842)	-1.637 (0.632)	-7.219 <b>(0.092)</b>	-9.589 <b>(0.092)</b>	-10.696 <b>(0.013)</b>
Prehypertension	Mid-30s	-0.176 <b>(0.066)</b>	-0.181 <b>(0.079)</b>	-0.049 (0.658)	-0.055 (0.645)	-0.060 (0.658)	-0.240 <b>(0.000)</b>	-0.268 <b>(0.000)</b>	-0.280 <b>(0.000)</b>
Hypertension	Mid-30s	-0.036 (0.487)	-0.097 (0.158)	0.083 (0.947)	0.143 (0.934)	0.045 (0.908)	-0.083 (0.289)	-0.151 (0.118)	-0.187 <b>(0.066)</b>
Hypertension Factor	Mid-30s	-0.293 <b>(0.092)</b>	-0.475 <b>(0.079)</b>	0.089 (0.947)	0.185 (0.921)	0.046 (0.895)	-0.447 <b>(0.039)</b>	-0.670 <b>(0.013)</b>	-0.703 <b>(0.013)</b>

Table C.73: Treatment Effects on Cholesterol, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	3.872 (0.184)	4.790 <b>(0.079)</b>	5.806 (0.132)	3.305 (0.395)	5.518 (0.145)	2.964 (0.276)	5.364 <b>(0.053)</b>	4.456 (0.118)
Dyslipidemia	Mid-30s	0.013 (0.763)	-0.003 (0.539)	0.035 (0.763)	0.082 (0.763)	0.031 (0.789)	0.032 (0.816)	-0.009 (0.526)	0.010 (0.737)
Cholesterol Factor	Mid-30s	-0.097 (0.434)	-0.141 (0.184)	-0.127 (0.342)	-0.000 (0.671)	-0.124 (0.355)	-0.049 (0.553)	-0.165 (0.184)	-0.117 (0.368)

Table C.74: Treatment Effects on Diabetes, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hemoglobin Level (%)	Mid-30s	0.003 (0.776)	0.075 (0.868)	0.032 (0.882)	-0.093 (0.658)	0.073 (0.882)	-0.029 (0.645)	0.086 (0.842)	-0.002 (0.711)
Prediabetes	Mid-30s	0.004 (0.789)	-0.039 (0.632)	-0.040 (0.697)	-0.110 (0.474)	-0.097 (0.566)	0.004 (0.671)	-0.045 (0.658)	-0.054 (0.579)
Diabetes	Mid-30s	-0.002 (0.724)	0.015 (0.868)	0.043 (1.000)	-0.003 (0.737)	0.049 (1.000)	-0.015 (0.605)	0.020 (0.868)	-0.007 (0.684)
Diabetes Factor	Mid-30s	-0.001 (0.737)	0.060 (0.868)	0.079 (0.934)	-0.065 (0.632)	0.104 (0.947)	-0.038 (0.645)	0.073 (0.842)	-0.017 (0.711)

Table C.75: Treatment Effects on Vitamin D Deficiency, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vitamin D Deficiency	Mid-30s	-0.153 <b>(0.039)</b>	-0.069 (0.237)	-0.264 <b>(0.000)</b>	-0.085 (0.342)	-0.233 <b>(0.000)</b>	-0.118 (0.132)	-0.031 (0.408)	-0.078 (0.263)

Table C.76: Treatment Effects on Obesity, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measured BMI	Mid-30s	0.999 (0.974)	3.094 (1.000)	-0.202 (0.842)	1.254 (1.000)	0.710 (0.921)	1.072 (0.974)	3.548 (1.000)	1.827 (0.987)
Obesity	Mid-30s	-0.050 (0.789)	0.063 (0.987)	-0.256 <b>(0.079)</b>	-0.144 (0.461)	-0.144 (0.513)	-0.013 (0.882)	0.113 (0.987)	0.011 (0.908)
Severe Obesity	Mid-30s	-0.126 (0.289)	-0.018 (0.816)	-0.093 (0.632)	-0.069 (0.855)	-0.065 (0.750)	-0.147 (0.276)	-0.011 (0.855)	-0.107 (0.500)
Waist-hip Ratio	Mid-30s	-0.006 (0.842)	0.000 (0.882)	-0.037 (0.434)	-0.047 (0.592)	-0.031 (0.579)	0.003 (0.947)	0.012 (0.987)	0.006 (0.947)
Abdominal Obesity	Mid-30s	-0.091 (0.579)	-0.021 (0.816)	-0.230 <b>(0.079)</b>	-0.139 (0.500)	-0.157 (0.355)	-0.041 (0.803)	0.031 (0.961)	-0.030 (0.816)
Framingham Risk Score	Mid-30s	0.348 (0.974)	-0.399 (0.645)	0.948 (1.000)	0.299 (1.000)	0.902 (1.000)	0.351 (0.974)	-0.704 (0.329)	0.088 (0.908)
Obesity Factor	Mid-30s	-0.022 (0.895)	0.162 (0.987)	-0.268 (0.579)	-0.138 (0.803)	-0.134 (0.750)	0.033 (0.947)	0.250 (0.987)	0.077 (0.961)

Table C.77: Treatment Effects on Mental Health, Pooled Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Somatization	21	-0.015 (0.855)	-0.068 (0.618)	-0.118 (0.618)	-0.240 (0.263)	-0.182 (0.355)	0.012 (0.974)	-0.046 (0.724)	-0.071 (0.632)
	34	-0.128 (0.539)	-0.119 (0.500)	-0.191 (0.724)	-0.202 (0.803)	-0.205 (0.711)	-0.122 (0.592)	-0.114 (0.579)	-0.123 (0.553)
Depression	21	-0.256 (0.118)	-0.207 (0.263)	-0.329 (0.263)	-0.373 (0.184)	-0.365 (0.211)	-0.184 (0.355)	-0.164 (0.342)	-0.202 (0.263)
	34	-0.165 (0.526)	-0.203 (0.355)	-0.061 (0.842)	0.001 (0.947)	-0.117 (0.803)	-0.211 (0.421)	-0.292 (0.184)	-0.236 (0.408)
Anxiety	21	-0.119 (0.474)	-0.110 (0.474)	-0.192 (0.526)	-0.344 (0.197)	-0.208 (0.474)	-0.077 (0.592)	-0.083 (0.632)	-0.102 (0.632)
	34	-0.256 (0.158)	-0.259 (0.197)	-0.224 (0.645)	-0.217 (0.724)	-0.262 (0.566)	-0.280 (0.132)	-0.286 (0.171)	-0.300 (0.118)
Hostility	21	-0.296 (0.105)	-0.279 (0.145)	-0.504 <b>(0.092)</b>	-0.666 <b>(0.066)</b>	-0.564 <b>(0.066)</b>	-0.191 (0.421)	-0.199 (0.329)	-0.229 (0.250)
	34	-0.145 (0.526)	-0.126 (0.447)	-0.090 (0.724)	-0.057 (0.895)	-0.129 (0.592)	-0.174 (0.513)	-0.144 (0.434)	-0.150 (0.539)
Global Severity Index	21	-0.147 (0.276)	-0.140 (0.329)	-0.221 (0.237)	-0.325 <b>(0.092)</b>	-0.258 (0.145)	-0.107 (0.500)	-0.107 (0.500)	-0.144 (0.329)
	34	-3.294 (0.368)	-3.487 (0.316)	-2.858 (0.750)	-2.506 (0.842)	-3.502 (0.711)	-3.674 (0.382)	-4.149 (0.211)	-3.952 (0.276)
BSI Factor	21 and 34	-0.559 <b>(0.000)</b>	-0.330 (0.289)	-0.600 (0.289)	-0.405 (0.579)	-0.485 (0.434)	-0.547 <b>(0.000)</b>	-0.371 (0.263)	-0.350 (0.211)

## C.7 Treatment Effects for Male Sample, Step Down

Table C.78: Treatment Effects on IQ Scores, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. IQ Test	2	9.528	11.036	6.875	10.704	7.944	10.286	11.497	11.080
		<b>(0.000)</b>	<b>(0.000)</b>	(0.105)	(0.171)	<b>(0.079)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3	13.410	14.873	13.896	17.254	15.474	13.271	14.229	14.302
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	3.5	8.756	7.994	6.354	5.904	6.784	9.443	8.690	9.043
		<b>(0.000)</b>	<b>(0.053)</b>	(0.158)	(0.382)	(0.118)	<b>(0.000)</b>	<b>(0.053)</b>	<b>(0.013)</b>
	4	12.089	11.867	8.950	7.899	9.668	12.986	12.604	13.487
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.053)</b>	(0.329)	<b>(0.066)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	4.5	8.508	8.914	10.411	11.349	10.639	7.964	8.518	7.804
		<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.092)</b>	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.039)</b>	<b>(0.026)</b>
	5	7.697	7.130	4.643	3.469	4.992	8.679	8.241	8.181
		<b>(0.000)</b>	<b>(0.013)</b>	(0.461)	(0.645)	(0.434)	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.000)</b>
	6.6	5.803	5.157	0.831	-1.794	1.479	5.916	5.542	5.886
		<b>(0.053)</b>	<b>(0.066)</b>	(0.803)	(0.882)	(0.776)	<b>(0.026)</b>	<b>(0.026)</b>	<b>(0.039)</b>
7	4.390	5.812	5.323	8.155	5.750	4.156	5.481	4.418	
	(0.118)	<b>(0.053)</b>	(0.526)	(0.395)	(0.461)	(0.145)	(0.105)	(0.171)	
8	4.160	2.644	-2.514	-4.445	-2.651	4.754	3.391	4.207	
	(0.197)	(0.355)	(0.882)	(0.882)	(0.895)	(0.105)	(0.224)	(0.132)	
12	0.686	-0.856	-0.343	-4.076	-1.018	0.943	-0.116	-0.794	
	(0.618)	(0.737)	(0.855)	(0.882)	(0.855)	(0.618)	(0.737)	(0.829)	
IQ Factor	2 to 5	0.831	0.845	0.708	0.766	0.759	0.867	0.870	0.874
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.053)</b>	(0.197)	<b>(0.066)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>
	6 to 12	0.268	0.227	0.308	0.251	0.304	0.258	0.245	0.226
		(0.197)	(0.289)	(0.658)	(0.711)	(0.645)	(0.211)	(0.250)	(0.276)

Table C.79: Treatment Effects on Achievement Scores, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. Achv. Test	5.5	5.108	5.375	10.088	11.148	10.628	3.863	4.183	4.171
		(0.105)	(0.132)	<b>(0.092)</b>	(0.224)	<b>(0.079)</b>	(0.224)	(0.316)	(0.184)
	6	3.091	3.803	2.271	3.803	3.354	3.312	3.662	3.629
		(0.145)	(0.105)	(0.474)	(0.408)	(0.434)	(0.132)	(0.132)	<b>(0.092)</b>
	7.5	0.019	1.103	-2.767	-2.380	-3.284	0.799	1.762	0.215
		(0.763)	(0.539)	(0.737)	(0.789)	(0.763)	(0.645)	(0.368)	(0.632)
8	2.309	4.117	-3.386	-0.625	-2.713	3.903	5.198	4.240	
	(0.329)	(0.105)	(0.829)	(0.645)	(0.763)	(0.145)	<b>(0.026)</b>	<b>(0.079)</b>	
8.5	3.910	5.586	-1.771	0.775	-1.498	4.199	5.497	4.117	
	(0.184)	<b>(0.092)</b>	(0.684)	(0.618)	(0.684)	(0.197)	<b>(0.066)</b>	(0.118)	
Achievement Factor	5.5 to 12	0.250	0.330	0.030	0.055	0.042	0.308	0.367	0.288
		(0.276)	(0.145)	(0.592)	(0.618)	(0.605)	(0.224)	(0.145)	(0.224)

Table C.80: Treatment Effects on HOME Scores, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOME Score	0.5	0.372 (0.750)	0.039 (0.868)	0.944 (0.645)	-0.416 (0.987)	0.431 (0.724)	0.143 (0.882)	0.069 (0.882)	-0.080 (0.947)
	1.5	-0.500 (0.987)	-0.874 (0.987)	0.431 (0.763)	-1.123 (1.000)	0.216 (0.816)	-0.766 (0.987)	-0.865 (0.987)	-0.872 (1.000)
	2.5	0.141 (0.829)	0.886 (0.658)	1.654 (0.566)	3.806 (0.342)	2.232 (0.461)	-0.292 (0.921)	0.543 (0.763)	0.145 (0.882)
	3.5	1.404 (0.632)	0.668 (0.842)	2.897 (0.566)	1.917 (0.671)	2.835 (0.553)	0.962 (0.776)	0.762 (0.776)	0.743 (0.776)
	4.5	1.146 (0.632)	1.534 (0.645)	3.312 (0.474)	3.894 (0.263)	2.807 (0.513)	0.527 (0.816)	1.279 (0.684)	0.232 (0.868)
	8	1.548 (0.579)	1.524 (0.645)	-0.898 (0.868)	-4.874 (1.000)	-0.503 (0.895)	2.062 (0.434)	2.245 (0.500)	2.634 (0.382)
	HOME Factor	0.5 to 8	0.266 (0.474)	0.180 (0.658)	0.130 (0.724)	-0.093 (1.000)	0.080 (0.737)	0.295 (0.434)	0.218 (0.592)

Table C.81: Treatment Effects on Parent Income, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parental Income	1.5	330 (0.658)	-97.199 (0.855)	-1,046 (0.776)	-2,384 (0.882)	-1,168 (0.816)	-9.245 (0.684)	-26.663 (0.829)	872 (0.658)
	2.5	673 (0.579)	-941 (0.908)	-1,167 (0.776)	-3,542 (0.934)	-1,858 (0.842)	478 (0.618)	-839 (0.908)	232 (0.737)
	3.5	1,036 (0.539)	223 (0.789)	3,085 (0.421)	-1,152 (0.763)	1,448 (0.553)	112 (0.684)	47.496 (0.816)	701 (0.711)
	4.5	821 (0.539)	1,677 (0.566)	1,561 (0.553)	3,815 (0.395)	-2,662 (0.961)	-81.743 (0.697)	917 (0.697)	-429 (0.829)
Parental Income Factor	1.5 to 15	0.130 (0.421)	-0.018 (0.855)	0.079 (0.645)	0.039 (0.618)	0.092 (0.618)	0.088 (0.579)	-0.062 (0.908)	0.116 (0.632)

Table C.82: Treatment Effects on Mother's Employment, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother Works	2	0.056 (0.487)	0.038 (0.658)	0.264 (0.092)	0.184 (0.342)	0.240 (0.171)	-0.004 (0.816)	0.008 (0.789)	-0.018 (0.855)
	3	0.150 (0.158)	0.125 (0.316)	0.261 (0.092)	0.184 (0.342)	0.240 (0.171)	0.117 (0.329)	0.117 (0.382)	0.117 (0.368)
	4	0.134 (0.118)	0.147 (0.197)	0.287 (0.092)	0.262 (0.250)	0.270 (0.105)	0.090 (0.342)	0.124 (0.329)	0.089 (0.382)
	5	0.111 (0.250)	0.095 (0.434)	0.311 (0.092)	0.259 (0.197)	0.289 (0.105)	0.061 (0.461)	0.071 (0.539)	0.054 (0.553)
Mother Works Factor	2 to 21	0.309 (0.224)	0.280 (0.434)	0.865 (0.079)	0.730 (0.289)	0.805 (0.145)	0.160 (0.461)	0.196 (0.539)	0.144 (0.553)

Table C.83: Treatment Effects on Father at Home, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father at Home	2	-0.018 (0.803)	0.160 (0.197)	-0.282 (1.000)	0.111 (0.697)	-0.223 (0.974)	0.057 (0.539)	0.187 (0.171)	0.171 (0.132)
	3	-0.076 (0.895)	0.007 (0.789)	-0.243 (1.000)	0.026 (0.842)	-0.197 (0.974)	-0.029 (0.763)	0.010 (0.763)	0.070 (0.487)
	4	-0.075 (0.895)	0.008 (0.789)	-0.339 (1.000)	-0.050 (0.934)	-0.287 (1.000)		0.036 (0.671)	0.103 (0.355)
	5	-0.057 (0.882)	0.051 (0.632)	-0.429 (1.000)	-0.081 (0.934)	-0.376 (1.000)	0.036 (0.618)	0.089 (0.500)	0.142 (0.237)
	8	0.037 (0.579)	0.070 (0.579)	-0.177 (0.974)	0.058 (0.776)	-0.160 (0.961)	0.123 (0.342)	0.084 (0.500)	0.168 (0.237)
Father at Home Factor	2 to 8	-0.160 (0.882)	0.034 (0.776)	-0.728 (1.000)	0.029 (0.855)	-0.617 (1.000)	0.039 (0.684)	0.073 (0.671)	0.308 (0.276)

Table C.84: Treatment Effects on Education, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduated High School	30	0.073 (0.526)	0.109 (0.539)	0.114 (0.553)	0.113 (0.697)	0.085 (0.605)	0.077 (0.526)	0.102 (0.592)	0.064 (0.645)
Attended Voc./Tech./Com. College	30	-0.099 (0.987)	-0.124 (0.987)	0.086 (0.605)	0.258 (0.303)	0.024 (0.789)	-0.138 (1.000)	-0.231 (1.000)	-0.233 (1.000)
Graduated 4-year College	30	0.170 (0.132)	0.221 (0.184)	0.124 (0.513)	0.323 (0.329)	0.118 (0.500)	0.179 (0.105)	0.210 (0.250)	0.170 (0.224)
Years of Edu.	30	0.525 (0.316)	0.708 (0.263)	0.857 (0.303)	1.302 (0.158)	0.791 (0.303)	0.385 (0.500)	0.540 (0.368)	0.347 (0.579)
Education Factor	30	0.338 (0.132)	0.437 (0.184)	0.298 (0.461)	0.722 (0.224)	0.249 (0.500)	0.346 (0.118)	0.385 (0.263)	0.288 (0.382)

Table C.85: Treatment Effects on Employment and Income, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	30	0.119 (0.395)	0.179 (0.276)	-0.029 (0.921)	-0.050 (0.961)	0.041 (0.868)	0.176 (0.276)	0.245 <b>(0.092)</b>	0.262 <b>(0.079)</b>
Labor Income	21	-1,672 (1.000)	-4,977 (1.000)	-3,951 (0.974)	-13,951 (1.000)	-4,585 (0.974)	-1,527 (1.000)	-3,973 (0.987)	-3,746 (1.000)
	30	19,810 (0.355)	24,902 (0.618)	17,909 (0.355)	21,069 (0.711)	24,012 (0.303)	20,065 (0.316)	28,483 (0.461)	21,170 (0.395)
Public-Transfer Income	21	315 (0.987)	456 (1.000)	1,376 (1.000)	2,654 (0.974)	1,543 (1.000)	-58,901 (0.947)	144 (0.908)	97,591 (0.947)
	30	-530 (0.592)	-176 (0.961)	287 (1.000)	722 (1.000)	548 (1.000)	-279 (0.855)	-82,612 (0.882)	-155 (0.816)
Employment Factor	21 to 30	0.351 (0.197)	0.433 (0.197)	0.053 (0.882)	-0.119 (0.961)	0.086 (0.842)	0.450 <b>(0.079)</b>	0.595 <b>(0.079)</b>	0.486 <b>(0.066)</b>

Table C.86: Treatment Effects on Crime, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Felony Arrests	Mid-30s	0.196 (0.895)	0.683 (0.961)	0.946 (1.000)	2.132 (0.987)	1.337 (1.000)	0.017 (0.789)	0.466 (0.921)	0.183 (0.908)
Total Misdemeanor Arrests	Mid-30s	-0.501 (0.368)	-0.239 (0.592)	-0.251 (0.724)	0.085 (0.961)	-0.040 (0.750)	-0.665 (0.289)	-0.343 (0.513)	-0.512 (0.303)
Total Years Incarcerated	30	0.347 (1.000)	0.489 (1.000)	0.553 (1.000)	0.811 (1.000)	0.701 (1.000)	0.338 (1.000)	0.489 (1.000)	0.470 (1.000)
Crime Factor	30 to Mid-30s	0.169 (0.961)	0.326 (0.974)	0.469 (1.000)	0.835 (1.000)	0.545 (1.000)	0.109 (0.868)	0.272 (0.934)	0.182 (0.987)

Table C.87: Treatment Effects on Tobacco, Drugs, Alcohol, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cig. Smoked per day last month	30	0.826 (0.961)	0.245 (0.882)	0.757 (0.921)	-2.228 (0.408)	0.624 (0.882)	1.429 (1.000)	1.083 (0.987)	1.215 (1.000)
Days drank alcohol last month	30	0.805 (0.947)	1.083 (0.947)	-0.186 (0.829)	-0.648 (0.724)	0.060 (0.763)	0.944 (0.987)	1.348 (0.987)	1.341 (1.000)
Days binge drank alcohol last month	30	0.500 (0.987)	0.474 (0.961)	0.543 (0.987)	-0.360 (0.671)	0.692 (1.000)	0.491 (1.000)	0.625 (0.987)	0.702 (1.000)
Self-reported drug user	Mid-30s	-0.333 <b>(0.066)</b>	-0.432 <b>(0.000)</b>	-0.500 (0.105)	-0.788 <b>(0.000)</b>	-0.555 <b>(0.092)</b>	-0.233 (0.355)	-0.326 (0.197)	-0.330 (0.158)
Substance Use Factor	30 to Mid-30s	0.337 (1.000)	0.232 (0.961)	0.180 (0.947)	0.014 (0.750)	0.219 (0.908)	0.403 (1.000)	0.273 (0.987)	0.292 (1.000)

Table C.88: Treatment Effects on Self-Reported Health, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported Health	30	-0.004 (0.868)	0.226 (1.000)	0.457 (0.987)	0.755 (0.934)	0.568 (1.000)	-0.145 (0.579)	0.096 (0.961)	-0.000 (0.868)
	Mid-30s	0.196 (0.987)	0.010 (0.855)	-0.852 <b>(0.013)</b>	-1.130 <b>(0.053)</b>	-0.925 <b>(0.013)</b>	0.382 (1.000)	0.207 (0.974)	0.170 (0.987)
Self-reported Health Factor	30 to Mid-30s	-0.035 (0.789)	-0.136 (0.618)	-0.563 (0.105)	-0.758 <b>(0.079)</b>	-0.596 (0.105)	0.065 (0.974)	-0.035 (0.829)	-0.044 (0.803)

Table C.89: Treatment Effects on Hypertension, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Systolic Blood Pressure (mm Hg)	Mid-30s	-9.791 (0.105)	-20.743 <b>(0.026)</b>	15.280 (1.000)	16.605 (1.000)	14.253 (1.000)	-19.920 <b>(0.066)</b>	-30.437 <b>(0.000)</b>	-30.587 <b>(0.026)</b>
Diastolic Blood Pressure (mm Hg)	Mid-30s	-10.854 <b>(0.039)</b>	-19.895 <b>(0.013)</b>	-8.640 (0.116)	-12.199 (0.246)	-8.150 (0.130)	-14.240 <b>(0.066)</b>	-22.740 <b>(0.000)</b>	-21.851 <b>(0.026)</b>
Prehypertension	Mid-30s	-0.137 (0.250)	-0.164 (0.184)	0.053 (0.826)	0.192 (0.884)	0.108 (0.913)	-0.280 <b>(0.000)</b>	-0.295 <b>(0.066)</b>	-0.296 <b>(0.000)</b>
Hypertension	Mid-30s	-0.291 <b>(0.079)</b>	-0.383 <b>(0.026)</b>	-0.053 (0.768)	0.090 (0.797)	-0.064 (0.768)	-0.420 <b>(0.039)</b>	-0.511 <b>(0.000)</b>	-0.517 <b>(0.000)</b>
Hypertension Factor	Mid-30s	-0.599 <b>(0.039)</b>	-1.063 <b>(0.013)</b>	0.074 (0.841)	0.112 (0.797)	0.076 (0.870)	-0.961 <b>(0.000)</b>	-1.401 <b>(0.000)</b>	-1.386 <b>(0.000)</b>

Table C.90: Treatment Effects on Cholesterol, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	7.753 <b>(0.039)</b>	4.139 (0.289)	-0.267 (0.739)	-8.079 (0.986)	-4.138 (0.913)	9.015 <b>(0.026)</b>	5.332 (0.211)	6.830 <b>(0.079)</b>
Dyslipidemia	Mid-30s	-0.094 (0.342)	-0.029 (0.513)	0.200 (1.000)	0.365 (0.986)	0.207 (1.000)	-0.108 (0.329)	-0.056 (0.513)	-0.130 (0.316)
Cholesterol Factor	Mid-30s	-0.333 (0.105)	-0.153 (0.408)	0.240 (0.957)	0.657 (0.986)	0.359 (0.971)	-0.385 (0.145)	-0.219 (0.368)	-0.348 (0.184)

Table C.91: Treatment Effects on Diabetes, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hemoglobin Level (%)	Mid-30s	0.322 (0.961)	0.435 (0.908)	0.240 (0.942)	0.469 (0.797)	0.355 (0.957)	0.286 (0.961)	0.386 (0.908)	0.396 (0.961)
Prediabetes	Mid-30s	-0.129 (0.513)	-0.190 (0.329)	-0.267 (0.464)	-0.478 (0.275)	-0.338 (0.261)	-0.139 (0.513)	-0.182 (0.395)	-0.173 (0.461)
Diabetes	Mid-30s	0.080 (0.987)	0.081 (0.921)	0.080 (1.000)	0.027 (0.768)	0.098 (1.000)	0.080 (0.987)	0.083 (0.921)	0.098 (1.000)
Diabetes Factor	Mid-30s	0.283 (0.987)	0.338 (0.921)	0.225 (0.971)	0.244 (0.797)	0.307 (0.971)	0.262 (0.974)	0.317 (0.908)	0.345 (0.987)

Table C.92: Treatment Effects on Vitamin D Deficiency, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vitamin D Deficiency	Mid-30s	-0.245 <b>(0.066)</b>	-0.185 (0.132)	-0.480 <b>(0.000)</b>	-0.216 (0.158)	-0.485 <b>(0.000)</b>	-0.172 (0.158)	-0.145 (0.263)	-0.189 (0.158)

Table C.93: Treatment Effects on Obesity, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measured BMI	Mid-30s	-0.125 (0.763)	-0.617 (0.789)	-0.684 (0.594)	-3.505 (0.797)	0.867 (0.797)	-0.627 (0.776)	0.025 (0.855)	-0.487 (0.829)
Obesity	Mid-30s		-0.103 (0.724)	-0.128 (0.493)	-0.331 (0.406)	0.031 (0.725)	-0.017 (0.776)	-0.045 (0.763)	-0.060 (0.737)
Severe Obesity	Mid-30s	-0.160 (0.434)	-0.132 (0.513)	-0.185 (0.464)	-0.319 (0.522)	-0.126 (0.493)	-0.185 (0.368)	-0.108 (0.711)	-0.131 (0.553)
Waist-hip Ratio	Mid-30s	0.004 (0.842)	-0.014 (0.724)	0.018 (0.783)	-0.043 (0.739)	0.030 (0.870)	-0.002 (0.789)	-0.008 (0.763)	-0.004 (0.829)
Abdominal Obesity	Mid-30s	0.003 (0.803)	-0.137 (0.500)	0.029 (0.710)	-0.377 (0.580)	0.100 (0.826)	0.029 (0.882)	-0.062 (0.750)	-0.031 (0.816)
Framingham Risk Score	Mid-30s	-0.766 (0.592)	-0.472 (0.750)	1.491 (0.942)	2.437 (1.000)	1.802 (0.971)	-1.202 (0.408)	-1.036 (0.592)	-0.705 (0.658)
Obesity Factor	Mid-30s	-0.040 (0.763)	-0.197 (0.724)	-0.036 (0.652)	-0.524 (0.739)	0.211 (0.812)	-0.101 (0.750)	-0.091 (0.803)	-0.116 (0.750)

Table C.94: Treatment Effects on Mental Health, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Somatization	21	0.035 (0.895)	-0.071 (0.658)	0.076 (1.000)	-0.196 (0.539)	0.035 (0.961)	0.032 (0.895)	-0.091 (0.671)	-0.043 (0.803)
	34	-0.239 (0.513)	-0.333 (0.368)	-0.025 (0.895)	-0.110 (0.658)	-0.075 (0.645)	-0.341 (0.395)	-0.370 (0.474)	-0.331 (0.395)
Depression	21	-0.023 (0.816)	-0.085 (0.684)	0.239 (1.000)	0.171 (1.000)	0.174 (1.000)	-0.024 (0.816)	-0.113 (0.592)	-0.074 (0.737)
	34	-0.145 (0.671)	-0.247 (0.566)	0.216 (1.000)	0.370 (1.000)	0.151 (1.000)	-0.295 (0.566)	-0.373 (0.526)	-0.273 (0.645)
Anxiety	21	0.145 (1.000)	0.008 (0.908)	0.307 (1.000)	0.012 (0.947)	0.249 (1.000)	0.105 (0.974)	-0.042 (0.776)	0.019 (0.908)
	34	-0.191 (0.553)	-0.293 (0.368)	0.099 (1.000)	0.117 (1.000)	0.040 (1.000)	-0.312 (0.447)	-0.364 (0.395)	-0.301 (0.434)
Hostility	21	-0.099 (0.632)	-0.148 (0.566)	0.126 (1.000)	-0.129 (0.816)	0.081 (0.974)	-0.134 (0.645)	-0.151 (0.592)	-0.144 (0.658)
	34	-0.157 (0.566)	-0.264 (0.421)	0.067 (1.000)	0.019 (0.974)	-0.007 (0.921)	-0.260 (0.526)	-0.322 (0.526)	-0.273 (0.500)
Global Severity Index	21	0.045 (0.908)	-0.061 (0.684)	0.229 (1.000)	-0.038 (0.882)	0.169 (1.000)	0.016 (0.868)	-0.089 (0.684)	-0.058 (0.750)
	34	-3.450 (0.566)	-5.240 (0.421)	1.741 (1.000)	2.260 (1.000)	0.696 (0.987)	-5.693 (0.513)	-6.643 (0.461)	-5.427 (0.513)
BSI Factor	21 and 34	-0.186 (0.684)	-0.150 (0.763)	0.016 (0.961)	0.328 (1.000)	0.042 (0.961)	-0.261 (0.724)	-0.237 (0.776)	-0.109 (0.829)

## C.8 Treatment Effects for Female Sample, Step Down

Table C.95: Treatment Effects on IQ Scores, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. IQ Test	2	10.700	9.200	13.950	16.963	13.910	9.431	6.680	9.341
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.053)</b>	<b>(0.000)</b>
	3	13.333	12.563	23.729	29.326	24.246	9.211	7.275	9.177
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.026)</b>	(0.118)	<b>(0.039)</b>
	3.5	8.049	6.310	16.188	20.843	16.047	5.049	1.970	4.960
		<b>(0.039)</b>	(0.132)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.171)	(0.671)	(0.197)
	4	6.035	4.483	14.812	19.509	15.321	3.007	0.543	3.475
		<b>(0.079)</b>	(0.276)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.382)	(0.829)	(0.342)
	4.5	8.162	6.607	16.058	20.187	16.348	5.318	1.706	5.810
		<b>(0.039)</b>	(0.118)	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.184)	(0.711)	(0.158)
	5	4.921	3.762	12.425	17.533	12.947	2.698	-0.158	2.994
		<b>(0.092)</b>	(0.395)	<b>(0.026)</b>	<b>(0.013)</b>	<b>(0.026)</b>	(0.421)	(0.882)	(0.395)
	6.6	6.127	5.322	7.339	12.597	8.174	5.773	3.462	6.408
		<b>(0.079)</b>	(0.132)	<b>(0.053)</b>	<b>(0.053)</b>	<b>(0.026)</b>	(0.158)	(0.447)	(0.132)
7	6.365	4.279	7.796	13.116	7.861	5.992	2.304	5.487	
	<b>(0.079)</b>	(0.276)	<b>(0.079)</b>	<b>(0.026)</b>	<b>(0.066)</b>	(0.132)	(0.671)	(0.158)	
8	5.906	4.111	7.857	12.487	8.186	5.360	2.282	5.550	
	<b>(0.053)</b>	(0.276)	<b>(0.066)</b>	<b>(0.039)</b>	<b>(0.066)</b>	(0.184)	(0.658)	(0.158)	
12	8.688	7.857	6.850	9.960	6.441	9.120	6.952	8.429	
	<b>(0.026)</b>	<b>(0.053)</b>	<b>(0.066)</b>	<b>(0.039)</b>	<b>(0.092)</b>	<b>(0.026)</b>	(0.118)	<b>(0.053)</b>	
IQ Factor	2 to 5	0.644	0.529	1.317	1.705	1.345	0.437	0.167	0.457
	<b>(0.026)</b>	<b>(0.092)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	(0.132)	(0.671)	(0.132)	
	6 to 12	0.545	0.475	0.448	0.976	0.506	0.570	0.362	0.601
	<b>(0.079)</b>	(0.132)	<b>(0.092)</b>	<b>(0.039)</b>	<b>(0.092)</b>	<b>(0.079)</b>	(0.395)	<b>(0.079)</b>	

Table C.96: Treatment Effects on Achievement Scores, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Std. Achv. Test	5.5	12.314	10.088	19.650	19.146	20.936	9.869	4.876	11.584
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.079)</b>	(0.421)	<b>(0.000)</b>
	6	6.269	6.756	10.379	14.081	10.986	5.018	4.220	5.887
		<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.039)</b>	(0.105)	<b>(0.039)</b>
	7.5	4.133	3.185	4.300	9.725	4.578	4.082	1.201	4.326
		<b>(0.092)</b>	<b>(0.092)</b>	(0.171)	<b>(0.039)</b>	(0.171)	(0.105)	(0.671)	(0.105)
8	6.619	5.350	7.125	11.989	7.738	6.465	3.618	7.036	
	<b>(0.066)</b>	<b>(0.039)</b>	<b>(0.053)</b>	<b>(0.013)</b>	<b>(0.039)</b>	<b>(0.092)</b>	(0.408)	<b>(0.092)</b>	
8.5	8.407	7.607	12.299	15.146	12.625	7.223	5.279	7.918	
	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.053)</b>	(0.184)	<b>(0.053)</b>	
Achievement Factor	5.5 to 12	0.839	0.838	1.183	1.277	1.264	0.706	0.648	0.809
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.039)</b>	<b>(0.039)</b>	<b>(0.000)</b>	

Table C.97: Treatment Effects on HOME Scores, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
HOME Score	0.5	1.581 (0.184)	0.749 (0.737)	1.684 (0.158)	1.131 (0.526)	1.067 (0.316)	0.980 (0.553)	0.498 (0.803)	0.438 (0.711)
	1.5	2.668 <b>(0.066)</b>	1.723 (0.342)	4.729 <b>(0.066)</b>	3.279 (0.237)	4.661 <b>(0.079)</b>	1.544 (0.447)	0.992 (0.684)	1.746 (0.408)
	2.5	0.762 (0.684)	0.832 (0.737)	4.434 <b>(0.013)</b>	5.502 <b>(0.092)</b>	4.637 <b>(0.000)</b>	-0.899 (0.974)	-0.932 (0.974)	-0.256 (0.934)
	3.5	2.858 (0.329)	2.561 (0.513)	13.719 <b>(0.000)</b>	15.918 <b>(0.039)</b>	13.454 <b>(0.000)</b>	-0.309 (0.934)	-1.851 (0.987)	-0.058 (0.868)
	4.5	2.736 (0.355)	1.288 (0.763)	12.957 <b>(0.026)</b>	11.926 (0.132)	13.245 <b>(0.026)</b>	-0.273 (0.934)	-1.462 (0.974)	0.468 (0.842)
	8	0.659 (0.776)	0.721 (0.842)	5.909 <b>(0.079)</b>	6.478 (0.145)	6.484 <b>(0.066)</b>	-0.773 (0.961)	-1.027 (0.961)	0.328 (0.842)
	HOME Factor	0.5 to 8	0.262 (0.474)	0.220 (0.605)	1.215 <b>(0.039)</b>	1.001 (0.276)	1.272 <b>(0.039)</b>	-0.010 (0.921)	-0.053 (0.947)

Table C.98: Treatment Effects on Parent Income, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Parental Income	1.5	4,516 (0.145)	7,539 <b>(0.079)</b>	5,105 (0.237)	8,682 (0.211)	9,681 <b>(0.079)</b>	4,309 (0.250)	7,091 (0.105)	7,337 <b>(0.026)</b>
	2.5	222 (0.711)	400 (0.579)	-3,712 (0.934)	894 (0.671)	1,769 (0.500)	1,598 (0.526)	894 (0.513)	3,227 (0.329)
	3.5	2,756 (0.355)	4,043 (0.263)	4,469 (0.263)	7,429 (0.237)	8,577 (0.118)	2,125 (0.487)	2,120 (0.447)	3,764 (0.316)
	4.5	4,039 (0.118)	8,141 <b>(0.066)</b>	6,443 (0.197)	9,836 (0.171)	9,337 <b>(0.079)</b>	3,202 (0.276)	7,852 <b>(0.092)</b>	6,022 <b>(0.026)</b>
Parental Income Factor	1.5 to 15	0.193 (0.421)	0.209 (0.329)	0.192 (0.487)	0.291 (0.526)	0.284 (0.395)	0.193 (0.434)	0.163 (0.447)	0.283 (0.316)

Table C.99: Treatment Effects on Mother's Employment, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mother Works	2	0.168 <b>(0.066)</b>	0.124 (0.237)	0.323 (0.132)	0.292 (0.237)	0.309 (0.132)	0.101 (0.250)	0.082 (0.342)	0.096 (0.289)
	3	0.087 (0.395)	0.047 (0.592)	0.177 (0.237)	0.145 (0.539)	0.160 (0.250)	0.066 (0.395)	0.021 (0.724)	0.057 (0.487)
	4	0.118 (0.158)	0.080 (0.434)	0.319 (0.132)	0.275 (0.289)	0.305 (0.145)	0.060 (0.395)	0.038 (0.566)	0.053 (0.461)
	5	0.067 (0.487)	0.017 (0.697)	0.367 <b>(0.066)</b>	0.268 (0.197)	0.397 <b>(0.039)</b>	-0.056 (0.908)	-0.097 (0.974)	-0.024 (0.842)
Mother Works Factor	2 to 21	0.200 (0.434)	0.076 (0.671)	0.578 (0.184)	0.484 (0.447)	0.564 (0.184)	0.101 (0.566)	-0.037 (0.789)	0.108 (0.539)

Table C.100: Treatment Effects on Father at Home, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Father at Home	2	-0.012 (0.776)	0.015 (0.737)	-0.115 (0.961)	-0.166 (1.000)	-0.094 (0.921)	0.034 (0.632)	0.095 (0.342)	0.087 (0.355)
	3	-0.079 (0.934)	-0.038 (0.882)	-0.337 (1.000)	-0.330 (1.000)	-0.315 (1.000)	0.034 (0.632)	0.095 (0.342)	0.087 (0.355)
	4	-0.071 (0.921)	-0.042 (0.882)	-0.330 (1.000)	-0.350 (1.000)	-0.306 (1.000)	0.041 (0.618)	0.091 (0.382)	0.096 (0.355)
	5	-0.139 (0.987)	-0.101 (0.961)	-0.333 (1.000)	-0.278 (1.000)	-0.328 (1.000)	-0.056 (0.868)	-0.025 (0.816)	-0.020 (0.763)
	8	0.056 (0.592)	0.078 (0.447)	-0.064 (0.868)	0.003 (0.908)	-0.055 (0.842)	0.092 (0.368)	0.125 (0.303)	0.095 (0.355)
Father at Home Factor	2 to 8	-0.211 (0.947)	-0.094 (0.882)	-0.899 (1.000)	-0.818 (1.000)	-0.884 (1.000)	-0.002 (0.724)	0.143 (0.474)	0.090 (0.645)

Table C.101: Treatment Effects on Education, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Graduated High School	30	0.253 <b>(0.026)</b>	0.148 (0.224)	0.642 <b>(0.000)</b>	0.519 <b>(0.026)</b>	0.595 <b>(0.000)</b>	0.137 (0.316)	0.004 (0.868)	0.066 (0.697)
Attended Voc./Tech./Com. College	30	-0.057 (0.947)	-0.115 (1.000)	-0.050 (0.627)	-0.098 (0.947)	-0.071 (0.697)	-0.041 (0.947)	-0.145 (0.987)	-0.051 (0.974)
Graduated 4-year College	30	0.134 (0.158)	0.102 (0.408)	0.217 <b>(0.013)</b>	0.097 (0.408)	0.210 <b>(0.000)</b>	0.106 (0.408)	0.073 (0.658)	0.095 (0.434)
Years of Edu.	30	2.143 <b>(0.000)</b>	1.695 <b>(0.013)</b>	4.025 <b>(0.000)</b>	2.984 <b>(0.053)</b>	3.918 <b>(0.000)</b>	1.567 <b>(0.026)</b>	1.155 (0.211)	1.409 <b>(0.053)</b>
Education Factor	30	0.661 <b>(0.000)</b>	0.461 (0.145)	1.277 <b>(0.000)</b>	0.799 <b>(0.092)</b>	1.212 <b>(0.000)</b>	0.447 (0.118)	0.261 (0.566)	0.361 (0.224)

Table C.102: Treatment Effects on Employment and Income, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Employed	30	0.131 (0.316)	0.080 (0.632)	0.333 (0.118)	0.363 (0.171)	0.340 <b>(0.092)</b>	0.056 (0.750)	0.003 (0.947)	0.070 (0.724)
Labor Income	21	1,741 (0.632)	1,694 (0.737)	6,932 <b>(0.053)</b>	10,072 <b>(0.053)</b>	8,248 <b>(0.000)</b>	496 (0.882)	-779 (0.974)	440 (0.961)
	30	2,548 (0.789)	2,391 (0.868)	14,356 <b>(0.053)</b>	8,166 (0.395)	14,835 <b>(0.079)</b>	-425 (0.934)	-803 (0.947)	-840 (0.987)
Public-Transfer Income	21	-1,424 (0.368)	-1,929 (0.118)	-1,322 (0.395)	-1,786 (0.289)	-2,879 <b>(0.039)</b>	-1,751 (0.276)	-1,144 (0.632)	-1,475 (0.434)
	30	-2,672 (0.158)	-1,560 (0.553)	-3,053 (0.145)	-2,783 (0.171)	-3,213 <b>(0.092)</b>	-2,269 (0.342)	-1,169 (0.750)	-2,620 (0.447)
Employment Factor	21 to 30	0.281 (0.329)	0.132 (0.776)	0.658 <b>(0.039)</b>	0.557 (0.171)	0.675 <b>(0.026)</b>	0.167 (0.671)	0.000 (0.947)	0.149 (0.724)

Table C.103: Treatment Effects on Crime, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Total Felony Arrests	Mid-30s	-0.328 (0.237)	-0.390 (0.289)	-1.345 (0.158)	-1.011 (0.289)	-0.968 (0.197)	-0.077 (0.421)	-0.022 (0.671)	0.005 (0.737)
Total Misdemeanor Arrests	Mid-30s	-0.973 (0.105)	-0.883 (0.355)	-2.708 (0.132)	-2.177 (0.316)	-2.453 (0.184)	-0.588 (0.158)	-0.142 (0.645)	-0.200 (0.487)
Total Years Incarcerated	30	-0.024 (0.263)	-0.025 (0.276)				-0.037 (0.224)	-0.032 (0.224)	-0.038 (0.289)
Crime Factor	30 to Mid-30s	-0.155 (0.158)	-0.176 (0.289)	-0.508 (0.171)	-0.523 (0.289)	-0.504 (0.197)	-0.071 (0.224)	-0.026 (0.605)	-0.041 (0.408)

Table C.104: Treatment Effects on Tobacco, Drugs, Alcohol, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cig. Smoked per day last month	30	-0.765 (0.618)	0.000 (0.855)	-2.337 (0.395)	0.463 (0.895)	-2.131 (0.434)	-0.530 (0.724)	0.253 (0.895)	-0.302 (0.803)
Days drank alcohol last month	30	-0.742 (0.645)	0.085 (0.882)	-0.567 (0.789)	0.274 (0.842)	-0.261 (0.908)	-0.918 (0.605)	0.043 (0.829)	-0.464 (0.763)
Days binge drank alcohol last month	30	-0.358 (0.645)	0.216 (0.947)	-1.062 (0.711)	-0.070 (0.776)	-0.922 (0.711)	-0.232 (0.724)	0.455 (0.974)	0.034 (0.855)
Self-reported drug user	Mid-30s	-0.033 (0.737)	-0.079 (0.605)	-0.116 (0.645)	-0.196 (0.368)	-0.101 (0.671)	-0.101 (0.816)	-0.043 (0.711)	0.033 (0.868)
Substance Use Factor	30 to Mid-30s	0.006 (0.842)	0.317 (1.000)	0.308 (1.000)	0.497 (0.987)	0.313 (1.000)	-0.076 (0.737)	0.264 (0.987)	0.067 (0.868)

Table C.105: Treatment Effects on Self-Reported Health, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Self-reported Health	30	-0.184 (0.461)	-0.215 (0.421)	-1.150 <b>(0.039)</b>	-0.988 <b>(0.079)</b>	-1.188 <b>(0.039)</b>	0.044 (0.961)	-0.027 (0.763)	-0.016 (0.855)
	Mid-30s	0.006 (0.855)	-0.108 (0.737)	-0.232 (0.447)	-0.171 (0.408)	-0.252 (0.421)	0.071 (0.974)	-0.087 (0.711)	0.010 (0.934)
Self-reported Health Factor	30 to Mid-30s	0.045 (0.908)	0.001 (0.868)	0.203 (0.882)	0.252 (0.934)	0.180 (0.921)	0.002 (0.921)	-0.071 (0.711)	-0.011 (0.855)

Table C.106: Treatment Effects on Hypertension, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Systolic Blood Pressure (mm Hg)	Mid-30s	-2.899 (0.447)	-3.901 (0.421)	1.065 (0.829)	2.263 (0.789)	-0.577 (0.763)	-3.980 (0.395)	-4.164 (0.421)	-6.433 (0.303)
Diastolic Blood Pressure (mm Hg)	Mid-30s	-0.002 (0.671)	1.317 (0.697)	4.725 (0.961)	7.497 (0.921)	4.337 (0.947)	-1.290 (0.500)	0.423 (0.566)	-1.783 (0.487)
Prehypertension	Mid-30s	-0.189 <b>(0.079)</b>	-0.224 <b>(0.092)</b>	-0.094 (0.618)	-0.194 (0.342)	-0.160 (0.447)	-0.215 <b>(0.013)</b>	-0.248 <b>(0.053)</b>	-0.275 <b>(0.013)</b>
Hypertension	Mid-30s	0.172 (0.974)	0.159 (0.961)	0.232 (0.987)	0.156 (0.908)	0.181 (0.974)	0.156 (0.947)	0.193 (0.908)	0.083 (0.829)
Hypertension Factor	Mid-30s	-0.033 (0.618)	-0.035 (0.566)	0.204 (0.961)	0.233 (0.868)	0.125 (0.908)	-0.097 (0.474)	-0.053 (0.553)	-0.202 (0.408)

Table C.107: Treatment Effects on Cholesterol, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High-Density Lipoprotein Chol. (mg/dL)	Mid-30s	2.884 (0.368)	5.555 (0.171)	10.514 <b>(0.013)</b>	6.966 (0.303)	13.319 <b>(0.013)</b>	0.802 (0.671)	4.446 (0.303)	4.243 (0.316)
Dyslipidemia	Mid-30s	0.051 (0.961)	0.043 (0.829)	-0.080 (0.222)	-0.072 (0.566)	-0.093 (0.250)	0.087 (1.000)	0.101 (0.987)	0.096 (1.000)
Cholesterol Factor	Mid-30s	-0.024 (0.645)	-0.111 (0.487)	-0.396 <b>(0.092)</b>	-0.284 (0.421)	-0.492 <b>(0.053)</b>	0.078 (0.908)	-0.011 (0.697)	-0.011 (0.671)

Table C.108: Treatment Effects on Diabetes, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Hemoglobin Level (%)	Mid-30s	-0.277 (0.368)	-0.101 (0.513)	-0.176 (0.184)	-0.088 (0.447)	-0.190 (0.197)	-0.304 (0.421)	-0.037 (0.711)	-0.355 (0.355)
Prediabetes	Mid-30s	0.088 (0.895)	0.163 (0.961)	0.076 (0.724)	0.176 (0.816)	0.035 (0.645)	0.091 (0.895)	0.161 (0.947)	0.029 (0.750)
Diabetes	Mid-30s	-0.071 (0.145)	-0.032 (0.342)				-0.091 (0.184)	-0.039 (0.368)	-0.095 (0.118)
Diabetes Factor	Mid-30s	-0.249 (0.237)	-0.086 (0.474)	-0.086 (0.408)	-0.029 (0.513)	-0.098 (0.316)	-0.294 (0.276)	-0.064 (0.618)	-0.334 (0.224)

Table C.109: Treatment Effects on Vitamin D Deficiency, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Vitamin D Deficiency	Mid-30s	-0.047 (0.303)	0.061 (0.658)	-0.094 (0.316)	-0.007 (0.368)	-0.034 (0.434)	-0.034 (0.355)	0.093 (0.711)	0.038 (0.553)

Table C.110: Treatment Effects on Obesity, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measured BMI	Mid-30s	3.545 (1.000)	5.649 (1.000)	1.937 (0.961)	3.117 (0.987)	1.969 (0.987)	3.983 (1.000)	6.692 (1.000)	4.705 (1.000)
Obesity	Mid-30s	-0.011 (0.921)	0.087 (0.974)	-0.261 <b>(0.039)</b>	-0.115 (0.592)	-0.199 (0.145)	0.057 (0.974)	0.178 (1.000)	0.109 (0.987)
Severe Obesity	Mid-30s	-0.045 (0.816)	0.065 (0.961)	0.015 (0.895)	0.076 (0.987)	0.019 (0.895)	-0.061 (0.829)	0.064 (0.961)	-0.040 (0.816)
Waist-hip Ratio	Mid-30s	-0.022 (0.697)	0.004 (0.882)	-0.076 (0.211)	-0.076 (0.526)	-0.072 (0.329)	-0.007 (0.882)	0.036 (1.000)	0.005 (0.961)
Abdominal Obesity	Mid-30s	-0.159 (0.487)	-0.016 (0.842)	-0.381 <b>(0.000)</b>	-0.207 (0.526)	-0.294 <b>(0.026)</b>	-0.095 (0.724)	0.054 (0.947)	-0.029 (0.842)
Framingham Risk Score	Mid-30s	-0.259 (0.368)	-0.228 (0.447)	-0.488 (0.289)	-0.551 (0.342)	-0.525 (0.329)	-0.197 (0.645)	-0.078 (0.750)	-0.220 (0.539)
Obesity Factor	Mid-30s	0.088 (0.987)	0.303 (1.000)	-0.289 (0.513)	-0.133 (0.842)	-0.267 (0.566)	0.196 (0.974)	0.474 (1.000)	0.255 (1.000)

Table C.111: Treatment Effects on Mental Health, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Somatization	21	-0.059 (0.632)	-0.170 (0.605)	-0.281 (0.355)	-0.431 (0.197)	-0.370 (0.289)	-0.006 (0.868)	-0.109 (0.711)	-0.100 (0.711)
	34	0.003 (0.789)	0.033 (0.974)	-0.199 (0.632)	-0.404 (0.526)	-0.198 (0.776)	0.058 (0.947)	0.110 (0.987)	0.075 (1.000)
Depression	21	-0.475 <b>(0.013)</b>	-0.483 (0.158)	-0.818 <b>(0.039)</b>	-0.905 <b>(0.092)</b>	-0.814 <b>(0.039)</b>	-0.346 (0.132)	-0.359 (0.368)	-0.317 (0.289)
	34	-0.125 (0.579)	-0.180 (0.421)	-0.149 (0.671)	-0.246 (0.605)	-0.226 (0.724)	-0.118 (0.776)	-0.217 (0.368)	-0.179 (0.645)
Anxiety	21	-0.387 <b>(0.079)</b>	-0.365 (0.211)	-0.643 <b>(0.092)</b>	-0.732 <b>(0.092)</b>	-0.606 (0.118)	-0.273 (0.250)	-0.240 (0.539)	-0.223 (0.539)
	34	-0.245 (0.316)	-0.218 (0.263)	-0.342 (0.474)	-0.474 (0.395)	-0.389 (0.474)	-0.218 (0.539)	-0.177 (0.526)	-0.265 (0.329)
Hostility	21	-0.475 <b>(0.053)</b>	-0.602 (0.132)	-1.036 <b>(0.039)</b>	-1.218 <b>(0.000)</b>	-1.110 <b>(0.039)</b>	-0.240 (0.539)	-0.403 (0.368)	-0.305 (0.461)
	34	-0.080 (0.632)	-0.058 (0.816)	-0.109 (0.632)	-0.217 (0.553)	-0.132 (0.711)	-0.072 (0.816)	-0.018 (0.921)	-0.017 (0.882)
Global Severity Index	21	-0.335 <b>(0.079)</b>	-0.351 (0.158)	-0.615 <b>(0.039)</b>	-0.665 <b>(0.013)</b>	-0.620 <b>(0.039)</b>	-0.234 (0.276)	-0.247 (0.500)	-0.225 (0.368)
	34	-2.197 (0.566)	-2.191 (0.605)	-4.138 (0.605)	-6.745 (0.513)	-4.884 (0.671)	-1.668 (0.789)	-1.704 (0.658)	-2.210 (0.684)
BSI Factor	21 and 34	-0.641 <b>(0.079)</b>	-0.347 (0.487)	-0.847 (0.276)	-0.799 (0.342)	-0.674 (0.408)	-0.583 <b>(0.092)</b>	-0.243 (0.658)	-0.396 (0.355)

## D Alternative Evaluation Methodologies

In the main paper, we develop and document the results from two methodologies to account for treatment substitution in the first phases of ABC and CARE. The objective of each method is to estimate the average treatment effect of ABC and CARE, holding fixed take-up of alternative preschool by the control groups. We want to construct a scenario in which the subjects in the control group do not attend alternative preschool. This allows us to evaluate center-based childcare relative to a counterfactual scenario in which subjects stay at home.

This section presents alternative methodologies to evaluate the first-phase treatment. Throughout the rest of this section, we refer to treatment in either ABC or CARE generically as center-based childcare, as we discard the family education treatment group of CARE—as in the main paper, the control group consists of the control groups in both ABC and CARE. We refer to take-up of treatment substitution as enrollment into alternative preschool. As in our main methodology, we choose the control sets to account for background variables using the method in Appendix B.

To illustrate the alternative methodologies, we present results for a small set of outcomes. First, we consider a time series of IQ scores. Second, we present results for three long-term outcomes: years of education, employment, and labor income. Both IQ scores and the chosen long-term outcomes follow predictable patterns and are straightforward to interpret. These characteristics allow us to evaluate the sensitivity of the estimated treatment effects using the different strategies.

Let the discrete choice to enroll in alternative preschool,  $P(\omega)$ , be defined as  $P(\omega) = \mathbf{1}[V(\omega) > 0]$  which is an indicator equal to 1 when the subject is enrolled. Let  $Q(\omega)$ , the

number of months in alternative preschool, be defined as  $V(\omega)T$ .  $\mathbf{X}(\omega)$  is a vector of observed individual- or household-level characteristics.  $P(\omega)$  and  $Q(\omega)$  are zero for participants of the treatment group. In this section, we allow for the take-up of the program to be imperfectly given by the randomization. Let  $D(\omega)$  be take-up of the ABC or CARE program.

## D.1 Instrumental Variables

The first alternative method we present is based on a standard, linear instrumental variable framework.

### D.1.1 Model and Conditions

Consider the following equation for an outcome  $Y(\omega)$ :

$$Y(\omega) = \alpha^D D(\omega) + \alpha^Q Q(\omega) + \mathbf{X}(\omega)\beta + \varepsilon(\omega), \quad (1)$$

where  $D(\omega)$  and  $Q(\omega)$  are endogenous. Selection into treatment or alternative preschool is likely correlated with characteristics not observed when estimating the coefficients characterizing the outcome equation, i.e.,  $Cov(D(\omega), \varepsilon(\omega)) \neq 0$  and  $Cov(Q(\omega), \varepsilon(\omega)) \neq 0$ . Estimating the coefficients in (1) by OLS yields inconsistent estimates.

A standard solution is to introduce a vector of instrumental variables,  $\mathbf{Z}(\omega)$ , satisfying two conditions: (i) the matrix  $\Pi$  of the coefficients in the population regression of  $D(\omega)$  and  $Q(\omega)$  on  $\mathbf{Z}(\omega)$  is full rank; and (ii)  $\mathbf{Z}(\omega)$  is uncorrelated with  $\varepsilon(\omega)$ .

### D.1.2 Instrumental Variables in Practice

To identify the coefficients in (1) using an instrumental variable strategy we need at least two instruments. We use randomization into ABC or CARE treatment as an instrument, denoted as an indicator,  $R(\omega)$ . We then consider up to three other instruments.

- (i) *Season of subject's birth.* This variable is coded as an indicator for being born in the fall (born between June and November). If preschools accepted children primarily in the fall, then children born in the fall could enter preschool when they were younger. Although there is evidence in the economics literature that being born during specific times in the year can influence a child's outcomes, the main channel for this effect seems to be the disproportionate selection into season of birth by high-income mothers (Buckles and Hungerman, 2013). Given that the mothers were very young and disadvantaged in the ABC and CARE samples, we assume that a negligible proportion of them selected their quarter of birth.
- (ii) *Presence of a grandmother in the county.* We hypothesize that the grandmothers of the subjects, who in this sample are relatively young and still working, could help the study mothers take care of their children. We assume that the mothers would have less influence over the presence of a grandmother in the county as opposed to the presence of a grandmother living in the home. We also assume that the grandmothers do not affect the subjects differently than any other avenue of informal care, such as care from a neighbor. In practice, the presence of a grandmother has a positive effect on preschool attendance, which could indicate that they helped take the subject to preschool. This variable is only available for ABC.
- (iii) *Number of relatives living in the household apart from the mother, subject, siblings, and male partner.* We assume that the relatives could take care of the subject while the mother was at work or in school, but that they do not affect the subjects differently than any other avenue of informal care. In practice, having additional relatives living at home decreases the probability that the subject goes into preschool.

### D.1.3 Sets of Instruments

We consider three combinations of the instruments. All of them include randomization into center-based treatments and birth in the fall. All of them include at least one variable representing access to informal care from relatives. The sets are the following:

- (i) Randomization, Born in the Fall, Grandmother in County, Number of Relatives at Home
- (ii) Randomization, Born in the Fall, Grandmother in County
- (iii) Randomization, Born in the Fall, Number of Relatives at Home

### D.1.4 Specifications for the Instruments

We test various specifications in which we allow for interactions of the potential instruments with the first-phase randomization indicator. In particular, we test three specifications for our instruments:

- (i) Instruments measured in levels:  $(R(\omega), \mathbf{Z}(\omega))$ .
- (ii) Instruments measured in levels and interactions of treatment with the instruments and the controls:  $(R(\omega), \mathbf{Z}(\omega), \mathbf{Z}(\omega)R(\omega), \mathbf{X}(\omega)R(\omega))$ .
- (iii) Interactions of the instruments and the controls with an indicator for being randomized into the control group:  $(R(\omega), \mathbf{X}(\omega)(1 - R(\omega)), \mathbf{Z}(\omega)(1 - R(\omega)))$ .

In practice, the specification in (iii) is the most stable across outcomes. This makes economic sense because the instruments are less likely to affect the participants of the treatment group, given that almost all the treatment families comply to the first-phase randomization protocol.

### D.1.5 Functional Forms of Enrollment in Alternative Preschool

We use a different parameterization of enrollment into alternative preschool in (1):

- (i) The number of months in alternative preschool,  $Q(\omega)$ .
- (ii) An indicator for take-up of alternative preschool,  $P(\omega)$ .
- (iii) The log of months in alternative preschool,  $\log Q(\omega)$ .

## D.2 Results

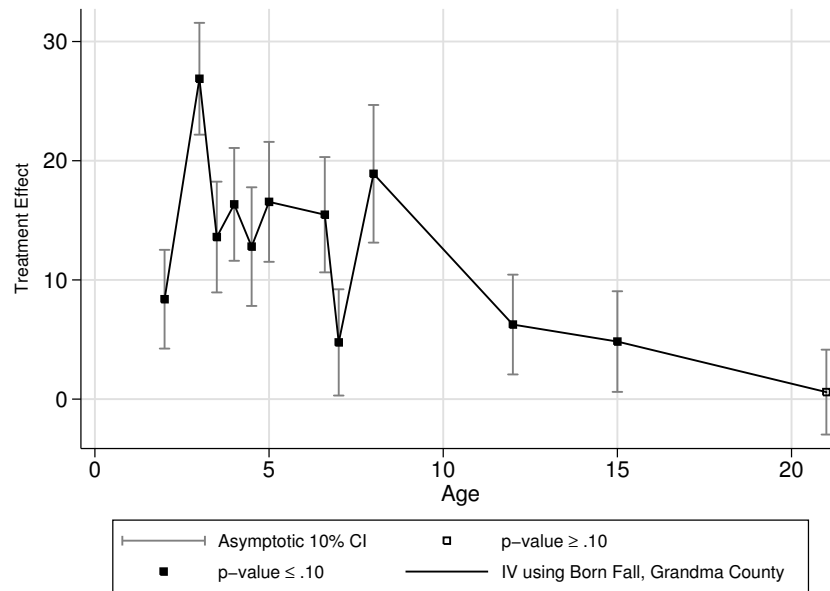
In this section, we present the results of the instrumental variable approach. We discuss the estimates of the coefficient for  $D(\omega)$  in (1), while accounting for endogenous take-up of alternative preschool. The results are roughly stable for all presented outcomes: the effect considering the take-up of alternative preschool in the control group is much stronger than the intent-to-treat effect (the mean difference between the treatment and control groups). At ages 15 and 21, the effects on IQ scores are close to zero.

### D.2.1 Main Specification

Our main specification uses three instruments: randomization into center-based childcare in either ABC or CARE, the presence of a grandmother in the county, and being born in the fall. We interact the latter two instrument with an indicator for being randomized into the control group ( $1 - R(\omega)$ ). The endogenous variables are  $D(\omega)$  and  $Q(\omega)$ .

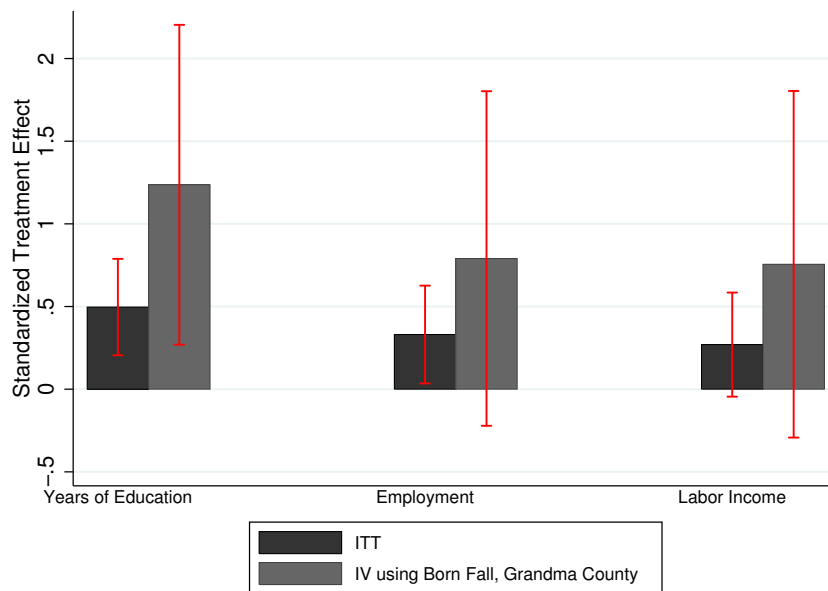
Figure D.1 and Figure D.2 display the estimates of the coefficient of  $D(\omega)$ . That is, the effect of attending the ABC or CARE treatment, fixing take-up of alternative preschool. The results indicate the following: (i) the effects are stronger compared to those in the main paper, even compared to those in the main paper that fix subjects to no preschool alternatives; (ii) the estimates are less stable across ages compared to those in the main paper. For example, while the effects on IQ scores from ages 0 to 5 average around 14 points in the main paper, they average 20 points using this specification of instrumental variables.

Figure D.1: Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Instrumental Variables



Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$  and  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are IQ tests at different ages, with a national standard deviation of 15 and a mean of 100.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.2: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Instrumental Variables

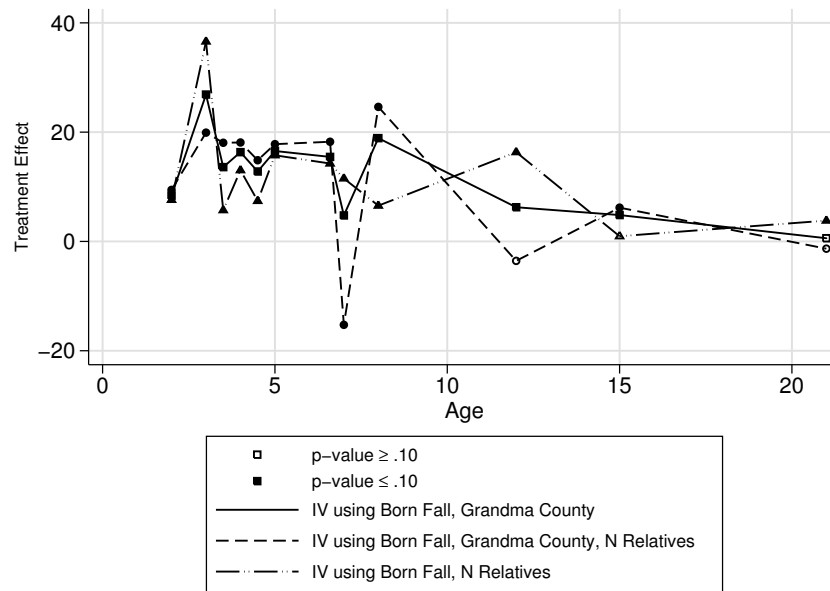


Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$  and  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are different adult outcomes labeled in the horizontal axis.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

### D.2.2 Varying the Sets of Instruments

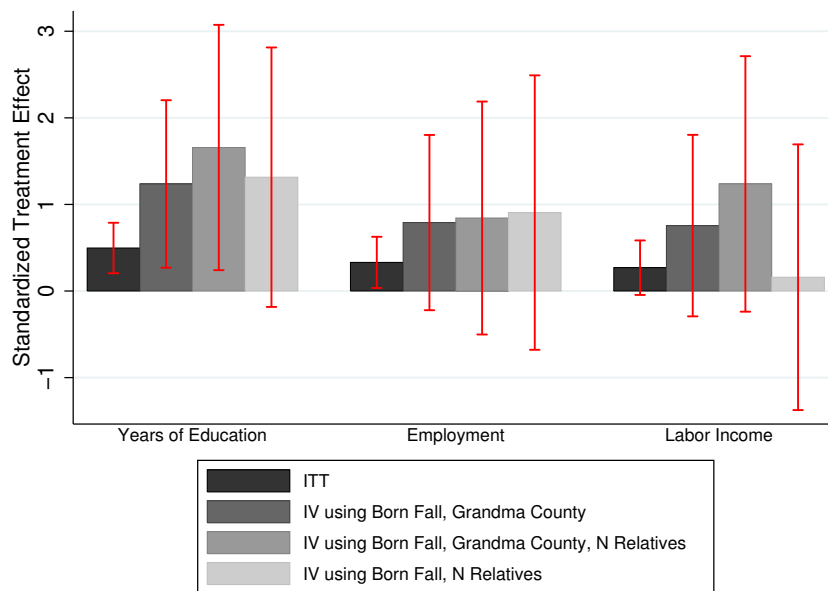
Figure D.3 and Figure D.4 explore the sensitivity of the estimates to different sets of instruments. The pattern of results indicates that the method is generally robust to the three sets of instrumental variables that we consider.

Figure D.3: Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables



Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$  and  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are IQ scores at different ages, with a national standard deviation of 15 and a mean of 100.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.4: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables



Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$  and  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are different adult outcomes labeled in the horizontal axis.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother’s IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

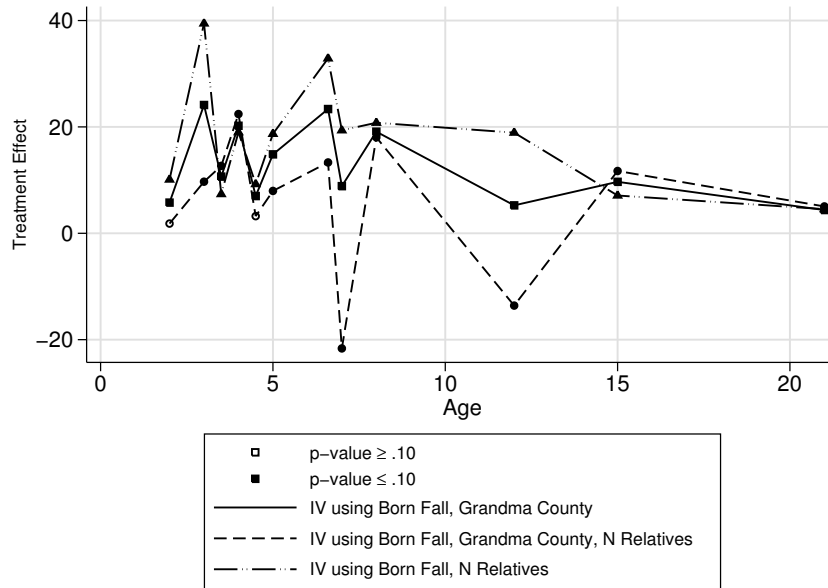
### D.2.3 Varying the Specification of the Instruments

We now present an exercise to evaluate the sensitivity of the results to different specifications of the instrumental variables. First, Figure D.5 and Figure D.6 present the results using the set of instruments that are not interacted with an indicator for randomization into the control group ( $1 - R(\omega)$ ). Figure D.7 and Figure D.8 present results not only interacting the instruments but also interacting the observed characteristics we control for. In both exercises, we use  $Q(\omega)$  as the endogenous variable, along with  $D(\omega)$ .

The results follow the same patterns as before, although they change when the instruments are not interacted. This makes economic sense because the interacted instruments better represent the economic intuition we offer before: the instruments other than  $R(\omega)$  are more

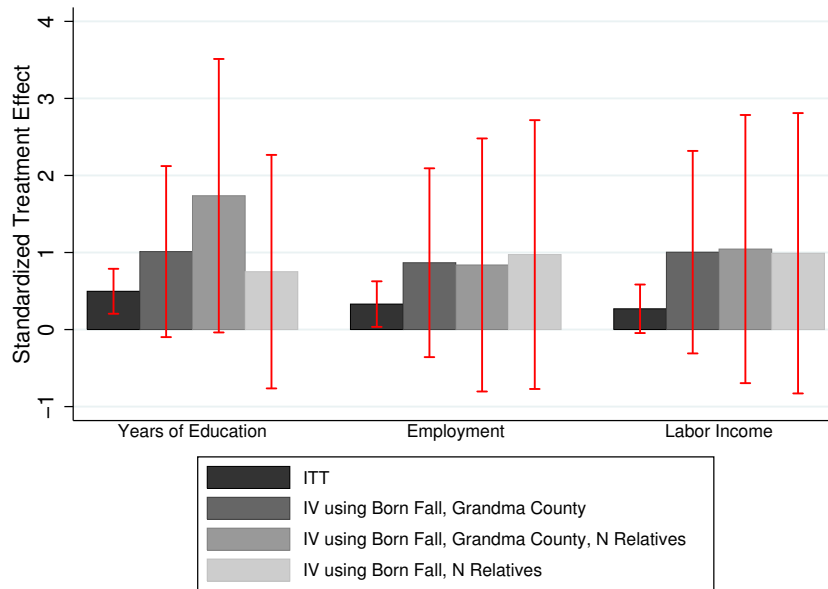
likely to shift the decisions of the families of the control-group subjects compared to those of the treatment-group subjects.

Figure D.5: Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications



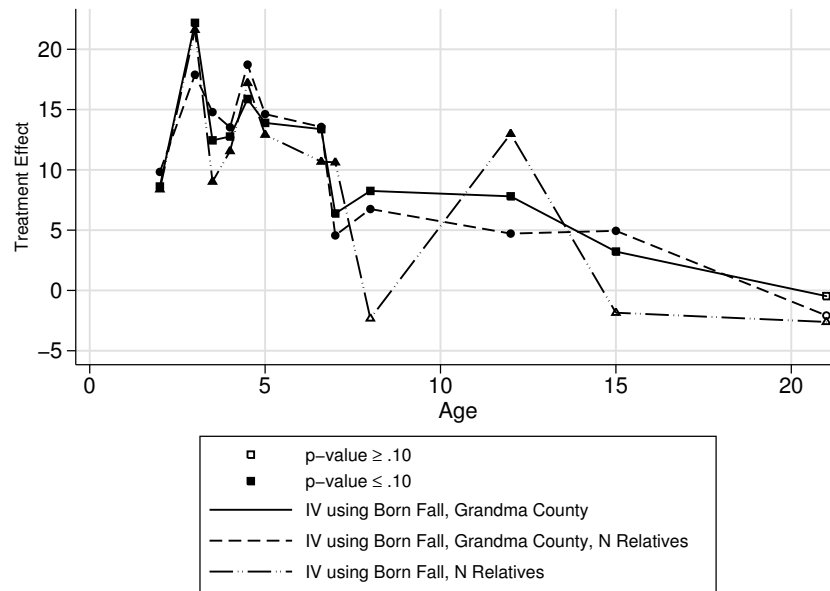
Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$  and  $\mathbf{Z}(\omega)$  as instruments.  $Y(\omega)$  is different IQ tests, with a national standard deviation of 15 and a mean of 100.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.6: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications



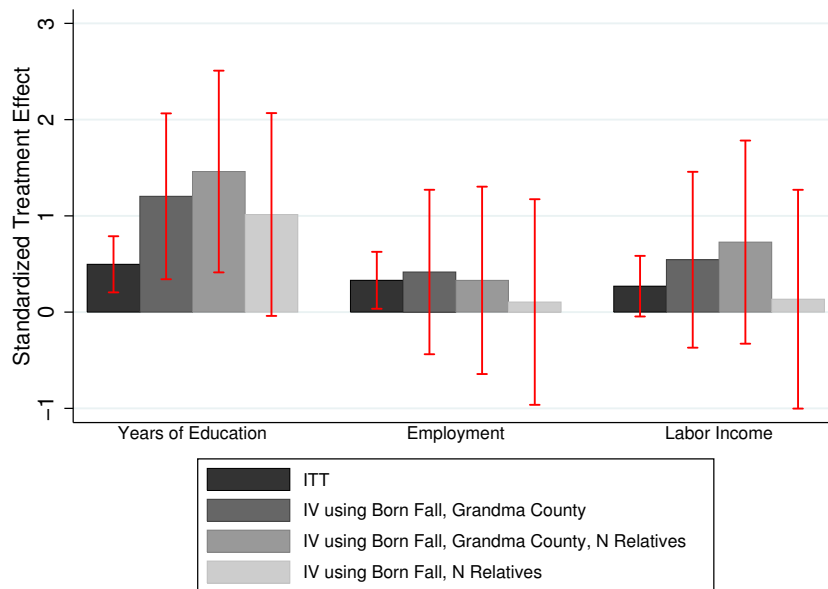
Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$  and  $\mathbf{Z}(\omega)$  as instruments. The outcomes ( $Y(\omega)$ ) are different adult outcomes labeled in the horizontal axis.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.7: Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications



Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$ ,  $\mathbf{X}(\omega)(1 - R(\omega))$  and  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are IQ scores at different ages, with a national standard deviation of 15 and a mean of 100.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.8: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous Take-up of Alternative Preschool Using Various Instrumental Variables Specifications

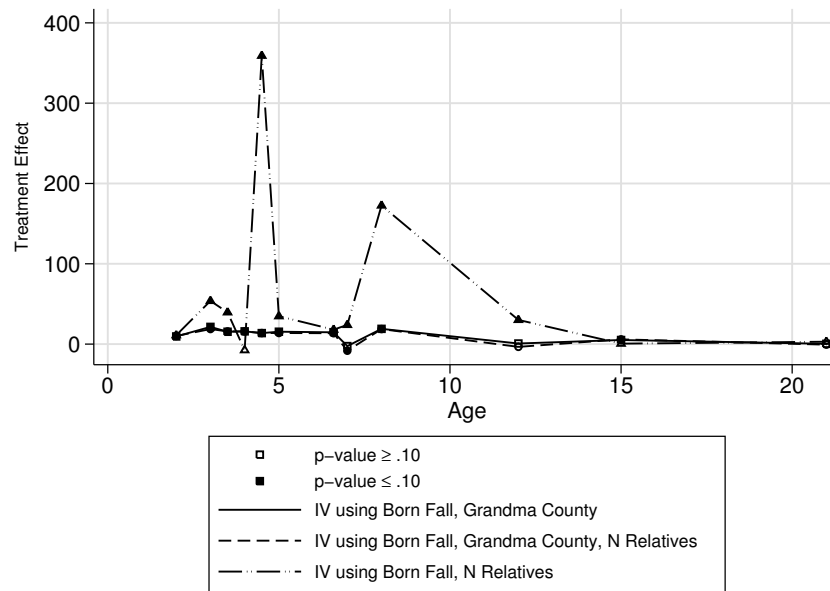


Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$ ,  $\mathbf{X}(\omega)(1 - R(\omega))$  and  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are different adult outcomes labeled in the horizontal axis.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

#### D.2.4 Varying the Parameterization of Alternative Preschool Take-up

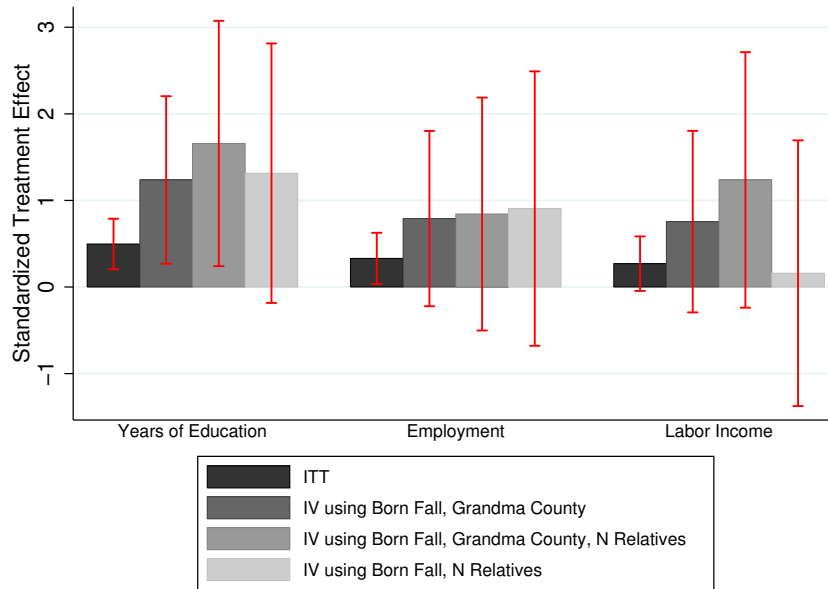
Now, we explore the sensitivity to the specification of  $Q(\omega)$  in (1). We consider two alternatives. First, we use an indicator of take-up of alternative preschool,  $P(\omega)$  (Figure D.9 and Figure D.10). Second, we take the log of  $Q(\omega)$  (Figure D.11 and Figure D.12).

Figure D.9: Effect of Center-based Childcare on IQ Scores, Accounting for an Endogenous Indicator of Take-up of Alternative Preschool



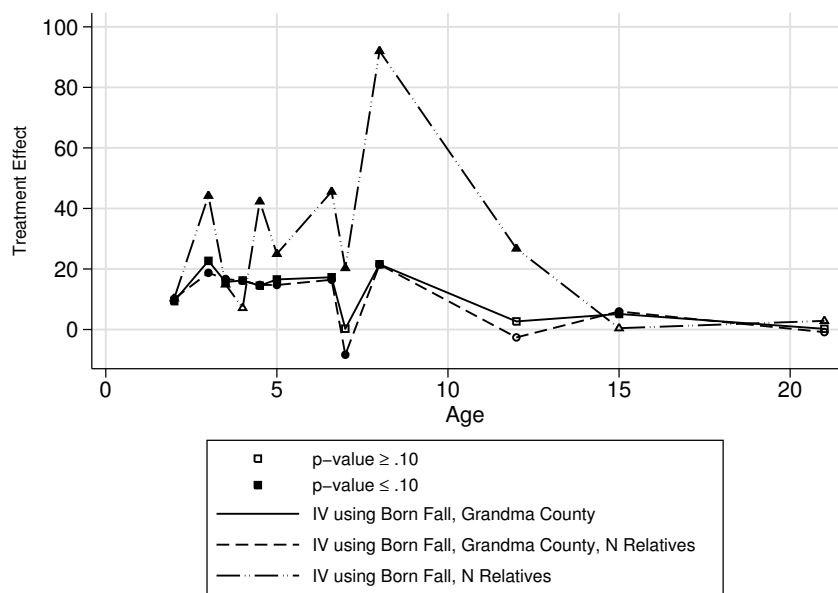
Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $P(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$ ,  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are IQ scores at different ages, with a national standard deviation of 15 and a mean of 100.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.10: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for an Endogenous Indicator of Take-up of Alternative Preschool



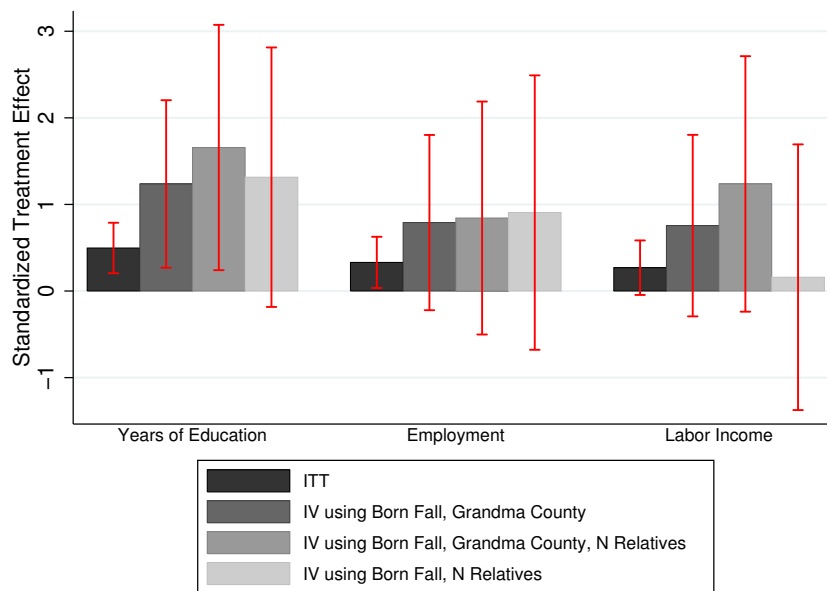
Note: This plot presents the parameter associated to  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $P(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$ ,  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are different relevant adult outcomes labeled in the horizontal axis.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.11: Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool



Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $\log Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$ ,  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are IQ scores at different ages, with a national standard deviation of 15 and a mean of 100.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

Figure D.12: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool



Note: This plot presents the parameter associated with  $D(\omega)$  from a regression of  $Y(\omega)$  on  $D(\omega)$ ,  $\log Q(\omega)$ , and  $\mathbf{X}(\omega)$ , using  $R(\omega)$ ,  $\mathbf{Z}(\omega)(1 - R(\omega))$  as instruments. The outcomes ( $Y(\omega)$ ) are different adult outcomes labeled in the horizontal axis.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother’s IQ score, High-risk Index, and Apgar Score at 1 minute. The confidence intervals are calculated at the 10% significance level.

### D.3 Control Functions

We now consider a control function approach. With control functions, the objective is also to simultaneously account for take-up of center-based childcare and alternative preschool.

#### D.3.1 Setup

The method we propose is an application of the selection correction in Heckman (1979). We model the selection into both endogenous variables of interest, center-based childcare and alternative preschool. The method involves three equations: (i) the outcome equation; (ii) the probability of participating in center-based childcare; (iii) a linear equation describing the number of months enrolled in preschool alternatives.

Let  $Y^0(\omega)$  be the counterfactual outcome of subject  $i$  when not participating in center-based childcare. Similarly, let  $Y^1(\omega)$  be her potential outcome if she participates. We model the outcome as:

$$\begin{aligned} Y^1(\omega) &= \alpha^1 + \mathbf{X}(\omega)\beta + \varepsilon^1(\omega) \\ Y^0(\omega) &= \alpha^0 + \mathbf{X}(\omega)\beta + \alpha^Q Q(\omega) + \varepsilon^0(\omega). \end{aligned} \quad (2)$$

The equation describing participation in center-based childcare is:

$$D(\omega) = \begin{cases} 0 & \text{if } D(\omega)^* \leq 0 \\ 1 & \text{if } D(\omega)^* > 0, \end{cases} \quad (3)$$

where we interpret  $D(\omega)^*$  as a latent continuous variable representing the household's interest in sending the subject to treatment. We write

$$D^*(\omega) = \mathbf{W}(\omega)\gamma^D + \varepsilon^D(\omega), \quad (4)$$

where  $\mathbf{W}(\omega)$  is a vector that includes  $\mathbf{X}(\omega)$  and  $R(\omega)$  and can include variables that shift the decision to enroll subjects into ABC or CARE without shifting the counterfactual outcome of interest,  $Y^d(\omega)$ .

We model the selection into months of alternative preschool as a linear equation with fixed coefficients, assuming homogeneous treatment effects:

$$Q(\omega) = \mathbf{W}(\omega)\gamma^Q + \varepsilon^Q(\omega), \quad (5)$$

In general, the unobserved variables in each of these equations are correlated. We assume that they are distributed as follows:

$$\begin{bmatrix} \varepsilon(\omega)^1 \\ \varepsilon(\omega)^0 \\ \varepsilon(\omega)^D \end{bmatrix} \sim \mathcal{N} \left[ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{1,0} & \sigma_{1,D} \\ \sigma_{1,0} & \sigma_0^2 & \sigma_{0,D} \\ \sigma_{1,D} & \sigma_{0,D} & 1 \end{pmatrix} \right], \quad (6)$$

where we normalize  $\text{Var}(\varepsilon^D(\omega)) = 1$ .

Further, we assume that

$$\mathbb{E} [\varepsilon(\omega)^0 | D(\omega) = 0, \mathbf{W}(\omega), \varepsilon^Q(\omega), Q(\omega) = q] = \sigma^{0,Q} \varepsilon^Q(\omega) + \mathbb{E} [\varepsilon(\omega)^0 | D(\omega) = 0, \mathbf{W}(\omega)]. \quad (7)$$

### D.3.2 Identification

The following steps identify the parameters of interest. First, we estimate the parameters characterizing the decision to enroll the subject in center-based childcare. We exploit the assumption that  $\varepsilon(\omega)^D \sim \mathcal{N}(0, 1)$  in (6) and estimate the parameters in (4) using a probit model.

Second, we approximate the unobserved term relevant to the choice of  $Q(\omega)$ . We take the coefficients in (5) to obtain an estimate for  $\varepsilon^Q(\omega)$ . By linearly conditioning on this term, we account for the correlation between the error term in the decision for  $Q(\omega)$  and the error term in the outcome equation,  $\varepsilon(\omega)^0$ .

Third, we estimate the coefficients in the outcome equation using the proxies for the unobserved components. We rewrite (2) using conditional expectations:

$$\begin{aligned} \mathbb{E} [Y(\omega)^1 | D(\omega) = 1, \mathbf{W}(\omega)] &= \alpha^1 + \mathbf{X}(\omega)\beta + \mathbb{E} [\varepsilon(\omega)^1 | D(\omega) = 1, \mathbf{W}(\omega)] \\ \mathbb{E} [Y(\omega)^0 | D(\omega) = 0, \mathbf{W}(\omega), \varepsilon^Q(\omega), Q(\omega) = q] &= \alpha^0 + \mathbf{X}(\omega)\beta + \alpha^Q q \\ &\quad + \mathbb{E} [\varepsilon(\omega)^0 | D(\omega) = 0, \mathbf{W}(\omega), \varepsilon^Q(\omega), Q(\omega) = q]. \end{aligned} \quad (8)$$

Once we condition on the proxy for  $\varepsilon^Q(\omega)$ , the error term in the outcome equations only depends on the selection into center-based childcare. The conditional error terms in (8) can be specified using control functions.

For subjects enrolled in treatment, the control function is:

$$\mathbb{E} [\varepsilon(\omega)^1 | D(\omega) = 1, \mathbf{W}(\omega)] = \sigma_1 \frac{\phi(\mathbf{W}(\omega)\gamma^D)}{\Phi(\mathbf{W}(\omega)\gamma^D)}. \quad (9)$$

For subjects not enrolled in the treatment, the control function is:

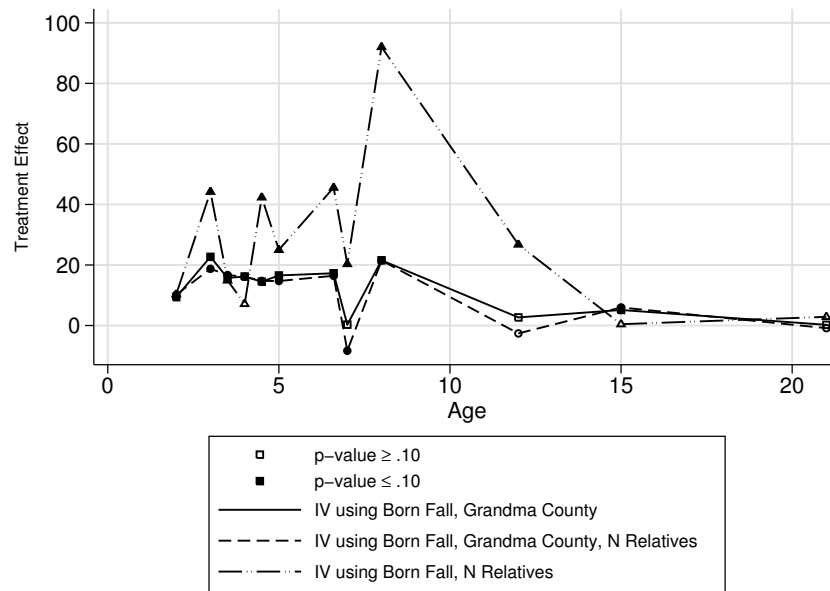
$$\mathbb{E} [\varepsilon(\omega)^0 | D(\omega) = 0, \mathbf{W}(\omega), \varepsilon^Q(\omega), Q(\omega) = q] = \sigma^{0,Q} \varepsilon^Q(\omega) - \sigma_0 \frac{\phi(\mathbf{W}(\omega)\gamma^D)}{\Phi(-\mathbf{W}(\omega)\gamma^D)}. \quad (10)$$

## D.4 Estimates

By including the control functions, we can recover consistent estimates of the parameters in (2) through a linear regression. The effect of center-based childcare is the difference of the intercepts in the two outcome equations.

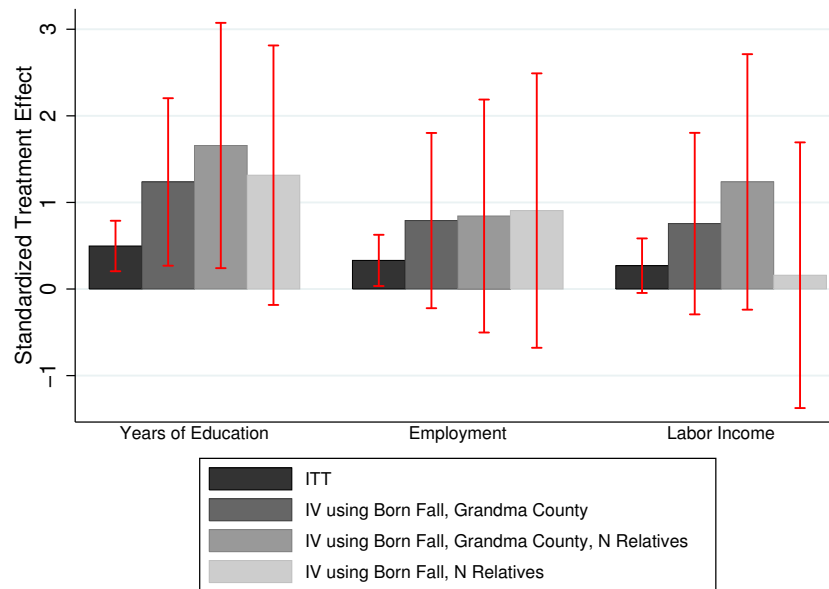
The charts below present the estimates for the parameter associated with  $D(\omega)$ . That is, the effect of participating in center-based childcare relative to a counterfactual of receiving no preschool alternative. As before, we present results for IQ scores at different ages and for a set of relevant adult outcomes. The results are not compelling, as they present irregularities over the life cycle that differ from the rest of results we present in the main paper and throughout this appendix.

Figure D.13: Effect of Center-based Childcare on IQ Scores, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool



Note: This plot presents the parameter associated with  $D(\omega)$  estimated using Control Functions as described in the text. The outcomes ( $Y(\omega)$ ) are IQ scores at different ages with a national standard deviation of 15 and a mean of 100.  $D(\omega) = 1$  for subjects that participate in ABC or CARE center-based childcare, and  $D(\omega) = 0$  for subjects who do not participate in treatment.  $Q(\omega)$  is the number of months attending preschool. It is coded as zero for subjects participating in ABC or CARE.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute.

Figure D.14: Effect of Center-based Childcare on Labor Market Outcomes, Accounting for Endogenous (log) Months of Take-up of Alternative Preschool



Note: This plot presents the parameter associated with  $D(\omega)$  estimated using Control Functions as described in the text. The outcomes ( $Y(\omega)$ ) are different adult outcomes labeled in the horizontal axis.  $D(\omega) = 1$  for subjects that participate in ABC or CARE center-based childcare, and  $D(\omega) = 0$  for subjects who do not participate in treatment.  $Q(\omega)$  is the number of months attending preschool. It is coded as zero for subjects participating in ABC or CARE.  $\mathbf{X}(\omega)$  includes a set of controls selected from all available baseline controls to maximize explanatory power across all outcomes tested in the paper: gender of the subject, mother's IQ score, High-risk Index, and Apgar Score at 1 minute.

## E Assessing Common Critiques to ABC

This section of the appendix assesses some of the critiques to ABC, which could apply to CARE as well. The critiques focus on the first phase of the program, to which we refer simply as ABC in this section of the appendix. We consider the following categories of critiques: (i) validity of treatment effects given early differences between treatment and control groups, (ii) fadeout of the treatment effect on cognitive outcomes, and (iii) compromised randomization.

## E.1 Treatment Effects vs. Early Differences Between Groups

The first critique we assess is that of [Spitz \(1992\)](#). The author asks if ABC prevented socio-cultural mental retardation. He focuses on different measures of cognition, and generically calls them IQ. His main focus is on the raw mean difference between the treatment and the control groups in the Bayley Mental Development Index (MDI) from the Bayley Scales of Infant Development and Infant Behavior (BSID) at 1 year of age.<sup>71</sup> He finds a noticeable disparity in the mean difference between cohorts 1 and 2 and cohorts 3 and 4, as early as six months after the treatment began. As [Spitz \(1992\)](#) argues, this difference, displayed in [Table E.1](#), is conspicuous. Importantly, the author fails to note that the mean differences have noticeably high standard errors.

Table E.1: Treatment - Control Mean Difference by Cohort, ABC

	(1) Bayley MDI 12 Months	(2) Bayley MDI 24 Months
Cohort 1	4.841 (6.013)	-2.901 (5.926)
Cohort 2	0.071 (6.013)	1.882 (5.830)
Cohort 3	9.143 (5.801)	10.038 (5.926)
Cohort 4	7.829 (5.801)	13.713 (5.830)
Observations	112	110
$R^2$	0.055	0.110

Note: This table displays the mean difference between the treatment and control groups in two measures of the Bayley Mental Development Index for ABC. Homoskedastic, asymptotic standard errors are in parentheses.

[Spitz \(1992\)](#) states that “an essential question is whether the differences at 6 months of age were due to the intervention or were preexisting” ([Spitz, 1992](#), p. 230). [Spitz \(1992\)](#) proceeds with speculative exercises that lead him to state that it is questionable to conclude that the

<sup>71</sup>The Bayley Mental Development Index (MDI) is a standard measure of cognition at taken between ages 0 and 2 ([Encyclopedia of Children’s Health, 2016](#)).

mean difference between the treatment and control groups as early as 6 months of age is a consequence of the treatment.<sup>72</sup>

His argument stems from not receiving information on the maternal characteristics, which he considers fundamental for his analysis to comparing test scores over the years.<sup>73</sup>

We present some annotations with respect to the comments in [Spitz \(1992\)](#).

**1. Mothers do not significantly differ in observed characteristics between the treatment and the control groups in any ABC cohort.** When jointly testing the hypothesis of differences in a battery of observed characteristics, we fail to reject any difference between the families of the subjects (see [Table A.11](#) to [Table A.14](#)). Even when the point estimates present some differences, they are minimal. For example, treatment-group mothers differ from control-group mothers, on average, by having one point higher IQ score (the population standard deviation is 15 points).

**2. Other early childhood education programs have generated gains in cognition very early in life, as soon as a year after the start of the program.** A more recent program, the Infant Health and Development Program (IHDP), treated children from ages 0 to 3. [Gross et al. \(1997\)](#) thoroughly describe this program. A lot of its elements were based on ABC and CARE. The program offered, on average, eight hours a week of center-based childcare from ages 1 to 3, weekly home visits from ages 0 to 1, and bi-monthly home visits from ages 1 to 3. Measures of the Bayley MDI at 12 and 24 months are available. The most intensive part of the treatment, center-based childcare, began in IHDP at 12 months of age

---

<sup>72</sup>The exact quote is: “Even if the differences between experimental and control groups at 6 months of age were a consequence of the first few months of intervention, which is questionable, the negligible additional effects after age 4.5 more years in the program should at least make one cautious about this kind of intervention’s potential [...]” ([Spitz, 1992](#), p. 235). We assess the former point before getting to the latter.

<sup>73</sup>He does not note that tests scores *do not have a well-defined scale* and are hard to compare.

while it began in ABC at birth. As in ABC, after a few months of center-based childcare, the effects on cognition were substantial in IHDP.

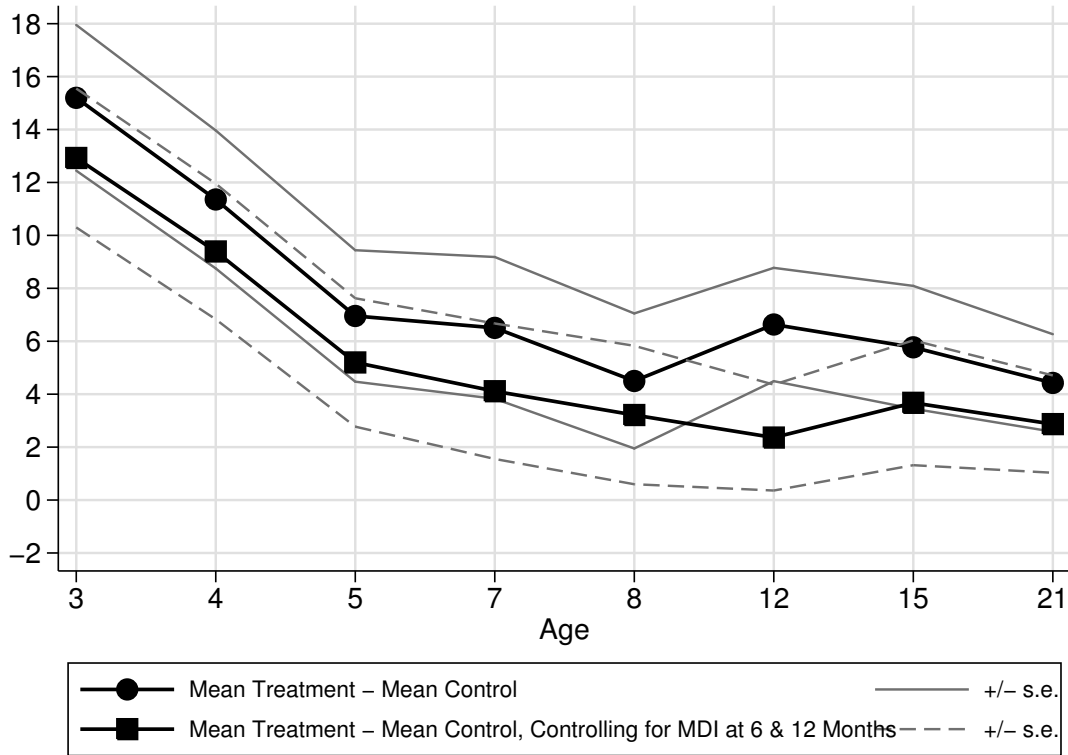
Table E.2: Treatment - Control Mean Difference, IHDP

	(1) Bayley MDI 12 Months	(2) Bayley MDI 24 Months
Treatment - Control Mean Difference	0.142 (0.968)	9.958 (1.498)
Observations	893	910
$R^2$	0.000	0.067

Note: This table displays the mean difference between the treatment and control groups in two measures of the Bayley Mental Development Index, for the Infant Health and Development Program using the principal analysis sample. Robust standard errors clustered by state and an indicator of weight below 2,000 grams in parentheses.

**3. The effects on cognition are sustained throughout childhood and up to age 21 even when partialing out the Bayley MDI at ages 12 and 24 months.** When partialing out from a linear regression the score in the Bayley MDI tests both at 12 and 24 months of age, the treatment effects on cognition, measured by the mean difference in IQ scores at different ages between the treatment and control groups, are still significant and are sustained up to age 21 (see Figure E.1).

Figure E.1: Treatment Effects on IQ, ABC



Note: This figure displays the mean difference in IQ scores at different ages between the treatment and control groups, pooling males and females. The line with circles represents the raw difference. The line with squares represents the difference when linearly controlling for the Bayley Mental Development Index at 12 and 24 months.

The author raises an additional concern. He states that even if the early-life effects on the Bayley MDI were valid, the effects were not long-lasting. When doing so, he exclusively refers to the effects the program had on cognition. As we document throughout the main paper, however, ABC had effects on a wide set of life-cycle outcomes, which perhaps are more important than the effects on cognition if measured by the economic gains they generate. Examples include: employment at age 30 for males, high-school graduation for females, and a variety of health measures at age 34 for males. We document these effects in the main paper.

## E.2 Fadeout of Treatment Effects

This sustained effect on life-cycle outcomes relates to another category of criticism of ABC—the fadeout of treatment effects after the actual treatment period.<sup>74</sup> Based on the fadeout of the effects on cognition and age-21 outcomes, some studies claim that ABC had no sustained long-term effects. In addition to the evidence we cite above, it is important to add the following.

**4. The treatment effects of ABC differed by gender.** For example, we find substantial effects on employment for males and high-school graduation for females. This does not mean that the program did not work. It means that market conditions could have affected the way in which the treated subjects benefited from the program’s effects.

## E.3 Compromised Randomization

The third and last critique we assess is that of compromised randomization.<sup>75</sup> The main critique to the program in this respect is that seven children in the treatment group did not comply to their initial assignment and one subject switched from control to treatment status (see Table A.1, which we reproduce from the main paper). Although we document more cases of compromised randomization, we start by assessing these eight cases.<sup>76</sup>

First, we propose a method to assess the non-compliance of six subjects for whom we have data available from ages 0 to 8. This allows for the possibility of adjusting the mean difference between the treatment and the control groups, using a standard Bloom estimator

---

<sup>74</sup>See, for example, Besharov et al. (2011).

<sup>75</sup>See for example Baumeister and Bacharach (2000) and the multiple studies cited there.

<sup>76</sup>Our main methodology assesses all the cases except the four subjects for whom we do not have data at all. We are not able to account for them because our method is based on observing at least some baseline characteristics. See Section 3 in the main paper for more details.

(Bloom, 1984). This estimator accounts for compliance, by weighting the mean difference by a measure of the probability of compliance. It takes the following form:

$$\Delta^{\text{Bloom}} = \frac{\mathbb{E}[Y|R = 1] - \mathbb{E}[Y|R = 0]}{\mathbb{E}[D|R = 1] - \mathbb{E}[D|R = 0]}, \quad (11)$$

where  $R$  indicates randomization into treatment,  $D$  indicates treatment take-up, and  $Y$  is an outcome of interest.

This estimator coincides with the instrumental variable estimator of the parameter in a linear regression of an outcome on a single explanatory variable, e.g. treatment take-up. As Angrist et al. (1996) show, under certain assumptions it can be interpreted as the average treatment effect for those individuals induced to take up treatment by the instrument. This interpretation has various weaknesses (Heckman et al., 2006; Heckman and Urzúa, 2010).

To compare the mean difference and the Bloom estimator we consider two measures of cognition, which are commonly used when evaluating ABC and are available for these six subjects: the Stanford-Binet Intelligence Scale at age 3 and the Wechsler Preschool and Primary Scale of Intelligence at age 5. Table E.3 presents the results and shows that the differences are minimal. **5. Adjusting for the non-compliance of these six subjects makes little difference** when assessing outcomes that we observe for these subjects.

Table E.3: Assessing Non-compliance in ABC, Exercise 1

	(1) IQ, Age 3	(2) IQ, Age 3	(3) IQ, Age 5	(4) IQ, Age 5
Treatment - Control Mean Difference	14.970 (2.794)		6.398 (2.494)	
Bloom		16.245 (2.877)		6.943 (2.643)
Observations	100	100	100	100
$R^2$	0.227	0.289	0.063	0.088

Note: This table displays the mean difference between the treatment and control groups and the same difference adjusted for non-compliance (Bloom estimator) in two measures of IQ, the Stanford-Binet Intelligence Scale at age 3 and the Wechsler Preschool and Primary Scale of Intelligence at age 5. Homoskedastic, asymptotic standard errors are in parentheses.

Second, we propose a method to account for non-compliance of the four subjects for whom we have no data at all. We reproduce the estimates in Table E.4, assigning the lowest score across all the subjects in the treatment group to each of these four children. For each of the tests we consider, the lowest score was 71 for both the age-3 and the age-5 measures.

**6. Even when assigning the lowest score to the subjects who did not comply to treatment and for whom no data are available, the mean difference between treatment and control is sizable and not significantly different from the baseline cases where we do not adjust for non-compliance cases (columns (1) and (3) in Table E.3).**

Table E.4: Assessing Non-compliance in ABC, Exercise 2

	(1) IQ, Age 3	(2) IQ, Age 3	(3) IQ, Age 5	(4) IQ, Age 5
Treatment - Control Mean Difference	12.908 (2.901)		4.284 (2.647)	
Bloom		15.111 (3.061)		5.016 (2.960)
Observations	104	104	104	104
$R^2$	0.163	0.306	0.025	0.093

Note: This table displays the mean difference between the treatment and control groups and the same difference adjusted for non-compliance (Bloom estimator) in two measures of IQ, the Stanford-Binet Intelligence Scale at age 3 and the Wechsler Preschool and Primary Scale of Intelligence at age 5. We impute the minimum test scores of four subjects initially assigned to treatment for whom we do not have data. The minimum of both the age-3 and age-5 tests was 71. Homoskedastic, asymptotic standard errors are in parentheses.

Finally, we note that our methodology in Section 3 and the results we present in Section 4 account for the remainder of randomization compromises, as when the families moved or abandoned the study later on. **7. The results are not greatly sensitive to adjusting for the rest of the randomization compromises**, as we show in Section 4.

## F Calculating the Benefit-to-cost Ratio and the IRR

We estimate the benefit-to-cost ratio and the internal rate of return of participation in ABC and CARE for the *average participant* of the program. Let  $B_{r,t}$  denote benefits and  $C_{r,t}$  denote costs for group  $r$  at age  $t$ . Then  $r$  takes either two values, 0 for children randomized into the control group or 1 for children randomized into the treatment group.

Both  $B_{r,t}$  and  $C_{r,t}$  are in 2014 USD and discounted to  $t = 0$ . Thus, the benefits  $B_t = B_{1,t} - B_{0,t}$  and costs  $C_t = C_{1,t} - C_{0,t}$  associated with program participation are the differences between the counter-factual benefits and costs. We use the counter-factual notation for costs because the costs for the control group are not zero: the control group children received some services

from the ABC and CARE centers and some attended alternative care programs.

The estimate for the benefit-to-cost ratio is

$$\mathbf{E} \left( \frac{\sum_{t=0}^T B_t}{\sum_{t=0}^T C_t} \right). \quad (12)$$

The estimate for the internal rate of return is the value  $\rho$  that solves:

$$\sum_{t=0}^T \frac{\mathbf{E}(B_t - C_t)}{(1 + \rho)^t} = 0. \quad (13)$$

## G Accounting for Attrition

We use a standard inverse probability weighting (IPW) scheme based on a logit model that predicts attrition using baseline characteristics.<sup>77</sup> We detail the identification of this scheme below. We describe our model selection process and the selected logit models in Tables H.1–H.4 in Appendix H.1. We use the Akaike Information Criteria (AIC) to select the model with the best predictive ability. We use the same notation as in Section 3.3 and a vector of background characteristics used in the randomization protocol;  $W$  are background characteristics not used in this protocol.

Formally, let  $A = 1$  denote the case where we observe  $Y$  and  $A = 0$  denote otherwise. We assume  $A$  is independent of  $Y$  conditional on  $Z$ ,  $R$ ,  $X$ , and  $W$ . Thus,

### Assumption G.1

$$A \perp\!\!\!\perp Y | W, Z, X, R.$$

---

<sup>77</sup>Horvitz and Thompson (1952).

Based on this assumption, we identify  $\mathbb{E}[Y_r]$  as follows:

$$\begin{aligned}
\mathbb{E}[Y_r] &= \int y f_{Y_r|Z}(y) f_Z(z) dy dz & (14) \\
&= \int y f_{Y|Z, R=r}(y) f_Z(z) dy dz \\
&= \int y f_{Y|R=r, W, Z, X}(y) f_{W, X|R=r}(w, x) f_Z(z) dy dx dw dz \\
&= \int y f_{Y|R=r, W, Z, X, A=1}(y) f_{W, X|R=r, Z}(w, x) f_Z(z) dy dx dw dz
\end{aligned}$$

where each component of (14) is straightforward to recover from the data. Using Bayes' Theorem, we can write an equivalent expression to make the IPW scheme explicit. That is, we apply Bayes' Theorem to  $f_{W, X|R=r, Z}(w, x)$  and  $f_Z(z)$  to obtain

$$f_{W, X|R=r, Z}(w, x) = \frac{f_{W, X|R=r, Z, A=1}(w, x) P(A = 1|R = r, Z)}{P(A = 1|R = r, W, Z, X)}$$

and

$$f_Z(z) = \frac{f_{Z|R=r, A=1}(z) P(R = r, A = 1)}{P(R = r, A = 1|Z)}.$$

Then, we substitute these expressions into (14) to get

$$\begin{aligned}
\mathbb{E}[Y_r] &= \int y f_{Y, X, W, Z|R=r, A=1}(y, x, w, z) \frac{P(R = r, A = 1) P(A = 1|R = r, Z)}{P(R = r, A = 1|Z) P(A = 1|R = r, W, Z, X)} dy dx dw dz \\
&= \int y f_{Y, X, W, Z|R=r, A=1}(y, x, w, z) \frac{P(R = r, A = 1)}{P(R = r|Z) P(A = 1|R = r, W, Z, X)} dy dx dw dz.
\end{aligned}$$

Assumption G.1 generalizes the matching assumption of [Campbell et al. \(2014\)](#). It conditions not only on pre-program variables but on fully observed post-treatment variables,  $X$ . This enables us to account for two types of selection: (i) selection into treatment; and (ii) selection

into item response. The corresponding sample estimator for  $\mathbb{E}[Y_r]$  is

$$\sum_{i \in \mathcal{I}} y(\omega) a_i w_{i,r} \mathbf{1}(r_i = r)$$

where  $\mathcal{I}$  indexes the individuals in the sample,  $a_i$  indicates whether we observe  $Y$  for individual  $i$ , and

$$w_{i,r} = \frac{1}{\pi_r(z_i) \alpha(r_i, w_i, z_i, x_i)} \frac{1}{\sum_k \frac{\mathbf{1}(r_i=r) \mathbf{1}(a_i=1)}{\pi_r(z_k) \alpha(r_k, w_k, z_k, x_k)}},$$

with  $\pi_r(z) := P(R = r | Z = z)$  and  $\alpha(r, w, z, x) := P(A = 1 | R = r, W = w, Z = z, X = x)$ . The weight  $\pi_r$  corrects for selection into treatment based on pre-program variables  $Z$ . The weight  $\alpha$  corrects for item non-response based on  $R, W, Z, X$ .

## H Details on Interpolation and Extrapolation

### H.1 Identifying Treatment Effects

In the main paper, we discuss three cases of attrition. Below, we elaborate on the methodology we use for each case.

#### H.1.1 Fully Observed Outcomes

In the case where we fully observe outcomes, we only need to address the compromised randomization of treatment status in our empirical estimates (see Section 3).

Randomization into treatment and control for each cohort occurred after the cohort was split into two subsamples balanced by gender, maternal IQ, and number of siblings. Subjects from each subsample were then matched on HRI, and randomized into their respective experimental groups. Sex, maternal IQ, number of siblings, and HRI therefore comprise  $Z$ . Thus, to obtain an unbiased estimate of the treatment effect, we regress the outcome variable on treatment status, controlling for those four baseline variables, as well as for cohort.

We perform this estimation on 1,000 bootstrap resamples of the original ABC and CARE data. We take the mean of the empirical bootstrap distribution as our point estimate, and the standard deviation of the distribution as the standard error. We estimate these effects on the pooled ABC and CARE samples, as well as the male and female subsamples. This allows us to generate rates of return and benefit-cost ratios for the male subsample, the female subsample, and the pooled sample.

### **H.1.2 Partially Observed Outcomes**

Given the small sample sizes of the ABC and CARE data, we consider outcome variables with fewer than 100 observations to exhibit a sizable rate of attrition. These variables include parental income at ages 2, 3, 4, 5, 9, 12, and 15, for which we observe no more than 91 subjects at any given age; and the health survey at age 34, for which we observe no more than 72 subjects.

We weight each subject by his/her propensity to select into treatment and propensity to respond. We estimate the former with a logit model for treatment status, controlling for baseline variables used to determine randomization into treatment status: cohort, maternal IQ, number of siblings, HRI, and sex (in the case of pooled estimates).

To estimate the propensity to respond, we similarly estimate a logit model of observing the outcome, controlling for treatment status and three additional covariates. The additional covariates are chosen amongst a set of variables likely to be correlated with non-response rates for the outcome variable. This set of variables includes baseline characteristics determining randomization, baseline variables other than randomization, and observed adult outcomes potentially affected by treatment. To select these variables, we estimate the aforementioned logit model controlling for treatment status and every three-variable combination from the described set of variables. The model with the lowest Akaike Information Criteria (AIC) value is used to estimate the propensity to respond. We limit the number of covariates in our logit models to three as this noticeably reduces the frequency of bootstrap resamples which exhibit perfect separation in the covariates.

To estimate the treatment effect, we regress the outcome of interest on treatment status, and weight the observations by the inverse of the product of the two propensity scores estimated with the method documented in Section 3 in the main paper. This procedure is performed on 1,000 bootstrap resamples of the original ABC and CARE data. We take the mean of the empirical bootstrap distribution as our point estimate, and the standard deviation of the distribution as the standard error. Again, we carry out this procedure for the pooled sample, male subsample, and female subsample.

#### **H.1.2.1 Parental Income**

Parental income is reported at ages 0, 2, 3, 4, 5, 9, 12, and 15, allowing us to directly estimate the treatment effect on parental income at those ages. Given the rate of attrition in parental income from age 2 onwards, we estimate these treatment effects accounting for attrition with IPW weights. To obtain a more complete picture of how the program affected parental income, we estimate the treatment effect in the missing ages by linearly interpolating the

treatment effects between the years for which we observe parental income. We assume there are no impacts from age 16 onwards (i.e. a treatment effect of 0).

### **H.1.3 Fully Missing Outcomes**

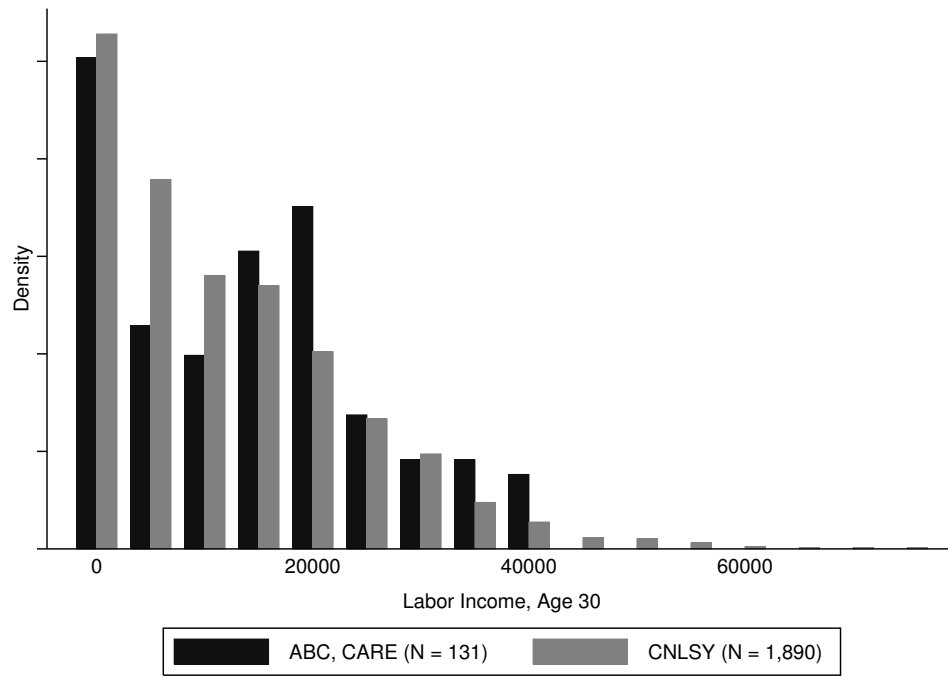
#### **H.1.3.1 Subject Income**

Labor and public-transfer income of subjects are observed intermittently at ages 21 and 30 only. In order to obtain a complete measure of the impact of the ABC and CARE programs on subject income, we interpolate the income streams of subjects for ages 22 to 29, and extrapolate income streams for ages 31 to 67 using three auxiliary datasets: NLSY79, CNLSY, and PSID.

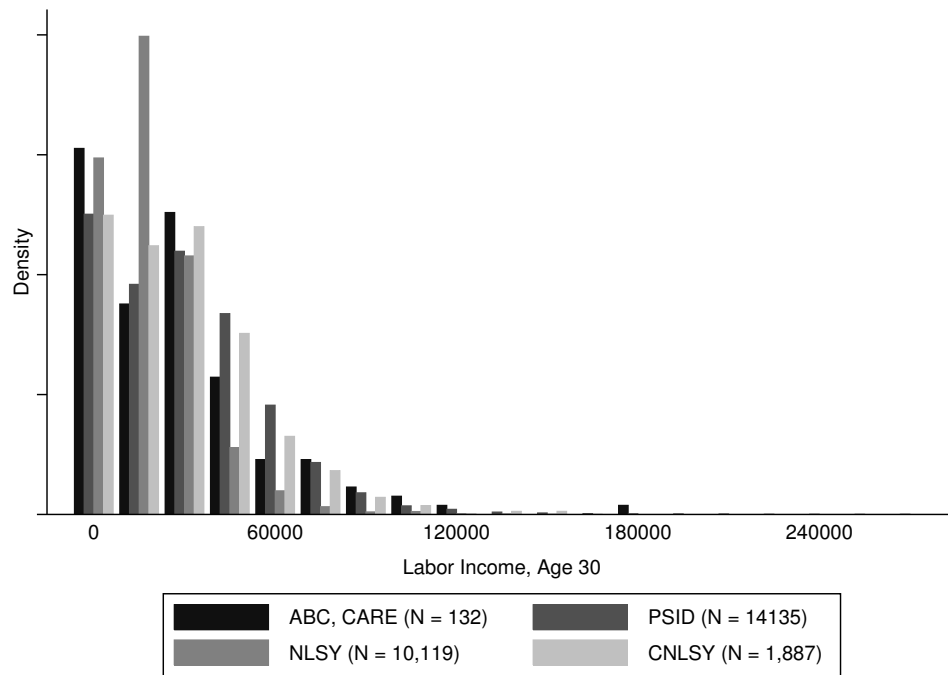
We assume common support between the auxiliary datasets and the analysis sample. This is required in order for us to reliably use external data to provide inference for the ABC and CARE samples. Figure [H.1](#) validates this assumption by displaying the overlapping support sets of ABC and our auxiliary data for the variables used to interpolate and extrapolate earnings. For details about these datasets, see Appendix [M](#).

Figure H.1: Support of ABC and Auxiliary Data

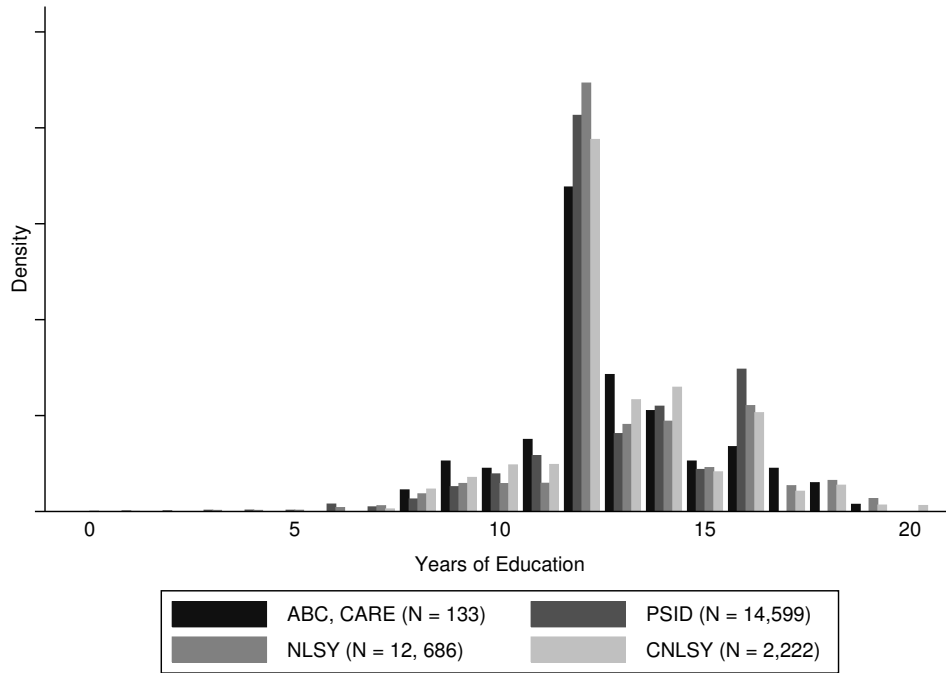
(a) Income at Age 21



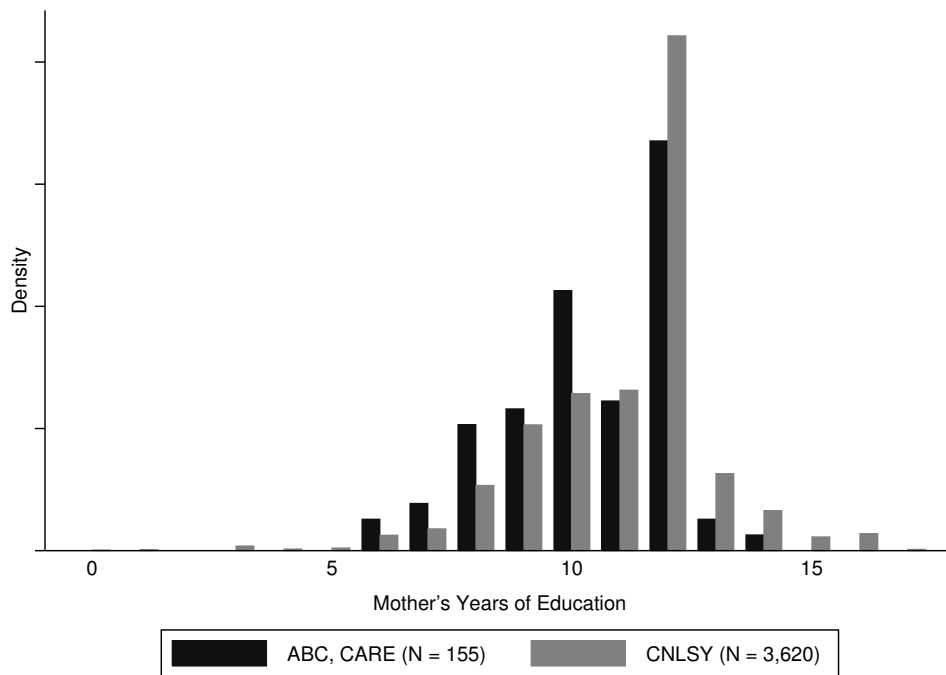
(b) Income at Age 30



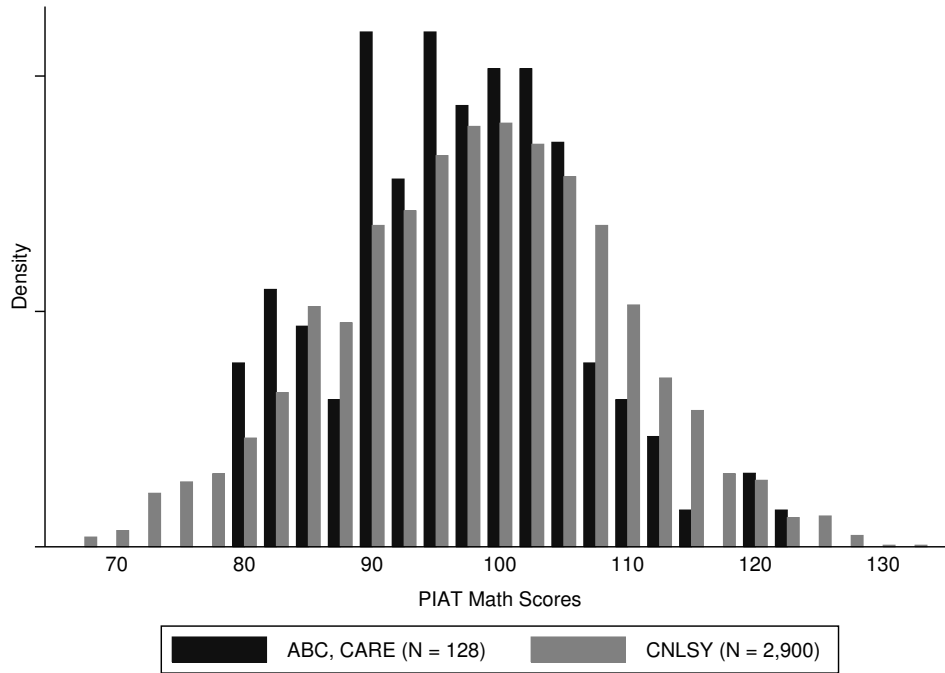
(c) Subject's Years of Education



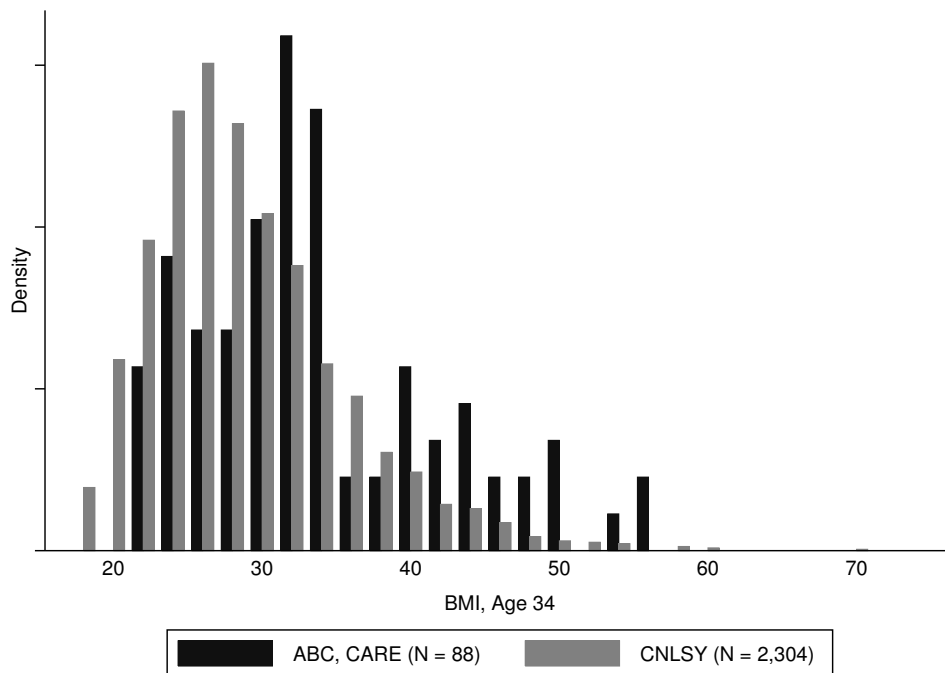
(d) Mother's Years of Education



(a) Average PIAT Math Scores, Ages 5–7



(b) Body Mass Index, Age 34



Note: These graphs display the support of ABC, PSID, NLSY, and CNLSY for variables we use to project future earnings. PIAT math scores are averaged over ages 5–7.

Both labor and public-transfer income are interpolated using the exact same methodology, so we restrict our attention to interpolating labor income. We employ the CNLSY dataset to model labor income at each age from 22 to 29. We use ordinary least squares and regress labor income of CNLSY subjects on background characteristics and early adult outcomes. Variables for background characteristics consist of an indicator for being black, mother’s education, and sex. Variables used to interpolate early adult outcomes consist of labor income at ages 21 and 30, years of education at age 30, average PIAT math scores from ages 5 to 10, and BMI at age 34. We estimate the model on 1,000 bootstrap resamples of the CNLSY to obtain a measure of prediction error, and denote the vectors of coefficients *for the outcome variables* by  $\beta_{a,s}$ , where  $a \in \{22, 23, \dots, 29\}$  denotes age, and  $s \in \{1, 2, \dots, 1,000\}$  indexes the bootstrap samples.

The outcome variables we use as independent variables to fit the model to the CNLSY data are also observed in the ABC and CARE data, and we are able to calculate the treatment effects on each of these outcomes using the procedures described in Appendix H.1.1 and Appendix H.1.2. We obtain these estimates for 1,000 bootstrap resamples of both the ABC and CARE data, and denote the vectors of coefficients by  $\gamma_r$ , where  $r \in \{1, 2, \dots, 1,000\}$  indexes the bootstrap samples.

To interpolate the treatment effect on subject labor income at age  $a$ , we take the dot product of  $\beta_{a,s}$  and  $\gamma_r$  for all  $s, r$  pairs, resulting in 1,000,000 scalars. We reintroduce prediction error by first randomly drawing residuals from our fitted model on the auxiliary data and assigning them to the ABC and CARE subjects. We then take the mean difference of these residuals across treatment and control groups, and add it to the previously described scalar. We do this for all 1,000,000 scalars. These scalars form the empirical bootstrap distribution of the treatment effect. Note that we bypass the estimation of the subjects’ actual incomes, of which we are unable to obtain an unbiased estimate. Instead, we directly obtain an unbiased

estimate of the treatment effect. This method implies that the impacts of the treatment on adult labor income can be expressed through the observed early-adult outcomes. In this case, the observed early-adult outcomes are labor income at ages 21 and 30, years of education at age 30, average PIAT math scores from ages 5 to 10, and BMI at age 34. Since  $\gamma_r$  has been adjusted to address selection into treatment and attrition in the ABC and CARE data, the bootstrap distribution is devoid of selection and attrition issues. In addition, the bootstrap distribution captures the error stemming from the ABC or CARE datasets, and prediction error of the model fitted to the CNLSY dataset. We take the mean of the empirical bootstrap distribution as our point estimate, and the standard deviation of the distribution as the standard error. Again, we carry out this procedure for the pooled, male, and female samples.

To extrapolate labor income, we employ the NLSY79 and PSID datasets. We execute the same estimation strategy as above, but with a reduced set of control variables in order to limit the number of dropped cases. Specifically, the set of background variables we control for include an indicator for being black and gender of the subject. As the NLSY79 and PSID are not as rich as the CNLSY, we also restrict our outcome variables to years of education at age 30 and subject labor income at age 30.

We estimate the vectors  $\beta_{a,s}$  for  $a \in \{31, 32, \dots, 67\}$  and  $\gamma_r$ , where  $s, r \in \{1, 2, \dots, 1,000\}$  index the bootstrap samples, and take the dot product of every pair of  $\beta_{a,s}$  and  $\gamma_r$  vectors to obtain the empirical bootstrap distribution of the treatment effect on labor income past age 30. We take the mean of the distribution to be the point estimate, and the standard deviation of the distribution to be the standard error.

The same procedure is carried out to interpolate public-transfer income, where we instead utilize public-transfer income as the dependent variable in the regressions on our auxiliary datasets. In the case of extrapolating public-transfer income from welfare, unemployment

benefits, AFDC, and food stamps, we also apply the same methodology as for labor income. However, in the case of disability insurance, social security, and supplemental security income, we apply an alternative approach. For these cases, we estimate the probability each ABC or CARE individual claims each of those benefits at ages 31–67, and multiply those probabilities by the average annual claims of each type of public transfer reported by the Social Security Administration in 2014. See Appendix [N.3.1](#) for an explanation of how we estimate these probabilities. For each age, we sum across our estimates of each type of public transfer received to obtain the total public-transfer income of both ABC and CARE subjects.

### **H.1.3.2 Health Costs**

We predict health care costs from age 30 up to age 79, and criminal costs from birth up to age 50. We then evaluate the predicted outcomes as if observed, and estimate the treatment effect on those data.

As health outcomes are simulated using a set of variables exhibiting a high rate of attrition, the predicted health outcomes exhibit an equal rate of attrition. We therefore employ the method described in Appendix [H.1.2](#) to estimate the treatment effect on health care costs at every age from 30 through 79, conditional on survival up to each age.

### **H.1.3.3 Crime Costs**

In the case of crime costs, we are able to predict costs for more than 100 individuals in the sample. We therefore utilize the method described in subsection [H.1.1](#) to estimate the treatment effect on the cost of criminal activity at every age from birth until age 50.

Table H.1: Variables Estimated without IPW

Variable	Age
IQ Standard Score	2, 3, 4, 5, 6, 7, 8, 12, 15, 21
PIAT Math Standard Score	7
Home Total Score	0.5, 1.5, 2.5, 3.5, 4.5, 8
Mother Works	2, 3, 4, 5
Bio. Mother's Education Level	2, 3, 4, 5, 9
Father is Home	2, 3, 4, 5
Ever Adopted	NA
Graduated High School	NA
Attended Vocation/Tech/Community College	NA
Earned Degree from 4-year College	NA
Years of Education	30
Works a Job	30
Total Felony Arrests	Mid-30s
Total Misdemeanor Arrests	Mid-30s
Total Years Incarcerated	30
Self-reported Health	30
Number of Cigarettes Smoked Per Day Last Month	30
Number of Days Drank Alcohol Last Month	30
Number of Days Binge Drank Alcohol Last Month	30
Program Costs	0-26
Control Contamination Costs	0-26
Education Costs	0-26
Medical Expenditure	8-30
Justice System Costs	0-50
Prison Costs	0-50
Victimization Costs	0-50

Note: The table above lists the variables for which we do not apply IPW when estimating treatment effects.

Table H.2: Model Selection for Attrition IPW Pooled

Variable	Age	Obs.	(1)	(2)	(3)
Parental Labor Income	2	72	Subject Birth, Year	Mother Works, age 5	Father works before pregnancy
Parental Labor Income	3	73	Subject Birth, Year	Mother Works, age 4	Father works before pregnancy
Parental Labor Income	4	69	Subject Birth, Year	Mother Works, age 4	Siblings in Household, age 3
Parental Labor Income	5	86	Mother Works, age 4	Father works before pregnancy	Siblings in Household, age 5
Parental Labor Income	9	82	Mother Works, age 3	Father home at birth	Siblings in Household, age 5
Parental Labor Income	12	91	Subject Birth, Year	Mother Works, age 4	Mother Works, age 5
Parental Labor Income	15	90	Mother Works, age 3	HRI 2: No Maternal Relatives	Siblings in Household, age 5
Subject Labor Income	21	96	HRI 10: Other special circumstances	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Subject Labor Income	30	100	Mother's Age at entry	Number of Siblings at entry	Number of times smoked marijuana
Subject Public Transfer Income	21	97	Mother's WAIS Block Design	ASR t-score: Attention Deficit and Hyperactivity	ASR t-score: Rule Breaking
Subject Public Transfer Income	30	101	Number of times smoked marijuana	ASR t-score: Substance Abuse	Works a Job
Self-reported Health	Mid-30s	70	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Self-reported Drug User	Mid-30s	67	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Systolic Blood Pressure (mm Hg)	Mid-30s	68	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Diastolic Blood Pressure (mm Hg)	Mid-30s	68	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Prehypertension, Sys. B.P. > 120 or Dys. B.P. > 80	Mid-30s	68	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Hypertension, Sys. B.P. > 140 or Dys. B.P. > 90	Mid-30s	68	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
High-density lipoprotein (HDL) Cholesterol (mg/dL)	Mid-30s	71	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Hemoglobin Level (%)	Mid-30s	71	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Prediabetes, Hemoglobin > 5.7%	Mid-30s	70	Gestational Age < 36	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Diabetes, Hemoglobin > 6.5%	Mid-30s	70	Gestational Age < 36	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Vitamin D Deficiency (< 20 ng/mL)	Mid-30s	71	Gestational Age < 36	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Measured BMI	Mid-30s	66	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Obesity (BMI > 30)	Mid-30s	69	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Severe Obesity (BMI > 35)	Mid-30s	62	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education
Waist-hip Ratio	Mid-30s	62	Gestational Age < 36	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Abdominal Obesity	Mid-30s	62	Gestational Age < 36	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Framingham Risk Score	Mid-30s	66	Gestational Age < 36	ASR t-score: Substance Abuse	Years of Education

Note: The table above lists the variables used to generate the component of the inverse probability weights we use to account for attrition. To obtain this component for attrition, we estimate a logit model for observing the outcome variable, controlling for treatment, and three additional covariates depending on the outcome variable. The additional covariates are chosen from a list of covariates possibly affecting response rates. The set yielding the lowest Akaike Information Criteria value is used to generate the IPW weights, and are listed in the table above.

Table H.3: Model Selection for Attrition IPW  
Females

Variable	Age	Obs.	(1)	(2)	(3)
Parental Labor Income	2	33	Subject Birth Year	Mother Works, age 5	Father works before pregnancy
Parental Labor Income	3	33	Subject Birth Year	Father Works, age 5	Births in Household, age 5
Parental Labor Income	4	31	Subject Birth Year	Mother Works, before pregnancy	Father works before pregnancy
Parental Labor Income	5	31	Mother Works, before pregnancy	Mother Works, age 5	Father works before pregnancy
Parental Labor Income	9	40	Mother Works, age 5	HRI 2: No Maternal Relatives	Father works before pregnancy
Parental Labor Income	12	45	Subject Birth Year	Mother Works, age 5	Father works before pregnancy
Parental Labor Income	15	47	Mother Works, age 5	HRI 2: No Maternal Relatives	Father works before pregnancy
Subject Labor Income	21	51	Mother's WAIS Performance IQ	Mother's Age at entry	ASR t-score: Attention Deficit and Hyperactivity
Subject Labor Income	20	53	HRI 10: Other special circumstances	ASR t-score: Substance Abuse	Years of Education
Subject Public Transfer Income	21	53	Mother's WAIS Comprehension	Number of cigarettes smoked per day	ASR t-score: Attention Deficit and Hyperactivity
Subject Public Transfer Income	30	53	HRI 10: Other special circumstances	ASR t-score: Substance Abuse	Years of Education
Self-reported Health	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Self-reported Drug User	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Systolic Blood Pressure (mm Hg)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Diastolic Blood Pressure (mm Hg)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Prehypertension, Sys. B.P. > 120 or Dys. B.P. > 80	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Hypertension, Sys. B.P. > 140 or Dys. B.P. > 90	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
High-Density Lipoprotein (HDL) Cholesterol (mg/dL)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Dyslipidemia (HDL < 40 mg/dL)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Hemoglobin Level (%)	Mid-30s	39	Length at birth	ASR t-score: Substance Abuse	Years of Education
Pre-diabetes, Hemoglobin > 5.7%	Mid-30s	39	Length at birth	ASR t-score: Substance Abuse	Years of Education
Diabetes, Hemoglobin > 6.5%	Mid-30s	39	Length at birth	ASR t-score: Substance Abuse	Years of Education
Vitamin D deficiency (< 20 ng/mL)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Obesity (BMI > 30)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Severe Obesity (BMI > 35)	Mid-30s	40	Length at birth	ASR t-score: Substance Abuse	Years of Education
Waist-hip Ratio	Mid-30s	37	Length at birth	ASR t-score: Substance Abuse	Years of Education
Abdominal Obesity	Mid-30s	37	Length at birth	ASR t-score: Substance Abuse	Years of Education
Framingham Risk Score	Mid-30s	39	Length at birth	ASR t-score: Substance Abuse	Years of Education

Note: The table above lists the variables used to generate the component of the inverse probability weights we use to account for attrition. To obtain this component for attrition, we estimate a logit model for observing the outcome variable, controlling for treatment, and three additional covariates depending on the outcome variable. The additional variables are chosen from a list of covariates possibly affecting response rates. The set yielding the lowest Akaike Information Criteria value is used to generate the IPW weights, and are listed in the table above.

Table H.4: Model Selection for Attrition IPW  
Males

Variable	Age	Obs.	(1)	(2)	(3)
Parental Labor Income	2	39	Subject Birth Year	Mother Works, age 3	Siblings in Household, age 3
Parental Labor Income	3	39	Subject Birth Year	Mother Works, age 3	Father works before pregnancy
Parental Labor Income	4	38	Subject Birth Year	Mother Works, age 4	Siblings in Household, age 3
Parental Labor Income	5	42	Mother Works, age 3	Mother Works, age 4	Siblings in Household, age 5
Parental Labor Income	9	42	Mother Works, age 3	Mother Works, age 5	Siblings in Household, age 5
Parental Labor Income	12	46	Subject Birth Year	Father home, age 4	Father home, age 5
Parental Labor Income	15	43	Mother Works, age 4	Father Works before pregnancy	Siblings in Household, age 5
Subject Labor Income	21	45	HRI 10: Other special circumstances	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse
Subject Labor Income	30	47	Father's Age at entry	Mother Married at entry	ASR t-score: Substance Abuse
Subject Public Transfer Income	21	48	Father's Age at entry	Gestational Age $\leq$ 36	Years of Education
Subject Public Transfer Income	30	48	Mother's WAIS Performance IQ	ASR t-score: Antisocial Personality Problems	Had physical exam for illness in last 2 years
Self-reported Health	Mid-30s	30	HRI 2: No Maternal Relatives	ASR t-score: Substance Abuse	ASR t-score: Antisocial Personality Problems
Self-reported Drug User	Mid-30s	27	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Systolic Blood Pressure (mm Hg)	Mid-30s	28	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Diastolic Blood Pressure (mm Hg)	Mid-30s	28	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Hypertension, Sys. B.P. $>$ 120 or Dys. B.P. $>$ 80	Mid-30s	28	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Hypertension, Sys. B.P. $>$ 140 or Dys. B.P. $>$ 90	Mid-30s	28	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
High-Density Lipoprotein (HDL) Cholesterol (mg/dL)	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Low-Density Lipoprotein (LDL) Cholesterol (mg/dL)	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Triglyceride Level (%)	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Prediabetes	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Diabetes, Hemoglobin $>$ 5.7%	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Diabetes, Hemoglobin $>$ 6.5%	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Vitamin D Deficiency ( $<$ 20 ng/mL)	Mid-30s	31	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Measured BMI	Mid-30s	26	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Obesity (BMI $>$ 30)	Mid-30s	28	Father works before pregnancy	ASR t-score: Substance Abuse	ASR t-score: Attention Deficit and Hyperactivity
Severe Obesity (BMI $>$ 35)	Mid-30s	29	HRI 2: No Maternal Relatives	ASR t-score: Substance Abuse	Number of times smoked marijuana
Waist-hip Ratio	Mid-30s	25	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Abdominal Obesity	Mid-30s	25	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years
Framingham Risk Score	Mid-30s	27	Gestational Age $\leq$ 36	ASR t-score: Substance Abuse	Had physical exam for illness in last 2 years

Note: The table above lists the variables used to generate the component of the inverse probability weights we use to account for attrition. To obtain this component for attrition, we estimate a logit model for observing the outcome variable, controlling for treatment, and three additional covariates depending on the outcome variable. The additional variables are chosen from a list of covariates possibly affecting response rates. The set yielding the lowest Akaike Information Criteria value is used to generate the IPW weights, and are listed in the table above.

## H.2 Internal Rate of Return

To estimate the internal rate of return, we must solve for  $\rho$  in the equation below,

$$\sum_{t=0}^T \frac{\mathbb{E}(B_t - C_t)}{(1 + \rho)^t} = 0, \quad (15)$$

where we let  $T = 79$ , define  $B_t$  and  $C_t$  to be the total benefits and costs of the program at time  $t$ , and define  $\mathbb{E}$  to be the sample mean.

All outcomes of the parents and subjects that we observe as having been affected by the program are treated as benefits. For this to make sense, we reverse the sign of the monetized effect of the program on specific outcomes. Costs of ABC and CARE consist only of the initial program costs from ages 0 to 5. Table H.5 provides a full list of the benefits and costs of ABC.

We take the sum of the treatment effects on each component of the benefits to be the total benefit,  $B_t$ , of the ABC and CARE programs (see Appendix H.1 on how we estimate these treatment effects). This includes parental income, subject labor income, and QALYs (quality-adjusted life years). Treatment effects on costs borne by the subject or society have their signs reversed and are included as benefits. We do this for subject public-transfer income, education costs, crime costs, control contamination costs, and health costs. To account for deadweight loss, we impose a marginal welfare cost of 50% by multiplying public costs by a factor of 1.5.<sup>78</sup> Public costs include education costs up until age 17, jail costs, justice system costs, Medicare costs, Medicaid costs, disability insurance claims, social security claims, and supplemental security income.

---

<sup>78</sup>There is no clear consensus on the marginal welfare cost of tax revenue. However, most researchers estimate the welfare cost per tax dollar to be between \$0.30 and \$0.50. See Feldstein (1999), Heckman and Smith (1998), and Browning (1987).

Having constructed our cash flow,  $\mathbb{E}(B_t - C_t)$ , solving for  $\rho$  reduces to an algebraic exercise.

As described in Appendix H.1, we estimate the treatment effect on each component of the benefits and costs at time  $t$  for the pooled, male, and female samples. We do this for 1,000 bootstrap resamples of the original ABC and CARE data. In the case of health costs and subject income, for which we employ auxiliary datasets to estimate the treatment effects, we also obtain 1,000 bootstrap estimates from the auxiliary data for every ABC or CARE bootstrap resample, resulting in a total of 1,000,000 estimates. By reusing each bootstrap estimate of the treatment effect on outcomes that do not require any auxiliary data set 1,000 times, we obtain a total of 1,000,000 estimates of the cash flow. We estimate the IRR on each of those cash flows. This is how we form our empirical bootstrap distribution of the IRR for the pooled, male, and female samples. We take the mean of the distributions to be the point estimates, and we take the standard deviations to be the standard errors. To construct the 80% confidence intervals, we take the 10<sup>th</sup> and 90<sup>th</sup> quantiles of each bootstrap distribution.

Table H.5: Components of Benefits and Costs

Variable	Sign Reversed	Welfare Cost Factor
<b>Benefits</b>		
Parent Income		
Subject QALY		
Subject Labor Income		
Subject Public-transfer Income	✓	0.5
Medicare Costs	✓	1.5
Medicaid Costs	✓	1.5
Out-of-pocket Medical Costs	✓	
Miscellaneous Medical Costs	✓	
Disability Insurance Claim	✓	1.5
Social Security Claim	✓	1.5
Supplemental Security Claim	✓	1.5
Control Substitution Costs	✓	
Education Costs	✓	1.5*
Justice System Costs	✓	1.5
Prison Costs	✓	1.5
Victimization Costs	✓	
<b>Costs</b>		
Program Costs		

Note: The table above lists the components of the costs and benefits of ABC. In order for some components to be categorized as benefits, we reversed the sign of the treatment effect. \*Only education costs up until age 18 are multiplied by 1.5 to account for welfare costs.

### H.3 Computing the Benefit-Cost Ratio

The benefit-cost ratio can be expressed as,

$$\mathbb{E} \left( \frac{\sum_{t=0}^T B_t}{\sum_{t=0}^T C_t} \right), \quad (16)$$

where we let  $T = 79$ , define  $B_t$  and  $C_t$  to be the total benefits and costs of the program at time  $t$ , and define  $\mathbb{E}$  to be the sample mean. See Table H.5 for a detailed list of the components to the benefits and costs of ABC and CARE. We take the sum of the treatment effects on each component of the benefits to be the total benefits of the ABC and CARE programs (see Appendix H.1 on how we estimate these treatment effects).

To account for deadweight loss, we assume a marginal welfare cost of 50% by multiplying public costs components by a factor of 1.5. For the same reason, we multiply public-transfer

income by a factor of 0.5. We discount each component of the benefits and costs by 3% every year to obtain their net present value at birth. We then sum up the discounted components of the benefits and find the ratio with the discounted costs.

As described in Appendix H.1, we estimate the treatment effect on each component of the benefits and costs at time  $t$  for the pooled, male, and female samples. We do this for 1,000 bootstrap resamples of the original ABC and CARE data. In the case of health costs and subject income, for which we employ auxiliary datasets to estimate the treatment effects, we also obtain 1,000 bootstrap estimates from the auxiliary data for every ABC and CARE bootstrap resample, resulting in a total of 1,000,000 estimates. By reusing each bootstrap estimate of the treatment effect on outcomes that do not require any auxiliary data set 1,000 times, we obtain a total of 1,000,000 estimates of the costs stream and benefits stream. We estimate the benefit-cost ratio for each of those streams. This is how we form our empirical bootstrap distribution of the benefit-cost ratio for the pooled, male, and female samples. We take the mean of the distributions to be the point estimates, and we take the standard deviations to be the standard errors. To construct the 80% confidence intervals, we take the 10<sup>th</sup> and 90<sup>th</sup> quantiles of each bootstrap distribution.

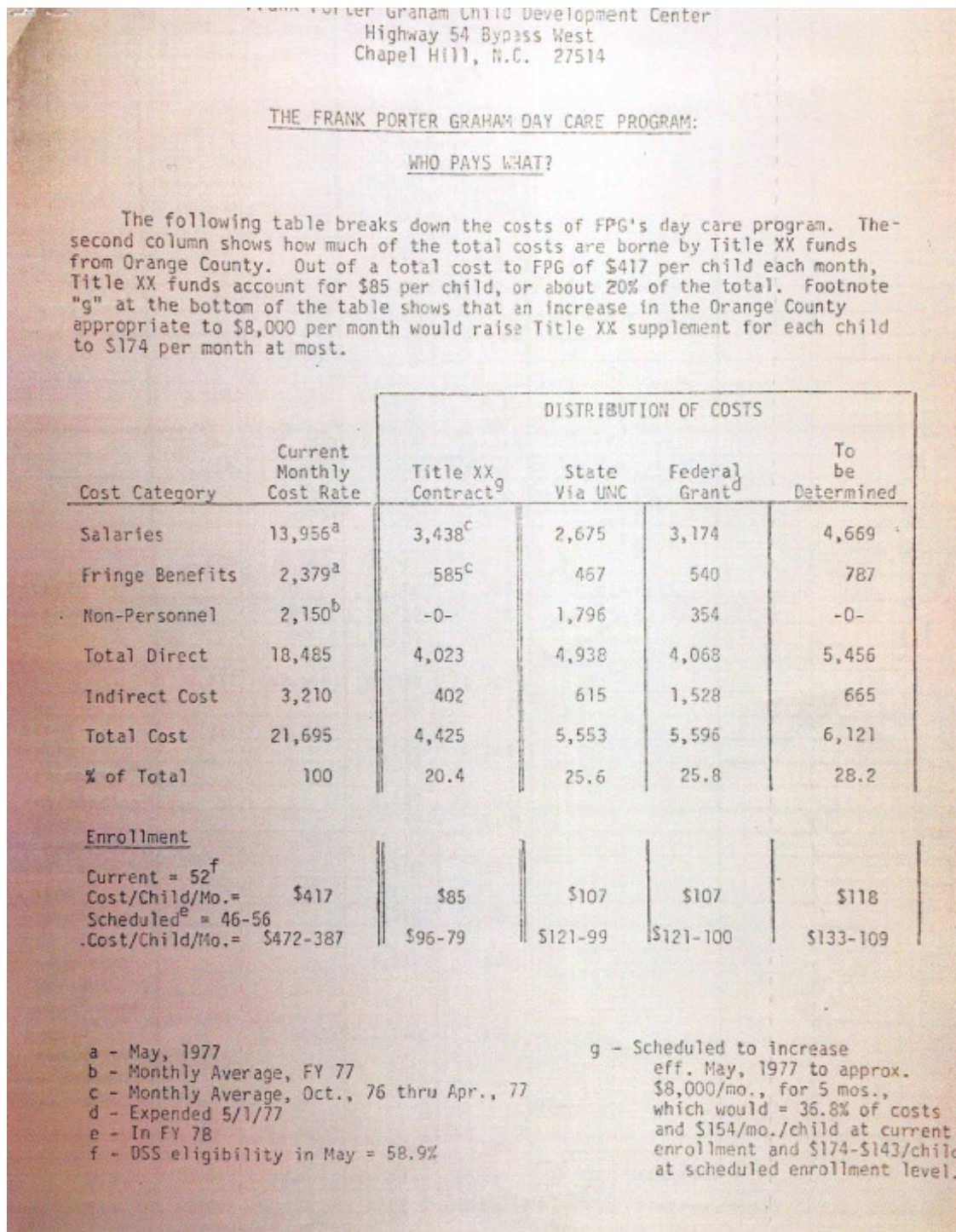
## I Program Costs<sup>79</sup>

In this appendix, we document the sources informing our programs' costs calculation in Section 2.4. We use a battery of primary sources obtained in the Archives of the University of North Carolina at Chapel Hill—Records of the Office of the Vice Chancellor [University of North Carolina Archives \(UNC\)](#). A copy of these battery of documents is available upon request. Figure I.1 exemplifies one of these sources, in which the monthly cost of treatment per child for the year 1977 is categorized by source of funding.

---

<sup>79</sup>This section of the appendix is based on joint work with Sylvi Kuperman.

Figure I.1: Primary-source Document Costs, Example



Note: This figure is a photograph exemplifying the primary-source document we use in the calculation of the programs' cost. It was obtained in the University of North Carolina at Chapel Hill—Records of the Office of the Vice Chancellor

Figure I.2 is another of our sources and it is the base of our calculation of personnel costs.

Figure I.2: Primary-source Document Costs, Personnel Wages

BUDGET ESTIMATE		
I. Personnel Services		
A. Personnel	1973- 1974	1974- 1975
1 Occupational Therapist	\$10,500	11,500
1 Physical Therapist	10,500	11,500
1 Nurse	11,000	12,100
1 Pediatrician	22,000	24,000
1 Social Worker	11,000	12,100
1 Speech Therapist	11,000	12,100
1 Special Educator	11,000	12,100
½ Psychiatrist @ \$25,000	12,500	13,700
½ Recreator @ \$9,500	4,750	5,200
½ Vocational Rehabilitation Counselor @ \$12,000	6,000	6,600
1 Psychologist	18,000	19,800
3 Secretaries @ \$5,500	16,500	17,325
Total Personnel	<u>144,750</u>	<u>157,925</u>
B. Social Security and Retirement	<u>21,134</u>	<u>22,346</u>
Total Personnel Services	\$165,884	180,271
II. Other		
Supplies	3,000	3,150
Travel	2,500	2,650
Postage, Communications	1,000	1,100
Total Other	<u>6,500</u>	<u>6,900</u>
Grand Total	\$172,384	187,171

Note: This figure is a photograph provides the estimates for the personnel wages we use. It was obtained in the University of North Carolina at Chapel Hill—Records of the Office of the Vice Chancellor

Interviews with the programs' staff [Kuperman and Cheng \(2014\)](#); [Kuperman \(2015\)](#), inform us about additional costs of the programs. An example is the salary of a social worker, who is not part of some of the of the costs estimates reported before but was part of the staff

implementing treatment. These interviews are available upon request.

Finally, a valuable source is a report written by the program staff [Frank Porter Graham Child Development Center \(1979\)](#). As we note in Section 2.4, this report produced an estimate of the cost in a completely independent way, although perhaps using a the same or similar primary sources. Our calculation of the costs comes very close to those in [Frank Porter Graham Child Development Center \(1979\)](#).

## J Costs of Education

We account for short- and long-term components of education costs. The short-term components include savings due to reductions in special education and grade retention. The long-term components include the type and level of the highest educational attainment at age 30. We do not calculate costs of education beyond age 30 because we do not have data on the subjects' later educational attainment. Instead of forming a projection, we do not add further modeling uncertainty through this component and document that, at the national level, education beyond age 30 increases marginally. To document this, we use the Panel Study of Income Dynamics (PSID) for a representative sample of individuals born between 1972 and 1982.

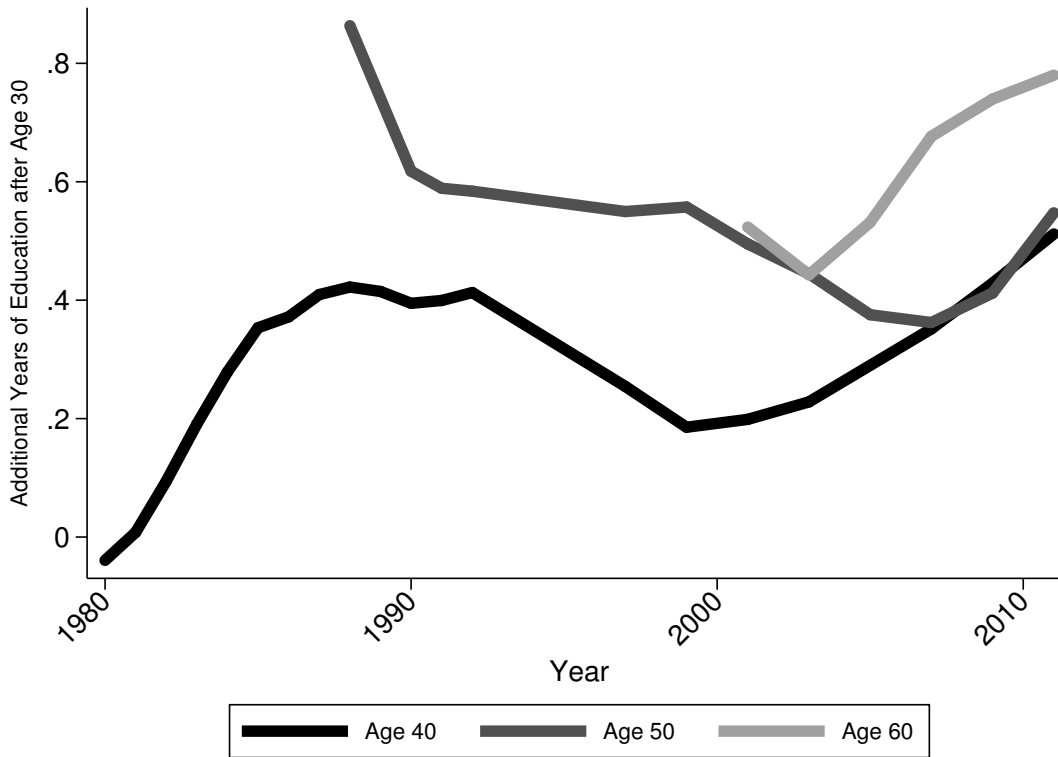
To estimate the costs of additional schooling, we combine various sources. Table J.1 describes the yearly cost of education at every level and the age and duration for which these costs are incurred. We apply these costs additively up to the highest level of educational attainment by age 30. Pooling males and females, the treatment groups had on average higher attainment and incurred a greater cost of education. To find the present value of the difference between the treatment and control groups, we first order educational attainment per Table J.1. We find a difference between the average educational attainment of the treatment groups (finish

community college) and the average educational attainment of the control groups (start community college). This can be represented as a cost that is \$12,586 higher for the treatment groups than for the control groups, as in Table J.1. The effect of the program on the educational attainment of females, however, is much greater than the effect on that of males.

## J.1 Measuring Lifetime Educational Attainment

Follow-up data on educational attainment were collected for ABC and CARE subjects up to age 30, on average. This may not necessarily be an accurate measure of lifetime educational attainment. Thus, we perform an exercise using nationally representative data from the Panel Study of Income Dynamics (PSID) to assess educational attainment after age 30. Figure J.1 shows the results of our exercise.

Figure J.1: Additional Education After Age 30



Note: This plot uses data from the Panel Study of Income Dynamics (PSID) to display the mean years of education beyond the education level at age 30, by years. We report this value for ages 40, 50, and 60. The domain of the trends is a function of data availability, i.e. of longitudinal span in the data. Each yearly calculation uses cross-sectional weights to attain nationally representative statistics.

Figure J.1 displays the average additional years of education at ages 40, 50, and 60. The series have different spans due to restrictions in data availability. Individuals, on average, never go a year beyond their education at age 30. We assume that imprecision generated by not observing educational levels beyond age 30 is of second order, and do not project additional educational attainment.

## **J.2 Cost of Education**

We apply the costs described in Table J.1 to subjects' educational attainment at age 30 to calculate the public and private costs of lifetime educational attainment. Costs up to high school are assumed to be public costs.

Table J.1: Yearly Individual Education Costs

Schooling level	Ages	Duration (Years)	Yearly Cost	Attainment & Notes
K-11	6-16	10	\$9,113	All subjects. Only 1 control individual dropped out before completing 9th grade; we assume completion up to 8th grade. Counted only for subjects who completed high school, 29:38 (control:treatment).
Grade 12	17-18	2	\$9,113	
GED (Started)	19	.5	\$155	GED is considered a one year program. No subjects identified as having starting a GED program without finishing.
GED (Completed)	19	1	\$155	8:6 (control:treatment).
Vocational or Technical Training (Started)	19	1	\$7,245	8:3 (control:treatment).
Vocational or Technical Training (Completed)	19-20	2	\$7,245	4:9 (control:treatment)
Community College (Started)	19	1	\$7,001	11:6 (control:treatment)
Community College (Completed)	19-20	2	\$7,001	9:8 (control:treatment)
College (Started)	19	1	\$7,685	Assume dropouts drop out at the end of the first year. 5:8 (control:treatment).
College (Completed)	19-22	4	\$11,886	3:8 (control:treatment)
Graduate School (Started)	23	1	\$9,704	Assume dropouts drop out at the end of the first year. 2 treated individuals.
Finished Masters	23-25	2	\$9,704	1 treated individual
Finished PhD	23-26	4	\$9,704	1 treated individual
Grade Retention	NA	1	\$9,113	
Special Ed.	NA	1	\$11,705	

Sources: [Snyder and Dillow \(2012\)](#); [Hoenack and Weiler \(1975\)](#); [Dhanidina and Griffith \(1975\)](#); [Freeman \(1974\)](#).

Note: This table reports the yearly cost and duration of each type of education, as well as the ages for which we evaluate them. All amounts are inflated to 2014 USD. We show the number of subjects who identified themselves as being in each education category (total number of respondents: 101/114). To compute the total cost of education for a subject, we applied these costs additively up to the highest level of educational attainment. Only K-12 education, special education, and grade retention costs account for deadweight loss. Because it gives costs that are applied across many years, this table does not show their present discounted value.

### J.3 Non-monetary Benefits of Education

There are many social and non-monetary benefits of education that our analysis cannot capture. These benefits impact the individual's quality of life, the general well-being of

society through positive peer effects as well as fewer costs and negative externalities, and even the well-being of future generations. Documenting them all may be impossible, but we briefly review some major benefits in this section. Vila (2000) documents private benefits with external effects, such as health (increases in longevity and better nutrition and preventative care choices). Higher education is also associated with decreased fertility rates linked with improved infant health and lower mortality rates. Moreover, higher education not only improves labor outcomes with respect to employment prospects and salary, but also with regard to how individuals perceive work and leisure, with more education leading to increased satisfaction from leisure. Furthermore, higher education is linked with better savings behavior and higher rates of return on savings. Higher education is also connected with social stability—better education promotes good citizenship and creates communities that are less likely to experience violent social conflict.<sup>80</sup>

## K Quantifying the Benefits in Crime Reduction

### K.1 Data Description

The crime data available in ABC and CARE come from four different sources provided by the program, which we supplement with auxiliary datasets. We summarize the ABC and CARE datasets and auxiliary datasets related to crime below.

#### K.1.1 ABC and CARE Datasets

1. Administrative Youth Arrests Dataset. This dataset is only available for ABC. This dataset records all arrests up to the age at which the data were obtained (April, 1996), when ABC subjects were about 21 years old. The categories of crimes in this dataset are coarser than the ones that we use in the other datasets: it categorizes crimes into

---

<sup>80</sup>Lochner (2011b) or Lochner (2011a).

property, violent, drug, and miscellaneous crimes. We assume some equivalences, as shown in Table K.1.

2. Administrative Adult Arrests Dataset. Gathered when ABC and CARE subjects were around 34 years old, this dataset includes individual data on arrests, with short descriptions of the associated offenses. This dataset includes some subjects for whom the arrest history is missing. To resolve this, we use a methodology (detailed below) involving the sentences data, which does not have missing values.
3. Administrative Sentences Dataset. Gathered when ABC and CARE subjects were about 34 years old, this dataset includes individual data on sentences, with descriptions of the crimes. It also includes total sentence length (projected and actual) and punishment type (jail, prison, parole).
4. Self-reported Adult Crimes Dataset. A module on crimes was included at the age-21 and age-30 interviews for both ABC and CARE. After matching all datasets, we use the information from the self-reported crimes whenever a particular crime does not appear in the other datasets.

### K.1.2 Auxiliary Datasets

1. National Crime Victimization Survey (NCVS). The NCVS is a survey (self-reported) on crime victimization reported on the household level. The NCVS does not cover crimes to businesses, and it might under-report crimes that might not be known to all members of a family, such as rape. It also does not give statistics for murders. The data are available from 1993 to 2013. We use NCVS to estimate the total number of victims per type of crime in the U.S., which is used to construct victims-to-arrests ratios.
2. Uniform Crime Reporting Statistics (UCRS). This dataset contains all crimes that are reported to the police. It contains crimes to households, individuals, and businesses

that are captured by most of the law enforcement agencies in the country. We used data from UCRS spanning the years 1996 to 2012. We use UCRS as a complement to NCVS when we estimate the total number of victims per type of crime in the U.S.

3. National Judicial Reporting Program (NJRP). We use the NJRP to get data for total number of sentences in the U.S. The data were taken from reports published biennially from 1986 to 2006. We use this dataset to construct arrest-to-sentences ratios for our crime categories.
4. North Carolina Department of Public Safety dataset (NCDPS). This dataset contains information since 1972 on each individual who is convicted of a crime and enters the state prison system or community supervision in North Carolina. The data do not include arrests, or sentences involving jail or unsupervised probation. Because the North Carolina Department of Public Safety was created 3 years ago to consolidate the state’s Department of Correction, Department of Juvenile Justice, and Crime Control, among other state agencies, these data were mostly constructed by the other agencies. We use this dataset to fit a prediction model that we use to predict future crimes for ABC and CARE subjects.

### **K.1.3 Crime Categories**

The administrative adult datasets have descriptions of all crimes committed by ABC and CARE subjects. We categorize these crimes to be as comparable as possible to the categories in the other data sources and in the literature on calculating the cost of crime. However, it is inevitable to have some crimes that do not clearly fit into the broad categories. As shown in Table K.1, the different experimental and auxiliary datasets and the literature use different crime categories. The categorization we present is our effort to make all sources comparable.

Table K.1: Crime Categories

Our Categories	Youth Data	Costs of Crime	Nat. Arrests Data	Nat. Sentences Data
Arson Assault	Violent	Arson Assault	Arson Total assaults	Aggravated assaults
Burglary Fraud		Household burglary Fraud	Burglary Fraud, Forgery, Embezzlement	Burglary Fraud, Forgery
Larceny Miscellaneous	Property Drug, Misc.	Larceny/theft	Larceny Drug abuse total	Larceny/theft Drug offenses
Vehicle Theft Murder		MV theft Murder	MV theft Murder, Non-negligent manslaughter	MV theft Murder, Manslaughter
Rape		Rape/sexual assault	Forcible rape, Sex offense	Rape
Robbery Vandalism		Robbery Vandalism	Robbery Vandalism	

Note: This table shows the various measures we have for our categories of crimes from each dataset. “Costs of Crime” are from [McCollister et al. \(2010\)](#).

## K.2 Methodology for Estimating Crime Costs

In this section we give a detailed explanation of the steps taken to calculate the total treatment effect on crime and the costs associated with that effect. We first give a more abstract and formal summary of the process, and then discuss the details for each step.

1. *Count Arrests and Sentences.* We count the total number of sentences for each subject,  $i$ , and category of crime (robbery, larceny, etc.),  $j$ , up to age 34, which we denote by  $S_{i,j}^{34}$ . We also match the data on adult arrests, juvenile arrests, and self-reported crimes, to construct total number of arrests for each crime type up to that age,  $A_{i,j}^{34}$ . For some subjects, the arrest data are missing. In those cases, we impute the missing data by assuming that the national arrests-to-sentences ratio for crime type,  $j$ , is valid for each subject. Let  $\bar{A}_j$  be the national total number of arrests for crime type,  $j$ , and let  $\bar{S}_j$  be the national total number of sentences. Then, we construct  $r_j = \frac{\bar{A}_j}{\bar{S}_j}$ , and we impute  $A_{i,j}^{34} = r_j S_{i,j}^{34}$ .
2. *Construct Predictions.* From our external data, we have a dataset to predict lifetime sentences. In that dataset, we use sentences up to age 34 in all types of crime to

predict future sentences for that crime type,  $\widehat{S}_{i,j}^{35-50}$ . This gives an estimate of the lifetime sentences as  $\widehat{S}_{i,j} = S_{i,j}^{34} + \widehat{S}_{i,j}^{35-50}$ . Given that we do not have an analogous dataset to predict lifetime arrests, we impute the predicted arrests as a linear function of the predicted number of sentences:  $\widehat{A}_{i,j}^{35-50} = r_j \widehat{S}_{i,j}^{35-50}$ . Then, we calculate  $\widehat{A}_{i,j} = A_{i,j}^{34} + \widehat{A}_{i,j}^{35-50}$ .

3. *Estimate Number of Victims.* Let the national number of victims of a given type of crime be  $\overline{V}_j$ . We construct a victimization inflation factor for each crime type:  $f_j = \frac{\overline{V}_j}{A_j}$ . It represents the number of times someone is arrested as a fraction of the number of victims of the crimes. Then, the estimated number of victims of subject,  $i$ , for crime type,  $j$ , based on arrests is estimated as  $\widehat{V}_{i,j}^A = A_{i,j} f_j$ . For sentences, we calculate an analogous estimate of victims based on the victimization inflation factor and the arrests-to-sentences ratio:  $\widehat{V}_{i,j}^S = S_{i,j} f_j r_j$ . Both estimates are similar, as we show below. We construct our final estimate of the lifetime victims of subject,  $i$ , for crime type,  $j$ , as the average of both estimates to achieve greater precision:  $\widehat{V}_{i,j} = \left( \widehat{V}_{i,j}^A + \widehat{V}_{i,j}^S \right) / 2$ .
4. *Find Total Costs of Crimes.* We use estimates of the cost of crimes for victims from the literature for each crime type  $j$ ,  $c_j^V$ . We impute the total victim costs of subject,  $i$ , for crime type,  $j$ , as  $\widehat{C}_{i,j}^V = \widehat{V}_{i,j} c_j^V$ . We also calculate different costs from the justice system (including police) associated with the different crime types, but only for the ones that included arrests or sentences (i.e. we do not consider the victimization inflation), as:  $C_{i,j}^{JS} = \widehat{A}_{i,j} c_j^{JS}$ . Finally, we also construct the total costs of incarceration for crime type,  $j$ ,  $\widehat{C}_{i,j}^P$  as the total time the subject was imprisoned for that type of crime,  $P_{i,j}$ , multiplied by the cost of a day in prison  $c_P$ . All of our cost estimates are discounted to birth.

### K.2.1 Count Arrests and Sentences

Unlike previous studies, we use several datasets to construct measures of criminal activity of the ABC or CARE subject. We generate a count of the number of arrests and of the number of sentences by crime category for each ABC and CARE subject. We start by counting sentences, which is trivial, as we have complete information on sentences for all ABC and CARE subjects in one dataset.

Unlike counting sentences, the count of arrests is more involved.<sup>81</sup> We now describe the procedure we follow to get a count of arrests that is as complete as possible. We begin by matching, based on crime description and date, the arrests from the adult administrative data, the youth administrative data, and the self-reported dataset. It is possible that the adult data are missing youth arrests that were expunged from the criminal records or crimes committed in states other than North Carolina, which were not obtained in the collection of administrative data. Because we observe the arrest dates, we confirm that we are not duplicating any arrests. To align the youth arrests data, which are categorized more broadly, we assume that all violent crimes are assaults (the most common category of violent crime in the sample) and that all property crimes are larcenies (the most common category of property crime). We categorize both drug and miscellaneous crimes in the miscellaneous category, for which we do not compute victim costs.

The main problem we confront with these data is that there are some individuals who are missing data on arrests in the data collected at age 34. We know this is the case because there are sentences for which we do not observe the corresponding arrests. While this is expected for some crimes, such as speeding or shoplifting, it is not possible for others. This seems to be due to the difficulty in collecting administrative information on arrests. Counting ar-

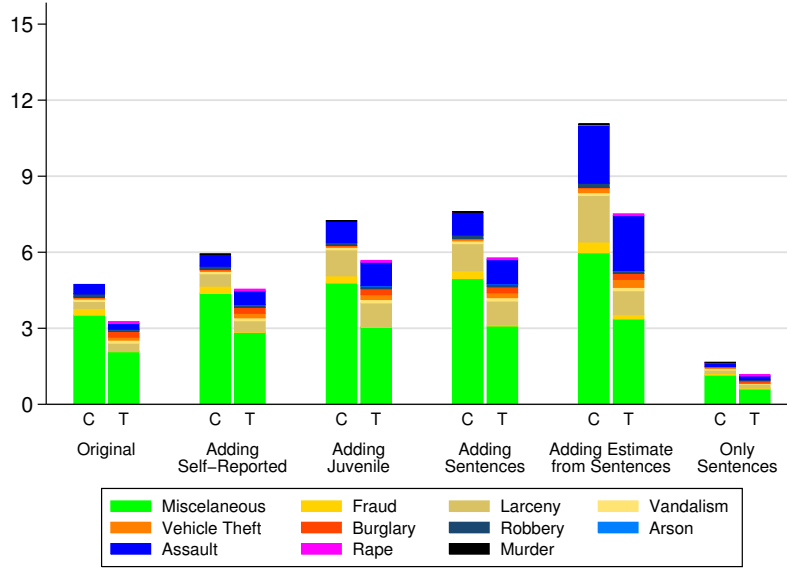
---

<sup>81</sup>Throughout this appendix, for all the calculations involving counting arrests, if the arrest was the result of more than one *offense*, we count the number and type of offenses separately, rather than counting all of them as one. An offense is the precise definition for the unit that we work with.

rests using many data sources eases of reconstructing the adult arrests data. We tackle this challenge by first identifying those individuals who should have arrests, because they have sentences that necessarily imply they were arrested at some point of time. Then we impute the number of arrests for each of those sentences, based on the national arrests-to-sentences ratio. We estimate that 11 individuals (10% of the ABC sample) have missing arrest data. We note that the final estimates of effects on reduction in crime costs using only sentences (which do not have any missing values) are very similar to the ones using arrests.

The effect of adding the different datasets, as well as the final counts of both arrests and sentences is presented in Figure K.1. Thes the first three pairs of bars present the number of additional crimes included in the data by the addition of the self-reported, juvenile, and sentences datasets. In the case of sentences, we first show the effect of imputing just one arrest per sentence, and only for individuals with missing arrests. Notice that the juvenile data only adds assaults, larcenies, and miscellaneous, as discussed above. The next pair of bars shows the effect of adding more than one arrest per sentence for the individuals with missing data, using the arrests-to-sentences ratios. The effect is significant, adding about 30% more crimes to the previous estimate. Importantly, some rape arrests are added to the control group in this step, because of one subject who presented a sentence for that crime. Finally, the total number of sentences is far lower than the total number of arrests, even if only the original arrests are counted. We also note in this chart that the volume of the crimes is mostly driven by miscellaneous crimes (which are mostly drug and traffic crimes). As these crimes are counted as victimless in our methodology, their effect on the final estimates is much smaller than what this chart might suggest.

Figure K.1: Counts of Arrests and Sentences



Note: This chart shows the arrests per capita for the control (C) and treatment (T) groups. The first pair of bars shows the original arrests data from the administrative adult dataset. The next pair adds the self-reported crimes that did not match with the original arrests data. The next pair adds data from the administrative juvenile dataset that did not match with the previous datasets. The next pair adds one arrest for every sentence that did not match with the previous datasets and one arrest for every sentence that had arrest data missing. The next pair adds  $n$  arrests instead of one for each sentence, where  $n$  is calculated using the arrests-to-sentences ratio obtained from auxiliary datasets. The final pair of bars, for reference, is the total number of sentences from the administrative sentences dataset.

## K.2.2 Construct Predictions

The data available describe the ABC and CARE subjects' crimes committed up to age 34. However, we believe that the effect of the programs on crime does not stop at that age. To predict the number of crimes that the study participants commit beyond age 34, we use data from the North Carolina Department of Public Safety (NCDPS). Because the crime data are obtained from the same state in which ABC and CARE were implemented, the prediction model is appropriate for the ABC and CARE samples. To the best of our knowledge, no study of the effects of an early childhood education program has ever used microdata to estimate a predictive model for future crimes. The estimations in Heckman et al. (2010) are based on national age ratios (crimes of a certain category committed by older people over

crimes of the same category committed by younger people at a specific time), rather than on microdata. However, age ratios only consider the same category of crime as an input to the estimation model and the results can reflect demographic transitions.

It would be ideal to use prediction models estimated from the same cohorts as the ABC and CARE subjects. However, it is not possible to predict crimes committed until older ages because the ABC and CARE subjects are currently about 46 years old. To predict crimes at older ages, it is necessary to use earlier cohorts. The data are available from 1972 (44 years ago as of 2016), and thus they do not contain a complete criminal history for any cohort of individuals. We assume that few crimes are committed after the age of 50.

We separately estimate predictions from ages 35-40, 40-45, and 45-50. We have plentiful observations to estimate crimes in all the age ranges. We calculate our predictions using a five-step procedure:

1. Find individuals that are at least 40 years old as of 2016 in the NCDPS dataset.
2. Regress the number of crimes of each type at ages 35–40 on those at ages 16–34.
3. Use the estimated prediction model for ABC individuals, attributing those crimes to age 40 for discounting purposes.
4. Repeat the three previous steps for ages 40–45 and 45–50.
5. For individuals with no criminal histories before age 34, assume that they commit no crimes after 34.

We perform the previous process separately for males and females. We use linear predictions, but replace negative predicted values of crime by zero. Table [K.2](#) through Table [K.7](#) show the estimated models. As expected, generally the most important prediction factor for a crime is the number of occurrences of the same crime type in a previous period. The

coefficients in some cases are substantial, which implies that considering predicted crimes is an important part of an assessment of the crime benefits of the program.

Table K.2: NCDPS Regressions of Ages 35–40 on Ages 16–35, Females

	Miscellaneous	Fraud	Larceny	Vandalism	Auto Theft	Burglary	Robbery	Arson	Assault	Rape	Murder
Miscellaneous	0.103 (27.71)**	-0.029 (5.90)**	0.009 (6.00)**	0.001 (3.88)**	0.001 (5.18)**	0.001 (2.25)*	0.000 (1.91)	-0.000 (0.62)	0.002 (3.16)**	-0.000 (1.27)	-0.001 (2.85)**
Fraud	-0.047 (21.71)**	0.104 (36.61)**	0.003 (3.81)**	-0.000 (2.31)*	0.000 (0.60)	0.000 (2.02)*	0.000 (0.02)	0.000 (1.76)	-0.001 (1.48)	-0.000 (0.75)	-0.000 (1.72)
Larceny	0.034 (8.16)**	0.002 (0.33)	0.105 (62.36)**	0.001 (2.37)*	0.002 (10.35)**	0.002 (7.80)**	0.001 (5.97)**	0.000 (2.99)**	0.004 (5.65)**	0.000 (1.30)	-0.000 (1.67)
Vandal	0.000 (0.01)	-0.036 (1.10)	-0.024 (2.43)*	0.012 (6.40)**	0.004 (2.86)**	0.007 (3.87)**	0.005 (3.34)**	0.003 (3.40)**	0.043 (9.27)**	0.000 (0.79)	-0.000 (0.23)
Auto Theft	0.059 (1.55)	-0.026 (0.52)	0.055 (3.57)**	0.004 (1.28)	0.012 (5.94)**	0.007 (2.48)*	0.005 (2.52)*	-0.002 (1.30)	-0.014 (2.02)*	-0.000 (0.23)	0.002 (1.02)
Burglary	0.049 (1.77)	-0.031 (0.86)	0.011 (1.01)	0.002 (0.95)	0.006 (3.86)**	0.044 (22.56)**	0.000 (0.20)	-0.001 (1.26)	-0.003 (0.55)	-0.000 (0.25)	-0.001 (0.34)
Robbery	0.030 (1.05)	0.054 (1.46)	0.040 (3.51)**	0.002 (1.03)	0.005 (3.43)**	-0.001 (0.28)	0.028 (17.69)**	0.002 (2.22)*	-0.000 (0.07)	-0.000 (0.29)	0.002 (1.50)
Arson	0.070 (1.29)	-0.087 (1.22)	-0.035 (1.61)	0.008 (1.98)*	0.011 (3.98)**	0.006 (1.62)	0.025 (8.19)**	0.010 (4.95)**	0.017 (1.67)	-0.000 (0.19)	0.005 (1.60)
Assault	0.064 (5.74)**	-0.060 (4.11)**	0.007 (1.63)	0.007 (7.82)**	0.001 (1.83)	0.002 (1.93)	-0.000 (0.14)	0.002 (4.52)**	0.048 (22.78)**	-0.000 (0.36)	0.000 (0.34)
Rape	-0.015 (0.15)	-0.131 (0.98)	-0.003 (0.08)	-0.003 (0.42)	-0.003 (0.49)	-0.004 (0.55)	-0.006 (0.99)	-0.001 (0.24)	0.001 (0.08)	-0.000 (0.06)	-0.002 (0.33)
Murder	0.006 (0.16)	-0.166 (3.44)**	-0.008 (0.54)	0.003 (1.19)	-0.002 (0.94)	-0.002 (0.75)	-0.002 (1.22)	0.003 (2.24)*	0.001 (0.12)	-0.000 (0.34)	-0.001 (0.37)
constant	0.069 (13.39)**	0.267 (39.01)**	0.041 (19.54)**	0.004 (10.27)**	0.001 (2.76)**	0.001 (3.17)**	0.001 (4.80)**	0.001 (3.52)**	0.023 (23.81)**	0.000 (3.38)**	0.004 (12.90)**
$R^2$	0.02	0.02	0.07	0.00	0.01	0.01	0.01	0.00	0.01	0.00	0.00
$N$	63,515	63,515	63,515	63,515	63,515	63,515	63,515	63,515	63,515	63,515	63,515

Note: This table shows how the crimes at ages 35–40 (in the columns) are predicted by crimes at ages 16–35 (in the rows). The estimations are predicted using the North Carolina Department of Public Safety dataset, which contains information on all individuals that have ever been sentenced in North Carolina. We use linear regressions in all cases. The sample for this model is limited to individuals who are at least 40 years old. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

Table K.3: NCDPS Regressions of Ages 35–40 on Ages 16–35, Males

	Miscellaneous	Fraud	Larceny	Vandalism	Auto Theft	Burglary	Robbery	Arson	Assault	Rape	Murder
Miscellaneous	0.095 (73.29)**	-0.002 (2.13)*	0.004 (8.97)**	0.000 (2.71)**	0.001 (6.47)**	0.001 (5.87)**	0.000 (1.50)	-0.000 (0.12)	0.005 (14.73)**	-0.001 (3.28)**	-0.001 (5.83)**
Fraud	-0.025 (13.76)**	0.107 (67.23)**	0.007 (11.68)**	0.000 (0.47)	0.001 (3.44)**	0.002 (5.27)**	0.001 (2.53)*	-0.000 (0.61)	-0.000 (0.34)	-0.000 (0.35)	-0.000 (1.96)*
Larceny	0.041 (14.34)**	0.012 (4.83)**	0.092 (97.06)**	0.001 (5.13)**	0.005 (14.85)**	0.010 (19.72)**	0.004 (12.30)**	0.000 (1.79)	0.009 (11.37)**	-0.000 (0.46)	-0.000 (0.59)
Vandal	0.041 (4.51)**	-0.017 (2.13)*	-0.007 (2.22)*	0.019 (23.49)**	0.002 (1.96)	0.002 (1.06)	0.002 (1.76)	0.002 (5.41)**	0.029 (11.26)**	0.001 (1.06)	0.000 (0.62)
Auto Theft	0.015 (1.58)	0.014 (1.70)	0.005 (1.48)	0.002 (1.86)	0.043 (39.51)**	0.013 (7.87)**	0.000 (0.14)	0.000 (0.13)	0.008 (3.12)**	0.001 (1.28)	0.001 (1.18)
Burglary	0.005 (1.26)	0.011 (3.06)**	0.010 (6.86)**	0.002 (4.33)**	0.004 (9.17)**	0.040 (52.89)**	0.003 (6.55)**	-0.000 (0.30)	0.003 (2.34)*	0.001 (2.37)*	-0.000 (0.73)
Robbery	-0.032 (5.06)**	0.000 (0.01)	0.011 (5.42)**	0.000 (0.67)	0.003 (4.61)**	0.006 (5.67)**	0.019 (26.50)**	-0.000 (1.00)	-0.000 (0.21)	0.000 (0.00)	0.001 (1.68)
Arson	0.008 (0.26)	-0.023 (0.82)	-0.010 (0.95)	0.012 (4.40)**	-0.004 (0.98)	0.008 (1.44)	0.004 (1.22)	0.007 (6.73)**	0.031 (3.49)**	-0.001 (0.21)	0.001 (0.39)
Assault	0.041 (11.14)**	-0.010 (3.18)**	-0.000 (0.09)	0.005 (15.94)**	-0.001 (1.60)	0.000 (0.12)	0.002 (4.47)**	0.000 (2.56)*	0.051 (49.57)**	0.001 (3.43)**	0.001 (3.29)**
Rape	-0.003 (0.33)	-0.016 (1.77)	-0.010 (2.86)**	-0.002 (2.32)*	-0.002 (1.72)	-0.006 (3.42)**	-0.002 (1.66)	-0.000 (0.54)	-0.004 (1.21)	0.011 (8.61)**	-0.001 (1.69)
Murder	-0.086 (6.33)**	-0.053 (4.45)**	-0.025 (5.51)**	-0.002 (1.71)	-0.004 (2.54)*	-0.011 (4.31)**	-0.001 (0.87)	-0.001 (1.53)	-0.007 (1.79)	-0.003 (1.56)	0.003 (3.35)**
constant	0.213 (69.25)**	0.124 (45.55)**	0.031 (29.74)**	0.005 (18.57)**	0.004 (10.95)**	0.011 (19.97)**	0.006 (17.78)**	0.001 (9.22)**	0.048 (55.54)**	0.009 (24.70)**	0.006 (27.76)**
$R^2$	0.03	0.02	0.05	0.01	0.01	0.02	0.01	0.00	0.02	0.00	0.00
$N$	230,706	230,706	230,706	230,706	230,706	230,706	230,706	230,706	230,706	230,706	230,706

Note: This table shows how the crimes at ages 35–40 (in the columns) are predicted by crimes at ages 16–35 (in the rows). The estimations are predicted using the North Carolina Department of Public Safety dataset, which contains information on all individuals that have ever been sentenced in North Carolina. We use linear regressions in all cases. The sample for this model is limited to individuals who are at least 40 years old. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

Table K.4: NCDPS Regressions of Ages 40–45 on Ages 16–35, Females

	Miscellaneous	Fraud	Larceny	Vandalism	Auto Theft	Burglary	Robbery	Arson	Assault	Rape	Murder
Miscellaneous	0.020 (4.33)**	-0.006 (5.18)**	-0.001 (0.85)	0.000 (0.78)	0.000 (0.46)	0.000 (1.29)	-0.000 (0.11)	0.000 (0.84)	0.001 (0.64)	0.000 (3.66)**	-0.001 (2.65)**
Fraud	0.046 (19.30)**	0.003 (4.41)**	0.002 (2.86)**	0.000 (1.50)	0.001 (4.86)**	-0.000 (0.56)	0.000 (1.62)	-0.000 (0.62)	-0.000 (1.13)	0.000 (2.50)*	-0.000 (0.44)
Larceny	0.025 (5.12)**	-0.002 (1.54)	0.070 (41.35)**	0.000 (0.06)	0.001 (3.61)**	0.002 (6.68)**	0.000 (0.75)	-0.000 (1.27)	0.004 (5.02)**	-0.000 (0.45)	-0.000 (0.45)
Vandal	-0.016 (0.52)	-0.009 (1.22)	-0.018 (1.71)	0.002 (1.02)	-0.001 (0.58)	0.005 (2.66)**	0.001 (0.38)	0.003 (3.06)**	0.019 (3.54)**	-0.000 (0.71)	-0.001 (0.49)
Auto Theft	0.012 (0.25)	-0.011 (0.87)	0.010 (0.56)	0.004 (0.95)	0.012 (4.90)**	0.013 (4.70)**	-0.001 (0.36)	-0.001 (0.49)	-0.017 (1.98)*	0.007 (11.24)**	0.001 (0.41)
Burglary	-0.034 (1.05)	0.019 (2.33)*	0.007 (0.64)	-0.003 (1.05)	0.005 (3.12)**	0.007 (3.66)**	0.004 (2.09)*	0.000 (0.22)	-0.004 (0.67)	0.001 (2.93)**	-0.001 (0.55)
Robbery	0.027 (0.81)	0.004 (0.47)	0.021 (1.81)	0.004 (1.52)	0.002 (1.21)	0.007 (3.98)**	0.015 (6.93)**	-0.001 (0.58)	0.018 (3.19)**	-0.000 (0.67)	-0.001 (0.82)
Arson	-0.042 (0.65)	-0.016 (0.97)	-0.025 (1.09)	0.013 (2.63)**	-0.001 (0.46)	-0.002 (0.57)	0.001 (0.31)	-0.001 (0.44)	0.001 (0.12)	-0.000 (0.06)	-0.002 (0.56)
Assault	-0.014 (1.00)	-0.008 (2.44)*	-0.007 (1.41)	0.002 (1.84)	-0.000 (0.20)	-0.001 (1.21)	0.002 (2.10)*	0.001 (2.34)*	0.028 (11.83)**	-0.000 (1.23)	-0.000 (0.41)
Rape	0.169 (1.36)	-0.014 (0.45)	-0.029 (0.67)	-0.003 (0.36)	-0.002 (0.29)	-0.003 (0.46)	-0.004 (0.57)	-0.000 (0.07)	-0.016 (0.75)	-0.000 (0.05)	0.011 (1.91)
Murder	-0.169 (4.25)**	-0.022 (2.18)*	-0.027 (1.93)	-0.004 (1.38)	-0.002 (0.84)	-0.002 (0.91)	0.001 (0.55)	-0.001 (0.64)	0.010 (1.46)	-0.000 (0.06)	-0.001 (0.50)
constant	0.304 (53.09)**	0.029 (20.43)**	0.046 (22.75)**	0.005 (10.59)**	0.002 (5.51)**	0.002 (4.90)**	0.002 (4.93)**	0.001 (4.52)**	0.022 (22.64)**	-0.000 (0.48)	0.003 (11.16)**
$R^2$	0.01	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$N$	49,738	49,738	49,738	49,738	49,738	49,738	49,738	49,738	49,738	49,738	49,738

Note: This table shows how the crimes at ages 40–45 (in the columns) are predicted by crimes at ages 16–35 (in the rows). The estimations are predicted using the North Carolina Department of Public Safety dataset, which contains information on all individuals that have ever been sentenced in North Carolina. We use linear regressions in all cases. The sample for this model is limited to individuals who are at least 45 years old. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

Table K.5: NCDPS Regressions of Ages 40–45 on Ages 16–35, Males

	Miscellaneous	Fraud	Larceny	Vandalism	Auto Theft	Burglary	Robbery	Arson	Assault	Rape	Murder
Miscellaneous	0.065 (40.59)**	-0.002 (4.98)**	0.003 (6.29)**	0.001 (4.35)**	0.001 (4.68)**	0.001 (3.18)**	-0.000 (2.61)**	0.000 (0.25)	0.003 (7.80)**	-0.000 (2.60)**	-0.001 (6.70)**
Fraud	0.032 (16.63)**	0.009 (18.85)**	0.004 (6.51)**	-0.000 (0.72)	0.001 (5.79)**	0.002 (5.19)**	0.000 (2.29)*	-0.000 (0.83)	0.000 (0.84)	-0.000 (1.28)	-0.000 (0.78)
Larceny	0.029 (9.25)**	0.001 (1.16)	0.055 (54.05)**	0.002 (6.64)**	0.003 (11.18)**	0.005 (9.45)**	0.002 (8.46)**	-0.000 (2.55)*	0.004 (4.35)**	-0.000 (0.64)	0.000 (0.70)
Vandal	0.003 (0.26)	-0.001 (0.46)	-0.009 (2.68)**	0.007 (8.19)**	-0.000 (0.04)	0.002 (1.05)	0.001 (1.02)	0.000 (1.29)	0.024 (8.44)**	0.001 (0.66)	-0.000 (0.37)
Auto Theft	0.053 (4.40)**	0.002 (0.76)	0.007 (1.82)	0.001 (0.57)	0.016 (14.05)**	0.017 (8.26)**	0.002 (1.85)	0.001 (1.36)	0.010 (3.12)**	0.001 (0.63)	0.000 (0.19)
Burglary	0.012 (2.76)**	-0.002 (1.27)	0.015 (10.20)**	0.001 (2.07)*	0.003 (6.28)**	0.025 (33.40)**	0.003 (6.14)**	0.000 (2.32)*	0.003 (2.77)**	0.000 (0.96)	-0.000 (0.10)
Robbery	-0.008 (1.10)	-0.000 (0.07)	0.022 (9.41)**	0.000 (0.65)	0.001 (1.84)	0.002 (1.89)	0.012 (19.34)**	-0.000 (0.27)	0.003 (1.68)	0.000 (0.27)	0.000 (0.70)
Arson	-0.020 (0.60)	-0.003 (0.40)	0.002 (0.21)	0.007 (2.48)*	0.006 (1.85)	-0.000 (0.03)	-0.000 (0.14)	0.006 (6.06)**	0.033 (3.60)**	0.004 (1.04)	0.000 (0.08)
Assault	0.018 (4.15)**	-0.004 (3.24)**	-0.003 (2.13)*	0.003 (7.43)**	0.001 (2.47)*	0.000 (0.48)	0.001 (3.43)**	0.000 (2.96)**	0.037 (31.56)**	0.000 (1.04)	0.000 (1.45)
Rape	-0.015 (1.26)	-0.003 (1.13)	-0.010 (2.56)*	-0.000 (0.49)	-0.003 (2.38)*	-0.004 (2.11)*	0.000 (0.13)	-0.000 (1.36)	-0.001 (0.31)	0.009 (7.53)**	-0.001 (1.32)
Murder	-0.086 (5.84)**	-0.005 (1.26)	-0.020 (4.17)**	-0.002 (2.01)*	-0.001 (0.89)	-0.008 (3.38)**	-0.000 (0.37)	0.001 (2.55)*	-0.003 (0.64)	-0.001 (0.44)	0.005 (5.35)**
constant	0.337 (104.31)**	0.018 (21.78)**	0.034 (31.87)**	0.005 (17.23)**	0.004 (12.30)**	0.011 (20.05)**	0.005 (15.95)**	0.001 (8.16)**	0.048 (54.59)**	0.007 (18.92)**	0.005 (24.05)**
$R^2$	0.02	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00
$N$	188,556	188,556	188,556	188,556	188,556	188,556	188,556	188,556	188,556	188,556	188,556

Note: This table shows how the crimes at ages 40–45 (in the columns) are predicted by crimes at ages 16–35 (in the rows). The estimations are predicted using the North Carolina Department of Public Safety dataset, which contains information on all individuals that have ever been sentenced in North Carolina. We use linear regressions in all cases. The sample for this model is limited to individuals who are at least 45 years old. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

Table K.6: NCDPS Regressions of Ages 45–50 on Ages 16–35, Females

	Miscellaneous	Fraud	Larceny	Vandalism	Auto Theft	Burglary	Robbery	Arson	Assault	Rape	Murder
Miscellaneous	-0.009 (1.75)	0.000	-0.003 (1.41)	-0.001 (1.98)*	-0.000 (0.48)	-0.000 (0.16)	-0.000 (0.58)	-0.000 (1.11)	-0.001 (0.70)	-0.000 (0.36)	-0.001 (1.71)
Fraud	0.012 (5.55)**	0.000	0.000 (0.36)	0.000 (1.63)	-0.000 (0.23)	0.000 (0.07)	-0.000 (0.44)	-0.000 (0.56)	-0.001 (2.18)*	-0.000 (0.13)	-0.000 (0.76)
Larceny	0.006 (1.28)	0.000	0.037 (21.43)**	-0.001 (1.44)	0.001 (2.02)*	0.000 (1.11)	0.000 (2.27)*	-0.000 (0.71)	0.001 (1.61)	-0.000 (0.11)	-0.000 (0.95)
Vandal	-0.043 (1.34)	0.000	-0.019 (1.52)	0.007 (2.53)*	-0.001 (0.58)	-0.001 (0.55)	-0.001 (0.95)	0.001 (0.91)	0.019 (3.07)**	-0.000 (0.05)	-0.002 (0.78)
Auto Theft	0.009 (0.16)	0.000	-0.022 (1.07)	-0.002 (0.51)	-0.001 (0.34)	-0.001 (0.33)	-0.001 (0.44)	-0.000 (0.13)	0.010 (0.99)	-0.000 (0.02)	-0.001 (0.29)
Burglary	-0.001 (0.03)	0.000	0.004 (0.31)	-0.002 (0.83)	-0.001 (0.53)	0.001 (0.31)	-0.001 (0.63)	-0.000 (0.28)	-0.013 (1.95)	-0.000 (0.03)	0.001 (0.36)
Robbery	0.015 (0.48)	0.000	-0.004 (0.30)	0.001 (0.27)	-0.001 (0.51)	-0.001 (0.50)	0.004 (2.58)**	-0.000 (0.31)	0.004 (0.66)	-0.000 (0.04)	0.000 (0.20)
Arson	-0.032 (0.57)	0.000	-0.027 (1.23)	-0.003 (0.59)	-0.001 (0.22)	-0.001 (0.24)	-0.001 (0.25)	-0.000 (0.24)	0.000 (0.01)	-0.000 (0.04)	0.008 (2.29)*
Assault	0.007 (0.50)	0.000	-0.009 (1.72)	0.005 (4.13)**	0.000 (0.11)	-0.000 (0.36)	0.001 (2.29)*	0.000 (0.36)	0.018 (6.56)**	-0.000 (0.11)	0.000 (0.12)
Rape	0.130 (1.02)	0.000	0.000 (0.01)	-0.002 (0.17)	-0.001 (0.08)	-0.001 (0.10)	-0.001 (0.12)	-0.000 (0.08)	0.038 (1.54)	-0.000 (0.02)	-0.002 (0.24)
Murder	-0.129 (4.07)**	0.000	-0.023 (1.95)	0.011 (3.99)**	0.004 (1.94)	0.002 (0.90)	-0.001 (0.52)	0.002 (2.09)*	0.014 (2.35)*	-0.000 (0.12)	-0.002 (1.02)
constant	0.264 (54.62)**	0.000	0.037 (20.07)**	0.003 (7.64)**	0.001 (4.01)**	0.001 (4.65)**	0.001 (3.15)**	0.001 (4.38)**	0.017 (18.65)**	0.000 (1.12)	0.003 (8.69)**
$R^2$	0.00		0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$N$	35,432	35,432	35,432	35,432	35,432	35,432	35,432	35,432	35,432	35,432	35,432

Note: This table shows how the crimes at ages 45–50 (in the columns) are predicted by crimes at ages 16–35 (in the rows). The estimations are predicted using the North Carolina Department of Public Safety dataset, which contains information on all individuals that have ever been sentenced in North Carolina. We use linear regressions in all cases. The sample for this model is limited to individuals who are at least 50 years old. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

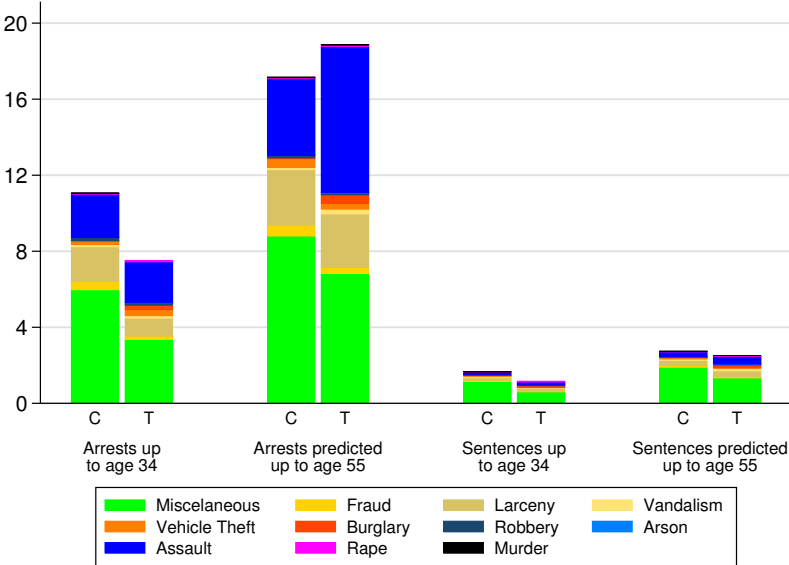
Table K.7: NCDPS Regressions from Ages 45–50 on Ages 16–35, Males

	Miscellaneous	Fraud	Larceny	Vandalism	Auto Theft	Burglary	Robbery	Arson	Assault	Rape	Murder
Miscellaneous	0.037 (19.83)**	0.000	0.003 (4.07)**	0.000 (2.77)**	0.000 (1.77)	0.001 (2.02)*	0.000 (0.64)	-0.000 (0.23)	0.001 (2.40)*	-0.001 (3.15)**	-0.001 (4.97)**
Fraud	0.020 (10.46)**	0.000	0.005 (8.12)**	-0.000 (0.04)	0.000 (1.44)	0.002 (5.94)**	-0.000 (0.68)	0.000 (0.97)	-0.000 (0.34)	0.000 (0.05)	-0.000 (1.14)
Larceny	0.023 (7.04)**	0.000	0.038 (34.47)**	-0.000 (0.33)	0.001 (3.21)**	0.004 (7.72)**	0.001 (2.74)**	0.000 (0.57)	0.002 (2.71)**	0.000 (0.72)	-0.000 (1.91)
Vandal	-0.000 (0.01)	0.000	-0.008 (2.12)*	0.007 (7.81)**	0.000 (0.00)	0.002 (0.97)	0.002 (1.87)	0.000 (0.17)	0.016 (5.00)**	-0.001 (0.39)	-0.001 (0.77)
Auto Theft	0.016 (1.07)	0.000	0.014 (2.76)**	0.001 (1.25)	0.015 (10.20)**	0.003 (1.12)	0.003 (2.24)*	-0.000 (0.99)	0.004 (0.88)	-0.002 (0.94)	-0.000 (0.34)
Burglary	0.003 (0.57)	0.000	0.013 (8.34)**	0.000 (0.08)	0.004 (8.14)**	0.014 (19.75)**	0.002 (4.39)**	0.000 (2.81)**	0.001 (0.72)	0.000 (0.02)	-0.000 (0.26)
Robbery	-0.017 (2.33)*	0.000	0.013 (5.15)**	0.001 (1.36)	0.000 (0.08)	0.001 (0.90)	0.008 (12.97)**	-0.000 (1.90)	0.005 (2.27)*	-0.001 (0.81)	-0.000 (0.63)
Arson	-0.036 (1.04)	0.000	-0.012 (1.02)	0.012 (4.70)**	-0.002 (0.59)	-0.003 (0.52)	-0.003 (0.83)	0.004 (3.73)**	0.027 (2.82)**	-0.004 (0.92)	-0.002 (0.98)
Assault	-0.004 (0.88)	0.000	-0.001 (0.72)	0.002 (6.45)**	-0.001 (1.28)	-0.002 (2.32)*	0.000 (0.06)	0.000 (0.56)	0.026 (19.80)**	0.000 (0.72)	-0.000 (0.18)
Rape	-0.025 (1.94)	0.000	-0.001 (0.25)	-0.000 (0.11)	-0.001 (0.91)	-0.001 (0.30)	0.001 (0.99)	0.000 (0.04)	-0.001 (0.40)	0.010 (6.70)**	-0.001 (1.25)
Murder	-0.068 (4.82)**	0.000	-0.015 (3.19)**	-0.001 (1.38)	-0.001 (0.83)	-0.004 (1.81)	-0.001 (1.03)	-0.000 (0.93)	-0.002 (0.42)	-0.000 (0.15)	0.000 (0.42)
constant	0.318 (103.46)**	0.000	0.026 (25.22)**	0.003 (14.03)**	0.003 (10.92)**	0.008 (16.47)**	0.003 (10.00)**	0.001 (7.13)**	0.041 (48.64)**	0.006 (16.51)**	0.004 (21.79)**
$R^2$	0.01	147,478	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
$N$	147,478	147,478	147,478	147,478	147,478	147,478	147,478	147,478	147,478	147,478	147,478

Note: This table shows how the crimes at ages 45–50 (in the columns) are predicted by crimes at ages 16–35 (in the rows). The estimations are predicted using the North Carolina Department of Public Safety dataset, which contains information on all individuals that have ever been sentenced in North Carolina. We use linear regressions in all cases. The sample for this model is limited to individuals who are at least 50 years old. \*  $p < 0.05$ ; \*\*  $p < 0.01$ .

This procedure gives a prediction of the number of sentences that the ABC and CARE subjects will receive after age 34. From the predicted number of sentences, we predict the number of arrests up to age 50. Figure K.2 shows the effect of our prediction methodology. The effect is quantitatively much larger for arrests than for sentences. For both arrests and sentences, including the predicted crimes, this effect adds 30-50% more crimes to our previous totals. The predictions are roughly proportional to the previous crimes, as discussed before.

Figure K.2: Constructed Predictions



Note: This figure continues Figure K.1. It shows, for the control (C) and treatment (T) groups, the effects of adding predictions. The first pair of columns is the same as the fifth pair of columns in Figure K.1. The second pair of columns includes the arrests that we predict. The third and fourth pairs of columns are the analogous pairs for sentences.

### K.3 Victimization Inflation

Even though we have administrative data on crimes, we only observe the crimes that had justice system consequences (arrests or sentences). However, it is possible that the subjects committed more crimes than what we observe. Victimization Inflation (VI) is a method to

capture benefits in crime reduction for crimes that did not result in justice system consequences.<sup>82</sup> For most types of crimes in the U.S., there are many more victims than arrests or sentences. Using arrests as an example, VI assumes that those “unpunished crimes” were committed by the same people who were arrested for crimes of the same type, and in the same proportion. The calculation of VI uses as an input the national ratios of total number of reported crimes over the number of arrests. VI assumes that those national ratios are also valid for each individual. Under those assumptions, it is possible to find the total number of crimes committed by a subject for a given type of crime as the total number of arrests for that type of crime multiplied by the estimated national ratios for that type of crime. We estimate the total number of victims using two methods, one based on arrests and one based on sentences. Given that the “unpunished” crimes are by definition unobserved, it is not straightforward to use a data-driven method to allocate them between those subjects with arrests, those with sentences, and those with neither arrests nor sentences. We calculate separate estimates for arrests and sentences and use the average of those estimates as our main estimate.

### **K.3.1 Construction of the Total Number of Victims in the U.S.**

The numerator of the VI ratio is an estimate of the total number of crimes of a certain crime type committed in the U.S. We construct this estimate using two datasets. First, we use the National Crime and Victimization Survey (NCVS). It has self-reported data on victimization of crimes reported on the household level. The data are available from 1995 to 2012. We also use the Uniform Crime Reporting Statistics (UCRS), which contains all crimes committed against households, individuals, and businesses that are reported to the police. These data are available from 1960 to 2013. Given that these two datasets independently underestimate the total number of crimes, but likely have significant overlap between them, we choose the

---

<sup>82</sup>Belfield et al. (2006); Heckman et al. (2010).

highest estimate among both datasets for each type of crime. We refer to this estimate,  $\overline{V_{j,t}}$ , as the total number of victims in the country for type of crime  $j$  in year  $t$ .

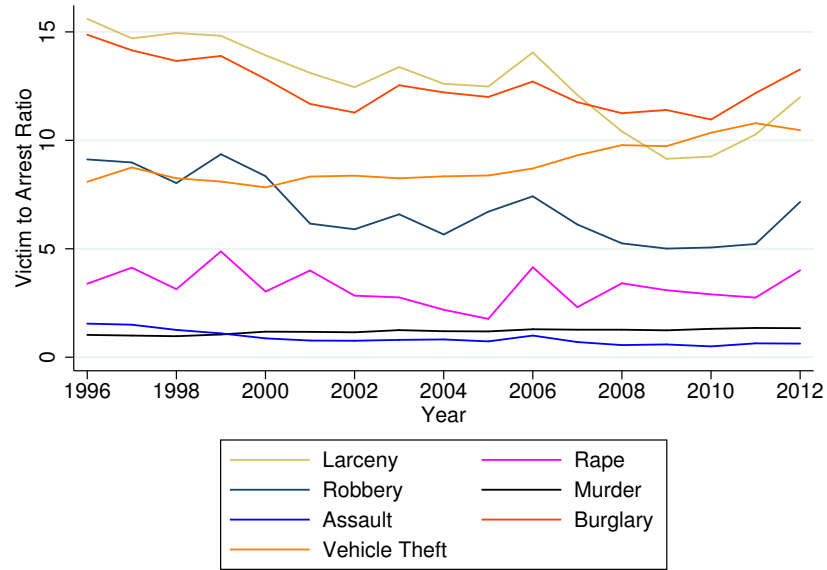
### K.3.2 Construction of the Total Number of Arrests in the U.S.

The denominator of the VI ratio is an estimate of the total number of arrests of a certain type committed in the U.S. We have data from the “National Arrests Analysis Tool” of the National Bureau of Justice Statistics. These data are available from 1980 to 2012, which spans the years of all crimes that we observe in the ABC data. There is one problem with this dataset that we consider relatively minor: not all law enforcement agencies report the number of crimes (there are dozens of agencies that can legally arrest in the U.S.). However, as a large majority of them report the numbers of crimes, and because we are using national estimates, this should not greatly affect our calculations. We use these data to create  $\overline{A_{j,t}}$ , the total number of arrests in the country for type of crime  $j$  in year  $t$ .

### K.3.3 Victimization Inflation Factors

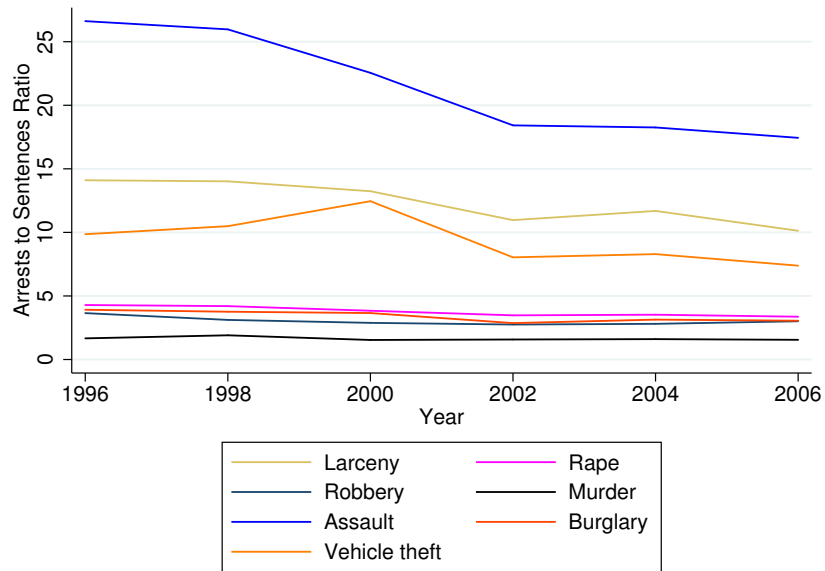
Figure K.3 and Figure K.4 show the VI factors calculated by year. The ratios in the charts are constructed as  $r_{j,t} = \frac{\overline{V_{j,t}}}{\overline{A_{j,t}}}$ . In practice, we use an average of all the yearly measures in our calculations given that this exercise imputes unobserved crimes that do not have a clearly defined date. This average of all the yearly measures is given by  $r_j = \sum_{t=t_0}^T r_{j,t} / T$ , and has more precise estimates of the ratio. The VI factors we use for sentences are equal to the factors used for arrests, multiplied by the arrests-to-sentences ratios discussed above. Below, we will discuss combining these different estimates as per our arrest-based and sentence-based methodology. For sentences, we have data from the National Judicial Reporting Program (NJRP). These data are available from 1986 onwards. Using this dataset, we construct  $\overline{S_{j,t}}$ , the total number of sentences in the country for type of crime  $j$  in year  $t$ .

Figure K.3: Victims-to-Arrests Ratios by Crime



Note: This figure shows, by year and type of crime, the number of victims (estimated from the NCVS and the UCRS datasets) divided by the number of arrests (estimated from the National Arrests Analysis Tool from the NIBRS). In practice, we use a single number for each type of crime, which is an average across years.

Figure K.4: Arrests-to-Sentences Ratio by Crime

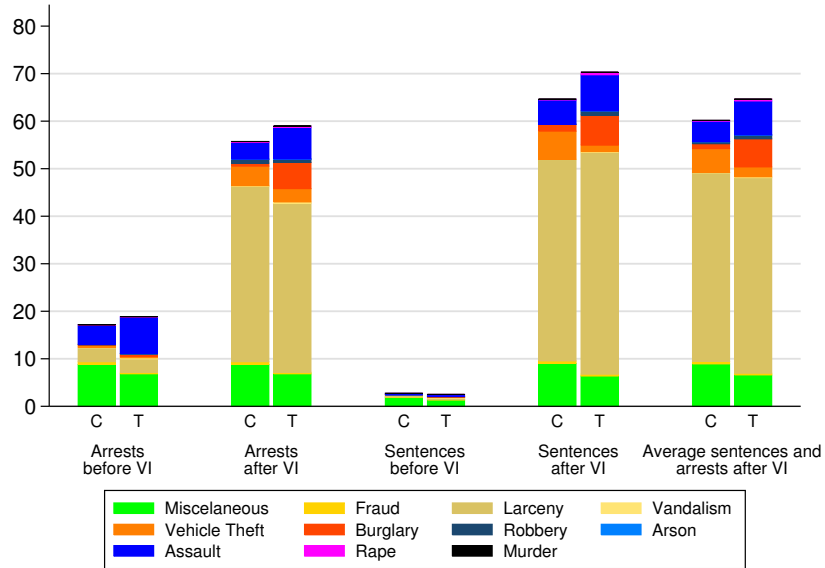


Note: This figure shows, by year and type of crime, the number of arrests (estimated from the National Arrests Analysis Tool from the NIBRS) divided by the number of sentences (estimated from the National Justice Reporting Program). In practice, we use a single number for each type of crime, which is an average across years.

### K.3.4 Effects on Number of Crimes, After Victimization Inflation

Figure K.5 shows the effects of VI on our estimates of the number of crimes committed. Note that the magnitudes in the axis are much larger than those of previous charts. The largest effects are for larceny, which is common in the data and has a victims-to-arrests factor of 12.6, the largest factor of all the categories of crime used in the paper. Given that the victim cost of larcenies is low, it affects the estimates less than what this chart suggests.

Figure K.5: Effects on Number of Crimes, After Victimization Inflation



Note: This chart continues Figure K.2. It shows, for the control (C) and treatment (T) groups, the effects of adding victimization inflation (VI). The first pair of columns is the same as the second pair of columns in Figure K.2. The second pair of columns expand the arrests to account for VI. The third pair of columns is the same as the fourth pair of columns in Figure K.2. The fourth pair of columns expand the sentences to account for VI. The last pair of columns averages the second and the fourth pairs of columns in this chart.

While the assumptions required for the victimization inflation methodology are strong, we argue that this is the best approximation for a total toll of crime's costs. The highest victims-to-arrests ratio shown in the figures are sensible and are not for the most costly categories of crime in the data, which stabilizes the estimates.

## K.4 Literature on Costs of Specific Crimes

There are many methods to estimate unit costs of representative crimes, and many studies presenting estimates.<sup>83</sup> In this document, we only review the literature related to the inputs necessary for this paper.

<sup>83</sup>Cohen and Bowles (2010) and McCollister et al. (2010) give comprehensive reviews of the state of the literature.

We start by classifying the costs of crime, which is necessary to later discuss the methods to estimate the costs. Then, we present the two general types of methodologies that are used to estimate the total costs of crime: the Top-down methodologies and the Bottom-up methodologies. The former attempts to quantify the value that people put into *ex-ante* prevention of crime, while the latter attempts to gather *ex-post* all sources of costs that crime generates. The difference between these two methods can be large: [Cohen and Bowles \(2010\)](#) show that for the particular case of estimates of the cost of rape, the top-down approach gives a value that is twice as large as the value given by most bottom-up studies. Other studies give cost estimates that are more homogeneous between these two approaches.

#### **K.4.1 Classifying the Costs of Crime**

Some methodologies used to estimate costs of crime are only able to capture some types of costs, and it might not even be clear what other methodologies are capturing. Some important types of costs are:

- Costs to the victim that can be directly quantified, such as medical bills, property losses, and lost productivity.
- Costs to the victim that cannot be observed, such as pain and suffering.
- Costs to the community in terms of prevention of crime, such as alarms, avoidance behavior, and police presence.
- Costs to the community in terms of fear.
- Costs to the community in terms of the criminal justice system, especially imprisonment.

- Costs to the offender in terms of lowered productivity, such as forgone wages.

#### K.4.2 Bottom-up (BU) Methodologies

These approaches sum each type of cost that is imposed after the crime has been committed. The most well-known studies combine direct (also known as tangible) costs of the crimes with intangible costs. Tangible costs are everything that can be directly measured by observation, such as foregone wages, hospital costs, and police expenditure. Intangible costs are subjective, like pain and suffering. One way to measure these costs is using jury awards. For example, a jury award given as a result of an arm broken at a construction site can be used as a proxy of the intangible cost of having an arm broken in an assault.<sup>84</sup> The problem of these approaches is that many of the costs of crime are not directly imposed on the victim and are hard to quantify, such as the “fear of crime,” the increased expenditure on crime prevention, and the negative impact of imprisonment on the community.

#### K.4.3 Top-down (TD) Methodologies

The other way to estimate the cost of crime is using TD methods, based on eliciting willingness to pay for avoiding crimes. The main advantage of these methods is that, in principle, they consider costs that are hard or impossible to measure directly, such as the cost of fear, avoidance behavior, and expenditures in preventative measures. There are three main methodologies for this approach, which we now briefly describe.

1. Stated Preferences. This basic method elicits the willingness to pay for hypothetical programs that would reduce crime nationwide for a sample of people.<sup>85</sup> Being an example of a TD methodology, it is expected that the costs obtained by this method

---

<sup>84</sup>This was first used in [Cohen \(1988\)](#), and has been extensively used in BU studies after that. [Miller et al. \(1996\)](#) improved on previous estimates by using jury awards specifically coming from criminal cases.

<sup>85</sup>[Cohen et al. \(2004\)](#).

would include factors that affect the community, and that are hard to capture, such as fear. However, it is unclear whether people consider factors like the cost of the justice system in their answers to these questions. An obvious caveat of this method is that people might not answer the real amount they would be willing to pay in these surveys.

2. **Revealed Preferences.** This method infers the value that individuals assign to crime reductions from market transactions. The most standard way to calculate these estimations is running regressions to explain the total price of houses with several factors, including the rates of crime in the area. Those parameters associated with the crime rate are considered the revealed valuation of avoiding crimes.<sup>86</sup> One weakness of this method is that it assumes that people are well-informed on the crime rates in an area. Another problem is that, in absence of extremely large and rich data on crimes and housing prices, it is not possible to separately identify the costs of different types of crimes. To the best of our knowledge, no paper has yet been able to convincingly obtain estimates per type of crime with this method.
3. **Life Satisfaction.** For this method, people are surveyed about their preferences between different life conditions, in which several different factors are considered. Some of those factors are income and rates of crime. By doing so, people implicitly associate monetary values to the levels of crime in the communities they would live in.<sup>87</sup>

#### **K.4.4 Costs Used in this Study**

To summarize, both approaches have strengths and weaknesses: the TD approaches are more likely to reflect costs to the community (e.g. fear and anxiety, avoidance behavior, and protective measures) and better capture the spirit of a prevention program. However, in practice TD estimates rely on strong assumptions, and there are methodological issues associated with obtaining detailed values for the different types of crimes. It is also possible

---

<sup>86</sup>Thaler (1978).

<sup>87</sup>Moore and Shepard (2006); Moore (2006).

that when people answer the survey used for TD calculations they include some costs that we are including separately, such as justice system costs, and risk of death from non-murder crimes, while BU does not include them. Given those considerations, and the lack of TD costs for some categories of crime, we use BU costs for our main estimates. For completeness, we present cost estimates using both approaches. We choose [Cohen et al. \(2004\)](#) as representative of the TD approaches, and [McCollister et al. \(2010\)](#) as representative of the BU approaches. In terms of timing, both of these studies match well with the ABC and CARE data. The bulk of crimes in the ABC and CARE data occurred between the late 1990s and early 2000s. While [Cohen et al. \(2004\)](#) do not report the exact year of their survey, they use Census 2000 figures for their estimates. Even though [McCollister et al. \(2010\)](#) is a more recent study, many of the productivity estimates that their costs are based on are taken from papers using data from years with more crimes the late 1990s and early 2000s. The costs in those studies are presented in [Table K.8](#). Notice that there are some strong differences in the cost of crimes, such as assault, burglary, and especially robbery.

Table K.8: Unitary Costs of Crime for Victims

Crime	Top-Down Approach Cohen et al. (2004)	Bottom-Up Approach McCollister et al. (2010)
Arson		12,093
Assault	95,200	16,132
Burglary	34,000	1,467
Fraud		0
Larceny		528
Motor Vehicle Theft		6,699
Murder	13,192,000	9,286,200
Rape	322,320	224,021
Robbery	315,520	7,273
Vandalism		0

Note: All amounts are in 2014 USD. The amounts reported in [McCollister et al. \(2010\)](#) for non-murder crimes have the extra cost for risk of death and the cost of a crime career removed (both were obtained from correspondence with the author). Risk-of-death costs do not apply, because we know the outcomes of the crimes. Crime-career costs do not apply, as we directly observe the income of the individuals. These costs also don't include police and legal system costs, as those are imputed separately and only for the cases for which individuals were arrested or sentenced.

#### K.4.5 Timing of Effects: Incidence vs. Prevalence

We would like to discount the costs of crime according to whether they were *incurred* during a particular age of ABC and CARE subjects, because those values should be discounted at a different rate than costs incurred later, even if both costs were *imposed* in the same year. Thus, the value of the imprisonment is discounted year-by-year. We have no information about the timing of costs for victims, so the value of the different crimes for the victims are discounted according to the time they were imposed (the time of the crime's occurrence).

#### K.4.6 Costs of Imprisonment

Unlike previous studies, we observe the sentences of the ABC and CARE subjects. This allows for a precise estimation of the costs of imprisonment. For the cost of jail and state prison, we use estimates from the [U.S. Department of Justice \(1988\)](#). In 2014 USD, these

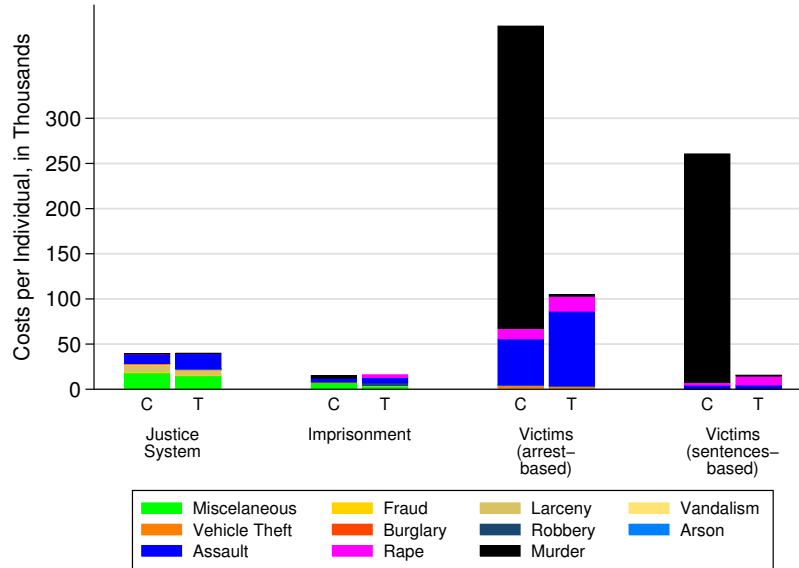
costs are \$25,338 for a year in a state prison, and \$21,939 for a year in jail. It is important to clarify that we only include costs of the justice system for the crimes that are known by the justice system, not for the crimes that we impute through victimization inflation.

## **K.5 Effects on Costs of Crime**

### **K.5.1 Effects on Costs Before Victimization Inflation**

Figure [K.6](#) presents the estimated costs per type of cost before victimization inflation. There are clear positive effects for the treatment group in terms of reductions in the costs of crime. Those reductions are almost exclusively given by the large effect of the murder case we observe in the control group (note that murder also appears in the treatment group costs because of the predictions). Comparing the bars in this figure, the costs from the justice system and from imprisonment are low compared with the victimization costs, even without victimization inflation. While the levels of the arrest-based estimates are higher than the levels of the sentence-based estimates for both the treatment and control groups, the impacts of the program are quite similar across both methods (Figure [K.6](#)).

Figure K.6: Costs of Crime Before Victimization Inflation

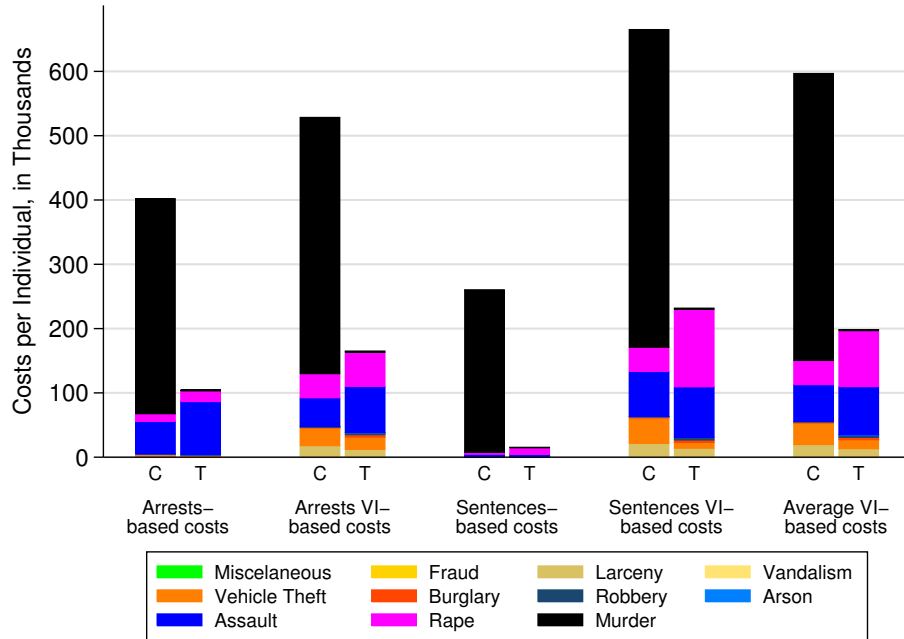


Note: This figure depicts the per capita cost for the different categories of costs and crimes we use, by control (C) and treatment (T). The first pair of columns adds up the justice system costs (including police) for all arrests inputted for each subject. The second pair of columns adds up the cost of Imprisonment. It is important to note that the costs are per capita, so there are individual cases that have much higher costs. The next two pairs of columns show the pre-victimization inflation estimates of number of crimes multiplied by the individual victim cost of the different crimes. The costs are taken from the Bottom-up approach in Table K.8. All costs are in thousands of 2014 USD.

### K.5.2 Effects on Costs After Victimization Inflation

Figure K.7 presents the data after applying the victimization inflation. As shown below, the inflation allows us to include a substantial amount of crime that otherwise would not have been considered. This chart shows that the treatment effects using arrest-based estimations are not substantively different from the ones using sentence-based estimations. Thus, to use all available information, we use the estimates based on the averages of the two for our analysis.

Figure K.7: Costs of Crime After Victimization Inflation



Note: This figure depicts the per capita cost for the different categories of costs and crimes we use, by control (C) and treatment (T). The first two pairs of columns adds up all of the arrest-based costs for each subject and compares the pre- and post-victimization inflation costs. The third and fourth pair of columns compare the pre- and post-victimization inflation sentence-based costs. It is important to note that the costs are per capita, so there are individual cases that have much higher costs. The next two pairs of columns show the post-victimization inflation estimates of number of crimes multiplied by the individual victim cost of the different crimes. The costs are taken from the Bottom-up approach in Table K.8. All costs are in thousands of 2014 USD.

We consider the impact on murder as a consequence of the program rather than a statistical coincidence. We use as precedent the cost-benefit analysis of Perry in which three control group individuals and one treatment group individual committed murders.<sup>88</sup>

Some of the sources of cost estimates, such as the more serious crimes, result in volatile estimates due to the small sample sizes. Our estimations of the standard errors associated with the objects of interest in this paper—the present value of the program and the internal rate of return—consider those sources of volatility. Ultimately, the benefit-cost analysis is a

<sup>88</sup>Heckman et al. (2010).

unidimensional summary of benefits of a program, and specific flows of benefits with high variability enter naturally into the process of aggregation.

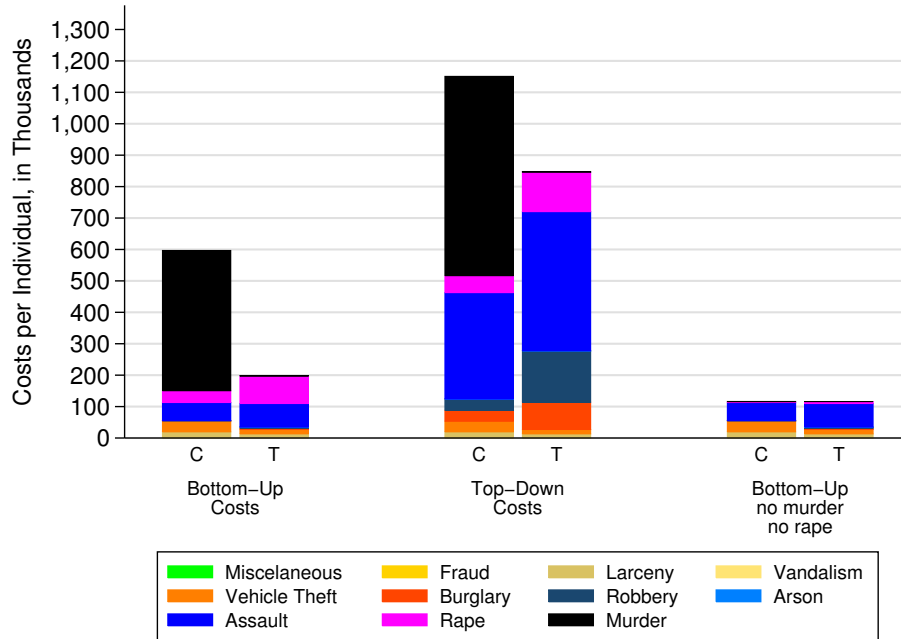
## **K.6 Sensitivity Analyses Using Alternative Cost Estimates**

So far, we study several ways to construct our estimates:

1. We show that not matching the different crime datasets could reduce the number of crimes, but that the general patterns are stable, and no dataset is especially influential.
2. We show that not including predictions up to age 50 noticeably reduces the number of crimes, but the general patterns are not modified.
3. We show estimates using arrest-based estimations versus sentence-based estimations, and find that the differences are large in terms of the number of crimes before victimization inflation, but small after it, and do not substantially change the total benefits calculations.

In Figure [K.8](#), we present additional deviations from our main estimates. In particular, we show how the estimations change when two different cost schedules are used: (i) Top-down costs, and (ii) Bottom-up costs, assuming that the costs of murders and rapes are identical to the cost of an assault. We also note that BU costs are a “conservative” option in the sense that the effects of the program are higher using TD costs. We can also see that with no murders and rapes, the effect of the program on crime is still positive, but much smaller than when those crimes are considered.

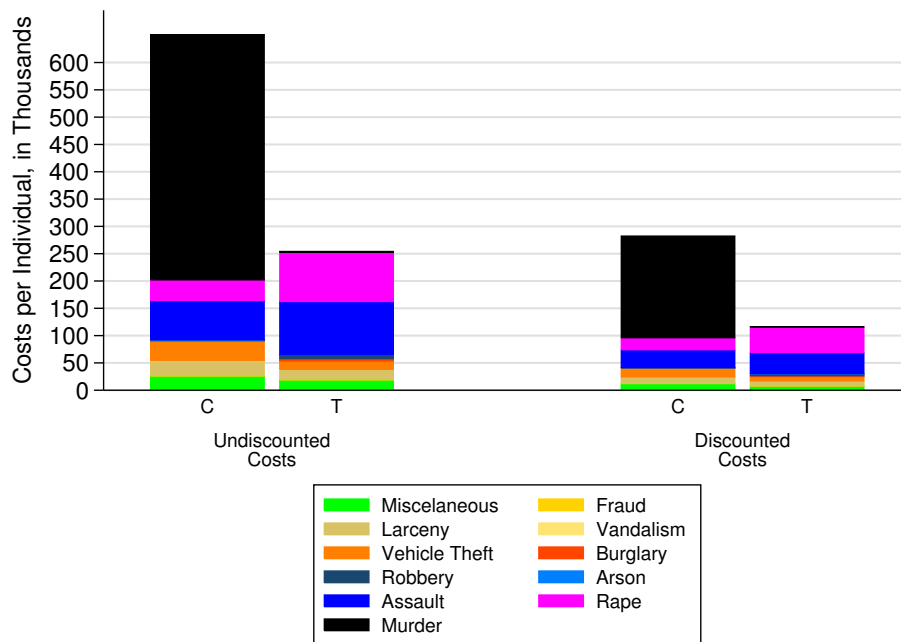
Figure K.8: Different Cost Schedules



Note: This figure depicts the per capita cost for the different categories of costs and crimes we use, by control (C) and treatment (T). It presents some sensitivity analyses. The first pair of bars represents costs using the Bottom-up approach in Table K.8 to determine individual costs of crime. The second pair of bars represents costs using the Top-down approach in Table K.8 to determine individual costs of crime. The third pair of bars uses the Bottom-up approach, but replaces the values of murders and rapes with that of assaults.

Finally, in Figure K.9, we present the effect that discounting and applying deadweight loss has on our estimates. The first pair of bars represents the cost estimates that are not discounted, the second pair of bars represents the discounted costs, and the third pair of bars represents the discounted costs including deadweight losses. It is clear that the effect of discounting is substantial, approximately halving the total cost estimates. The effect of applying deadweight loss barely changes the estimates. The effect of deadweight loss is most noticeable in the value of miscellaneous crimes, which are only public costs. Regardless, its impact in the overall costs is minor.

Figure K.9: Effect of Discounting Crime Costs



Note: This figure depicts the per capita cost for the different categories of costs and crimes we use, by control (C) and treatment (T). The discounted costs use 3% as a discount factor, and are discounted to birth. The deadweight loss (DWL) adjustments increases the costs from justice system and imprisonment by 50%. All costs are in thousands of 2014 USD.

## L Interpolation of Parental Income

We observe labor income of the mother and, if applicable, her partner at ages 0, 2, 3, 4, 5, 9, 12, and 15. To obtain our measure of parental income, we simply sum the labor income of the respondent and her partner.

## M Extrapolation of Individual Income

### M.1 Data Sources

Income of the ABC and CARE subjects is observed in detail at ages 21 and 30. For both ages, we separate labor income from public-transfer income. At age 21, public-transfer income includes Aid to Families with Dependent Children (AFDC) subsidies, food stamps, survivor benefits, disability benefits, social security, rent subsidies, and fuel subsidies. At age 30, public-transfer income includes food stamps, welfare, housing assistance, workman's compensation, disability, social security, supplemental security income, unemployment benefits, worker's compensation insurance, fuel subsidies, educational and aid grants, and other forms of welfare. At both ages 21 and 30, labor income is observed. Also at both ages, income from other family members are excluded, and all values are in to 2014 USD. We utilize the NLSY, CNLY, and PSID datasets. We provide a brief overview of them below.

#### M.1.1 NLSY

The National Longitudinal Survey of Youth (NLSY) is a longitudinal survey beginning in 1979 that follows individuals born between 1957 and 1964. The initial interview included 12,686 respondents aged 14 to 22. The survey was designed to include 6,111 individuals representing the non-institutionalized civilian population, a supplemental sample of 5,295 civilian Hispanics, Latinos, blacks, non-blacks/non-Hispanics, and economically disadvantaged youth, and a sample of 1,280 who served in the military as of September 30, 1978. When appropriately weighted, the NLSY is nationally representative of the youth living in the U.S. on January 1, 1979. We include individuals from all three subsamples in our analysis.

The NLSY collected data on labor market participation, education, family background, fam-

ily life, health issues, assets and income, government program participation, and measures of cognitive skills. We use these data to estimate a prediction model for subject income for ages 31 through 67.

We restrict the NLSY sample to individuals with labor income less than \$300,000 (2014 USD) at any given year. With the mean labor income (2014 USD) in the ABC sample being \$12,232 at age 21 and \$32,782 at age 30, and the maximum reported being \$189,938, the cut-off we impose on the auxiliary data is high enough so that everyone in the ABC and CARE samples are represented, yet low enough to exclude high-earning individuals in the auxiliary sample that do not reflect the ABC or CARE samples well.

We do not impose a restriction on birth year on the NLSY as all respondents are aged between 47 and 55 at the time of the last interview (conducted in 2012). This age range is within the 31–67 range for which we extrapolate the income of the ABC and CARE subjects.

Given the biennial nature of the NLSY, we only observe each subject at either odd or even ages. Not only does this reduce the size of the sample on which we fit our prediction model at each age, but it can introduce biases associated with the odd-aged and even-aged cohorts. To address this issue, we perform a linear interpolation on the variables in the NLSY data that enter into our prediction model. This allows us to estimate our prediction model on all subjects of the NLSY satisfying our restrictions at every age.

### **M.1.2 CNLSY**

The Children of the National Longitudinal Survey of Youth (CNLSY) is a survey of the children of the mothers from the NLSY, beginning in 1986. At the time of the initial interview, the ages of the children surveyed ranged from 0 to 23. As of 2010, the CNLSY sample

includes 11,504 children born to NLSY mothers. With appropriate weights, the CNLSY may be considered nationally representative of children born to women who were aged 14 to 22 during 1979. Interviews were conducted annually between 1986 and 1994, and biennially thereafter.

Similar to the NLSY, the CNLSY collected data on cognitive ability, motor and social development, home environment, health information, education, attitudes, employment, income, family decisions, and more. We use these data to estimate a prediction model for subject income for ages 22 through 29.

As we did with the NLSY, we restrict the CNLSY sample to individuals with labor income less than \$300,000 (2014 USD) at any given year. In addition to this, we limit the sample to subjects born between 1978 and 1983. Because the CNLSY data extend to 2012, this implies that we use the most recent data from the CNLSY in which individuals are aged 29 to 34. Finally, given the biennial nature of the survey, we perform a linear interpolation on the variables that enter into our prediction model. This allows us to use as much of the CNLSY data as possible at every age when interpolating subject income.

### **M.1.3 PSID**

The Panel Survey of Income Dynamics (PSID) is a longitudinal household survey containing between 5,000 and 8,500 families in each wave. It began as yearly survey in 1968 and has been fielded biennially since 1996. When appropriately weighted, the PSID is designed to be representative of U.S. households. The PSID provides extensive information concerning demographics, economic outcomes, health outcomes, marriage and fertility, and more. In addition to using the PSID to forecast future earnings of ABC and CARE subjects, we use the PSID to forecast health outcomes. See Appendix [N.2.1](#) for details on how the PSID

relates to health outcomes.

Similar to the CNLSY, we restrict the PSID to individuals born between 1945 and 1981. Because the data extend to 2013, this means we use the most recent subsample of individuals aged 30 to 67. We also exclude all individuals with labor income exceeding \$300,000 (2014 USD) at any given year. Finally, given the biennial nature of the survey, we perform a linear interpolation on the variables that enter into our prediction model. This allows us to use as much of the PSID data as possible at every age to interpolate subject income.

## N Health Outcomes in the Future America Model (FAM)

### N.1 Background and Description of FAM

In this appendix, we explain our methodology to measure the projected differences in health outcomes and medical expenditure over the adult life for the treated and control groups in ABC and CARE. Health outcomes and behaviors of ABC and CARE subjects are measured at the age-30 interview and in a health follow-up conducted when the subjects were in their mid-30s. To project the life-cycle path of health outcomes and medical expenditures, we use a dynamic microsimulation model to track the treatment and control cohorts from age 30 until death.<sup>89</sup>

---

<sup>89</sup>This microsimulation model is an extension of the model used by Prados et al. (2015); the technical details are described in Goldman et al. (2015). Both models are related to the Future Elderly Model (FEM), which is a microsimulation tool originally developed to examine the health and health care costs among the elderly Medicare population (Goldman et al., 2004). It has been used extensively to assess health and disease prevention scenarios: FEM has been used to assess the future costs of disease, the benefits of preventing disease among older population, the consequences of new medical technologies, trends in disability, and the fiscal consequences of worsening population health (see Goldman et al. (2004), Lakdawalla et al. (2004), Goldman et al. (2005), and Zissimopoulos et al. (2014)). The main differences with FEM are that the model we use starts cohorts of individuals at age 30 instead of 50, and that it simulates more outcomes than FEM, because they are important to explain health outcomes and medical expenditure at younger ages, like evolution of partnership and marital status, work status, and family size.

The defining characteristic of this approach is the modeling of real rather than synthetic cohorts, all of whom are followed at the individual level. This allows for more heterogeneity in behavior than would be allowed by a cell-based approach. The core of the model can be described as follows: first, the cohort starter module contains the health outcomes of the ABC or CARE subjects at age 30 and other variables that are used as input for the simulation of individual trajectories. This module uses ABC or CARE data. Missing variables in the ABC and CARE data are imputed probabilistically from models estimated using PSID data. Next, the transition module calculates the individual probabilities of transiting across various health states and other outcomes relevant to health. The transition probabilities are estimated from the longitudinal data in the PSID, taking as inputs risk factors such as smoking, weight, alcohol consumption, gender, race and ethnicity, age and education, along with lagged health, personal, and economic states.<sup>90</sup> This scheme allows for a great deal of heterogeneity and fairly general feedback effects. Finally, the policy outcomes module aggregates projections of individual-level outcomes into policy outcomes such as taxes, medical expenditures, and disability benefits.

The cohort starter module includes the following variables for each ABC and CARE subject:

- Individual characteristics: year of birth, gender, treatment status, education of the mother, self-reported “poor” as a child, race, and education at age 30.
- Economic outcomes at age 30: working status, earnings, and health insurance status.
- Health outcomes and health behaviors at age 30: body mass index (BMI), smoking, binge drinking, physical activity, psychological distress, asthma, high blood pressure, heart disease, cancer, lung disease, diabetes, and stroke.

Health conditions and health behaviors are derived from survey questions about doctor-

---

<sup>90</sup>Section N.2 provides details about the data sources used in the estimation.

diagnosed conditions and self-reported health behaviors. For details about how we deal with missing variables and assumptions, see Section [N.3.1](#).

The core of the microsimulation is a set of models of disease conditions designed to predict future health and functional status of each individual from his or her current health state at age 30. To predict health and economic outcomes over time, the model calculates the transition probabilities between various health states and other outcomes. Health states include diabetes, heart disease, stroke, cancer, hypertension, lung disease, and number of difficulties in physical and instrumental activities of daily living (ADLs and IADLs); health risks include BMI (defined as weight in kilograms divided by height in meters squared, which is used to measure incidence of obesity), binge drinking (defined as binge drinking at least three times per month<sup>91</sup>), smoking behavior and (lack of) physical activity; other outcomes include health insurance status, changes in family structure (partnership or marriage, child-bearing), labor market participation, working status, receipt of Social Security, participation in public programs, and medical expenditures. We also estimate quality-adjusted life years (QALYs), a measure of the quality of life that adjusts for the burden of disease.<sup>92</sup>

The likelihood of developing a health condition depends on key risk factors including age, gender, education, race and ethnicity, obesity (BMI greater than 30 kg/m<sup>2</sup>), smoking status, physical activity, age of asthma diagnosis, and lagged health outcomes. By incorporating lagged variables into the model, we account for the likelihood that past behaviors may influence risks far into the future. This capacity is important because prior research indicates that past health behaviors, such as recency of prior smoking and 10-year history of obesity, can influence current health outcomes.<sup>93</sup> Furthermore, because transition probabilities vary

---

<sup>91</sup>Where binge drinking behavior is defined as drinking more than five alcoholic drinks in an instance for males and more than four for females.

<sup>92</sup>A QALY equals one year of life in the absence of disease. This measure has been widely used in the literature to evaluate the value of medical interventions and improvements.

<sup>93</sup>[Tong et al. \(1996\)](#); [Moore et al. \(2008\)](#).

depending on demographic characteristics, such as race and ethnicity, education, gender, and age, the model tracks outcomes by socioeconomic subgroups, and it allows for responses to policies to be subgroup-specific.

Health conditions are treated as absorbing states, i.e., once a person has a disease she is assumed to have it forever. A person can transition out of an obese state and back into it. A person can quit smoking and resume smoking. We model transitions of all health conditions and risk factors with a first-order Markov process. A number of restrictions are placed on the way a disease or condition is associated with other conditions (see Section N.3.3 for details). Like the previous literature that uses FEM and FAM, we assume that risk factors and health conditions only from the prior period determine future health, disability, and mortality (a Markov process). This Markov assumption may be a potential limitation to the accuracy of disease dynamics, but given the available follow-up data it would be difficult to implement a higher order process. Moreover, to evaluate the performance of the estimated model, we validate it using various techniques, including comparing model results from early years with actual data available for later years.<sup>94</sup>

Family formation models estimate transition probabilities between the following relationship statuses: single, cohabiting, married, separated/divorced, and widowed. We use multivariate regression models to estimate the number of children born separately for women and men. Economic models are developed to estimate labor force participation and employment status (possible states allowed by the model are: unemployed, out of the labor force, working part-time, or working full-time), and the take-up of government social insurance programs such as disability insurance. Transitions of labor earnings are projected outside of the simulation (Section M describes the methodology). Because there is no information about assets in the age 30 ABC or CARE data, we do not simulate wealth transitions.<sup>95</sup>

---

<sup>94</sup>Goldman et al. (2015).

<sup>95</sup>Additional details of the transition models are provided in Goldman et al. (2015).

We simulate outcomes for cohorts that have the initial characteristics of the ABC or CARE treatment and control groups at age 30. In each year, we use the health, family, and economic transition models to predict obesity, smoking behavior, health status, economic status, family characteristics, and disability. We then use the models of health care spending to calculate medical costs for Medicare, other public sources excluding Medicare, and medical costs for private sources. We repeat the simulation each year until everyone in the cohort would have died.

## **N.2 Data Sources**

FAM uses data from ABC and CARE follow-up surveys to build the initial state of the cohort. The transition model parameters are estimated from the 1997 to 2013 waves of the Panel Study of Income Dynamics (PSID). We supplement the PSID with data from the Health and Retirement Study (HRS). We use the National Health and Nutrition Examination Survey (NHANES) to account for differences between measured and self-reported BMI. To estimate medical care costs associated with health conditions, we use the Medical Expenditures Panel Survey (MEPS) and the Medicare Current Beneficiaries Survey (MCBS).

### **N.2.1 PSID**

The Panel Study of Income Dynamics (PSID) provides extensive information concerning demographics, economic outcomes, health care access, health outcomes, and health behaviors (such as smoking history, alcohol consumption, and exercise habits). Health outcome variables include diagnosis of diabetes, heart disease, hypertension, lung disease, and cancer, among others. See Section [M.1.3](#) for additional details on the PSID.

We estimate the transition models using waves from 1997 to 2013. We create a dataset of respondents who have formed their own households, either as single heads of households, cohabiting partners, or married partners. These heads, wives, and husbands respond to the richest set of PSID questions, including the health questions that are critical for our purposes. We use all respondents aged 25 and older.<sup>96</sup> The length of the PSID is a significant advantage, because we can include past health behaviors as explanatory variables for current health outcomes. This dataset provides adequate sample sizes to explore health outcomes of specific groups. PSID does not follow individuals who are institutionalized in nursing homes or other long-term care facilities. To overcome this weakness, we pool the PSID sample with the HRS sample when estimating mortality models.

## **N.2.2 HRS**

The Health and Retirement Study (HRS) is a longitudinal panel that surveys a nationally representative sample of individuals over the age of 50 and their spouses every two years. When appropriately weighted, the HRS in 2010 is representative of U.S. households where at least one member is at least 51 years old. This study collects in-depth information about income, work, health, and medical expenditures. In our model, waves from 1998 to 2012 are pooled with the PSID for estimation of mortality and widowhood models. The HRS data are harmonized to the PSID for all relevant variables. Because PSID does not follow respondents into nursing homes, we also use the HRS to estimate the model for nursing home residency. We use all cohorts in the dataset created by RAND (RAND HRS, version O) as the basis for our analysis.

---

<sup>96</sup>While we use the full sample, we explored using a few different subsamples to better adapt to the demographics of the ABC and CARE subjects.

### **N.2.3 MCBS**

The Medicare Current Beneficiary Survey (MCBS) is a nationally representative sample of aged, disabled, and institutionalized Medicare beneficiaries. The MCBS attempts to interview each respondent twelve times over three years, regardless of whether he or she resides in the community, a facility, or transitions between community and facility settings. The disabled (under 65 years of age) and very elderly (85 years of age or older) are over-sampled. The first round of interviewing was conducted in 1991. Originally, the survey was a longitudinal sample with periodic supplements and indefinite periods of participation. In 1994, the MCBS switched to a rotating panel design with limited periods of participation. Each fall, a new panel is introduced, with a target sample size of 12,000 respondents. Each summer, a panel is retired. Institutionalized respondents are interviewed by proxy. The MCBS contains comprehensive self-reported information on the health status, health care use and expenditures, health insurance coverage, and socioeconomic and demographic characteristics of the entire spectrum of Medicare beneficiaries. Medicare claims data for beneficiaries enrolled in fee-for-service plans are also used to provide more accurate information on health care use and expenditures. MCBS data from 2007 to 2010 are used for estimating medical costs and enrollment models.

### **N.2.4 MEPS**

The Medical Expenditure Panel Survey (MEPS), which began in 1996, is a set of large-scale surveys of families and individuals, their medical providers, and employers across the U.S. The Household Component (HC) of the MEPS provides data from individual households and their members, which is supplemented by data from their medical providers. The HC collects data from a representative subsample of households drawn from the previous year's National Health Interview Survey (NHIS). Since NHIS does not include the institutionalized population, neither does MEPS; this implies that we can only use the MEPS to estimate

medical costs for the non-elderly (ages 25–64) population. Information collected during household interviews include: demographic characteristics, health conditions, health status, use of medical services, sources of medical payments, and body weight and height. Each year the household survey includes approximately 12,000 households, or 34,000 individuals. Sample size for those aged 25-64 is about 15,800 in each year. MEPS has comparable measures of socioeconomic status as those in PSID, including age, race and ethnicity, educational attainment, census region, and marital status. We estimate medical expenditure and utilization using data from 2008 to 2010. We use waves from 2001 to 2003 to estimate models of quality-adjusted life years (QALYs), due to availability of EQ-5D instrument in these waves.<sup>97</sup>

## N.2.5 NHANES

The National Health and Nutrition Examination Survey (NHANES) targets a nationally representative sample of approximately 5,000 individuals in each year since 1999. The data collected includes responses to interview questions about demographics, disease conditions, height, and weight, as well as physical measurement of BMI. We use NHANES years 2002 to 2010 to estimate a model for imputing measured BMI from self-reported BMI. The methodology is described in Section [N.3.1](#).

## N.2.6 ABC and CARE

FAM uses ABC and CARE data to initialize the state of each ABC or CARE subject when they enter into the simulation. These data are taken from the the age-21 parent interview; age-30 subject interview; age-30 tobacco, alcohol, and drugs interview; and age-34 health interview. The goal is to have each subject’s initial state match their status at the age-30 subject interview. Because several key FAM inputs are not available at the age-30 interview,

---

<sup>97</sup>Section [N.3.2](#) explains the estimation of the QALY model.

we use PSID or ABC and CARE surveys corresponding to other ages to impute missing elements. These imputations are discussed in Section [N.3.1](#).

## **N.3 Methods and Analysis**

### **N.3.1 ABC and CARE Data Assumptions and Imputations**

Marital status transitions and childbearing in FAM are affected by the subject’s mother’s education level. The ABC and CARE age-30 subject interview did not ask about mother’s education, but the ABC age-21 parent interview did. For ABC subjects, we assume that each subject’s mother had the same education level at the age-30 subject interview as what was reported in the age-21 parent interview. For CARE subjects, we impute mother’s education from an ordered Probit model using race, ethnicity, education, disease conditions, employment status, presence of a health-related work limitation, and a self-report of whether or not the subject was “poor” as a child. The model is estimated using age 30 to 31 PSID subjects. Each of the model covariate values are taken from the CARE age 30 interview. At the beginning of each simulation repetition, an education level is randomly drawn from the probability distribution for each CARE subject and assigned to be the mother’s education level.

Many FAM transition models depend on a three-level measure of parents’ economic status when the subject was a child. This is based on the PSID question: “Were your parents poor when you were growing up, pretty well off, or what?” The three possible responses are “poor,” “average”/“it varied”, or “pretty well off.” This question is not included in the ABC or CARE interviews, but because preliminary eligibility for the program focused on children from high-risk backgrounds, based on socioeconomic factors, the value of this variable is set to “poor” (when growing up) for all ABC and CARE subjects.

All FAM transition models depend on demographics of the subject, including whether or not the subject is Hispanic. This information is not available in the ABC or CARE data, but it is assumed that none of the ABC or CARE subjects are Hispanic.<sup>98</sup>

Most FAM models depend on smoking status. Employment status affects FAM transitions in marital status, childbearing, claiming of disability insurance (DI) and supplemental security income (SSI), and type of health insurance. One male in the ABC control group was missing smoking status and, although known to be not working, was also missing specific employment status (unemployed or out of the labor force). We use a multinomial logit model to jointly estimate the probability of each combined smoking and employment category among 25- to 35-year-olds in the PSID who were not working. At the beginning of each simulation repetition, we use a Monte Carlo random draw generated from this distribution to assign this subject's smoking and employment statuses. This same subject is also missing information about binge drinking. A separate binary Probit binge drinking model was estimated using the age 25–35 PSID data. A Monte Carlo random draw is taken according the Probit probability to predict binge drinking behavior at the beginning of the simulation.

BMI affects FAM transitions in health, functional status, employment, and smoking. The FAM transition models are estimated with BMI computed from self-reported height and weight in the PSID. The only BMI data in ABC and CARE come from height and weight measured during the health interview. This interview took place at roughly age 30 for CARE subjects, and at age 34 for ABC subjects. This poses two challenges. First, self-reported BMI can be biased by factors such as actual height and weight, gender, and race.<sup>99</sup> Second, it is possible that BMI could increase or decrease systematically in the years between the

---

<sup>98</sup>Census data on Hispanics in North Carolina were not available for 1970 and 1980, but Hispanic migration into this state is more recent than in other regions, and as late as 1990, only 2% of the North Carolina poor were Hispanic (Johnson, 2003).

<sup>99</sup>Cawley (2004).

age-30 subject interview and the age-34 health interview.

To address the first BMI imputation challenge, we used a variation on the method of [Courtemanche, Pinkston, and Stewart \(2015\)](#) to impute measured BMI in the PSID. While the method in [Courtemanche et al. \(2015\)](#) works with height and weight, we apply the specification to directly model BMI. Using respondents aged 30 to 40 in the 2002-2010 NHANES waves, we predict measured BMI from percentile ranks of self-reported BMI using the model specification in [Courtemanche et al. \(2015\)](#). Three variations on the spline interactions of [Courtemanche et al. \(2015\)](#) were also considered. After estimating this model using NHANES data, covariate values from the PSID age 30–34 data in years 2002–2013 were used to impute measured BMI values for PSID respondents. A Kolmogorov-Sminov (K-S) test and a visual inspection of smoothed histograms were used to compare the distribution of PSID imputed values to the distribution of observed values in the NHANES estimation sample. The model specification selected for imputation had the smallest K-S distance between the two distributions. The smoothed histogram of the distributions for the entire samples and the black subgroup in each data set appeared reasonably close.

After imputing values of measured BMI for PSID respondents age 30–34, we turned to the second problem: accounting for systematic trends in BMI from the age 30 interview to the health interview. The goal was to have a model that could map from measured BMI at the health interview around age 34 to self-reported BMI at the age 30 interview. Employing the longitudinal structure of PSID, we matched each respondent’s first interview between age 30–32 with their imputed measured BMI at between ages 33–40. We then estimated a model using self-reported BMI between ages 30–32 as the response variable and imputed measured BMI at ages 33–40, the age when BMI is measured, along with other variables observed at age 30 as explanatory variables. This imputation model was applied to any ABC or CARE subject who has their health interview at least one year after their age 30 interview.

For ABC and CARE subjects who have their health interview within one year of the age 30 interview, we assumed that any systematic time trends in BMI are too small to have any practical significance. However, we still needed to convert the imputed measured BMI to a self-reported value for compatibility with other transition models estimated in PSID. This model was estimated on ages 30–32 in the PSID and uses covariates from the age 30 interview along with imputed measured BMI to predict self-reported BMI.

At the beginning of each simulation repetition, we chose the appropriate model to impute self-reported BMI for each ABC or CARE subject based on the time between their age 30 interview and their health interview. Their expected BMI was estimated from this model. A Monte Carlo Normal random draw was generated using the subject’s expected BMI and the estimated variance from the model. This Monte Carlo draw was then assigned to be the subject’s initial self-reported BMI in the simulation. Using BMI from the health interview limited the ABC and CARE subjects simulated in FAM to only those who had height and weight measured in the health interview.

Subjects’ health insurance coverage affects their medical costs. FAM uses three categories of health insurance: none, public only, and some private. Five ABC subjects and three CARE subjects were missing health insurance status. Three cases were logically imputed by assuming that subjects have no health insurance if they do not know their insurance status and either go to an emergency room or community health clinic or do not go anywhere when they need health care. In order to impute the insurance category for the remaining five cases, we use age 25–35 PSID data to estimate a Probit model for whether or not a subject had insurance. The predictors were gender, earnings, marital status, self-reported health, employment status, and whether or not the subject had any biological children. We use this model to compute the probability of having insurance at the start of the simulation (at the

age-30 interview). Then, we generate a Monte Carlo binary random variate according to this probability. If the outcome is positive, the subject is assigned to have some private insurance.

FAM uses six Activities of Daily Living (ADLs) about which there is data in PSID: walking, dressing, eating, bathing or showering, getting in and out of bed or a chair, and using the toilet, including getting to the toilet. FAM simulates the number of these ADLs in which the subject has difficulty. ADL difficulties predict FAM transitions in benefits claiming, mortality, employment status, insurance category, and nursing home residency. FAM also transitions the count of difficulties among six Instrumental Activities of Daily Living (IADLs) from PSID: preparing one's own meals; shopping for personal toilet items or medicines; managing one's own money, such as keeping track of expenses or paying bills; using the phone; doing heavy housework, like scrubbing floors or washing windows; and doing light housework, like doing dishes, straightening up, or light housecleaning. Both ADLs and IADLs are components of FAM's model for quality-adjusted life years (QALYs). The ABC and CARE age-30 subject interview does not ask about ADLs or IADLs, but it does ask if the subject has a physical or nervous condition that keeps them from working. PSID respondents are also asked this question. We create an imputation model for each of these two measures using an ordered Probit model estimated on PSID respondents aged 25 to 35. We use these models to compute the probabilities for each number of ADLs and IADLs. At the start of the the simulation, we generate Monte Carlo random draws according to these probabilities and use them to assign the corresponding counts.

When a subject claims DI benefits, it affects their FAM transitions in employment status, insurance category, and Medicare enrollment. DI claiming also affects medical costs. SSI claiming affects FAM transitions in employment status. Lastly, claiming Social Security retirement benefits affects FAM transitions in employment status and insurance category. The ABC age-30 subject interview has a single yes/no question about claiming which asks:

“Currently are you receiving income from workman’s compensation, disability, or Social Security benefits including Supplemental Security Income?” CARE asks a similar question. The PSID has separate questions for each benefit type. We use a multinomial logit model to estimate the joint probability of each combination of DI and SSI claiming. The estimation uses PSID respondents aged 25 to 35 who were claiming at least one of the following benefits: workman’s compensation, DI, or SSI. A Monte Carlo random draw generated from this distribution is used to assign each ABC and CARE subject’s DI- and SSI-claiming status at the start of the simulation. One ABC subject is missing data about whether or not they were claiming and was assumed to not be claiming any benefits.

As discussed in Section [N.3.2](#), FAM uses different models to estimate medical costs depending on whether or not a subject is Medicare-eligible. Subjects can enroll in Medicare before the age of 65 if they are claiming DI. The cost estimates for Medicare-eligible subjects depend on the subjects’ current disease status at the age-30 interview and their disease status two years prior to the interview. Unfortunately, ABC and CARE do not have disease data two years before the age-30 interview. It is assumed that all subjects did not have their disease conditions in the previous period. In other words, for any subjects who reported a disease condition in the age-30 interview, their costs in the first simulation time step is estimated as if it were their incident year of the disease. Section [N.3.2](#) describes the implications of this assumption.

### **N.3.2 FAM Models and Estimation**

We develop models to estimate the determinants of transitions between health outcomes, labor market outcomes, educational attainment, and family formation, for individuals aged 25 and older. Additionally, we estimate transition probabilities by gender, race and ethnicity, and educational attainment as a function of individual characteristics (see below). Each

transition model includes a subset of variables and relevant interactions from the following list: age, gender, race and ethnicity, education, parents' education, self-reported body mass index (BMI), smoking history, physical activity, binge drinking, lagged health conditions, asthma diagnosis before age 30, number of biological children, past earnings and work status, partnership status (single, cohabiting, married, separated/divorced, or widowed), disability status, and health insurance status. We consider three racial and ethnic groups (black non-Hispanic, white non-Hispanic, and Hispanic), and four educational groups (less than high school degree; high school graduate, including some college or associate's degree; college; and more than college).

The health transition models estimate the probability that a person transitions between health states, e.g. obesity or heart disease, as a function of current health status, demographic characteristics (including race, gender, age, and education), and risk factors (including weight, smoking status, physical activity, asthma, and number of births if female for BMI transitions), enabling us to age the cohorts. This mechanism to model health transitions accounts for the fact that certain health conditions increase the likelihood of comorbidities. We estimate a transition model for each of the following health conditions: heart disease, blood pressure, stroke, lung disease, diabetes, and cancer. Each disease model includes gender, race and ethnicity, and educational group as covariates. We select conditions that are prevalent in the U.S. and are characterized by significant disparities in outcomes across education, race and ethnicity, and income. The reason for this is that the incidence and progression of these conditions can potentially be reduced by preventive services, education policies, and modifications in health behaviors. These chronic conditions are treated as absorbing states, i.e., once the individual transitions into a chronic condition, the condition persists until death.

Additionally, we allow individuals to transition in and out of risk factors, such as smoking, binge drinking, and BMI. These transitions are estimated as a function of demographics,

past health, and risk factors. In the estimation, changes in risk behaviors alter future health outcomes and risk factors (e.g., smoking cessation may impact changes in BMI, and continuing smoking may affect incidence of lung disease). We also transition mental health, approximated by the Kessler mental distress scale, which is one of the predictors of the medical costs models.

Because the PSID sample covers a broad age range, it is smaller than the HRS sample at older ages where mortality becomes more likely. Also, the PSID does not follow respondents into nursing homes. Therefore, the FAM mortality model is estimated using a pooled PSID and HRS sample. The mortality model includes these covariates: age, gender, race and ethnicity, education, disease conditions, ADL count, binge drinking and current smoking status. Similarly, a partner mortality model is estimated from the pooled PSID and HRS data for transitions into widowhood. The covariates in the partner mortality model are age, gender, race and ethnicity, and education. These covariates are characteristics of the respondents, not the partners who are facing mortality (FAM does not simulate the characteristics of partners).

Since the PSID does not follow respondents into nursing homes, the model for nursing home residency is estimated using only HRS data. It includes these covariates: age, gender, race and ethnicity, education, disease conditions, ADL and IADL counts, and widowhood.

The marriage transition model estimates transitions between partnership status (we distinguish between single, cohabiting, and married). The model is a function of demographics, past employment status, earnings, mother's education and number of children. There are also childbearing models that estimate new births for each gender. Childbearing is modeled as an ordered probit model that is a function of past health and birth history, demographics, education, past work status, and past partnership status.

The employment status model estimates the probability that a person transitions into different employment states (unemployed, out of the labor force, working part-time, or working full-time). This transition is a function of demographics, marital or partnership status, education, health status and behaviors, past earnings and benefits claiming.

The combination of transition models allows us to address the aspects of the life-cycle that are most relevant for the proposed analysis. To complete the analysis, there are also models to estimate QALYs, medical expenditure, and Social Security participation.

We compute a QALY model based on the EQ-5D instrument, a widely-used, health-related quality-of-life (HRQoL) measure. The scoring system for EQ-5D was first developed using a U.K. sample.<sup>100</sup> Later, a scoring system based on a U.S. sample was generated.<sup>101</sup> The PSID does not ask the appropriate questions for computing EQ-5D, but the MEPS does.

We map EQ-5D scores from the MEPS onto the PSID data using common measures between the MEPS and PSID.<sup>102</sup> We then predict the EQ-5D scores for all PSID members running a linear regression using the variables that are transitioned in FAM (including ADL counts, IADL counts, and diseases). The microsimulation uses this linear regression to compute QALYs.

FAM has two versions of each medical cost model. For individuals who are not Medicare-eligible, the cost models are estimated from MEPS data. Once an individual becomes Medicare-eligible, their costs are estimated from MCBS data. Both sets of models include the following covariates: age, gender, race and ethnicity, education level, relationship status,

---

<sup>100</sup>Dolan (1997).

<sup>101</sup>Shaw et al. (2005).

<sup>102</sup>The main variables in this mapping are self-reported health and requiring help with ADLs.

disease conditions, and earnings. The MEPS models also include type of health insurance as a covariate. Because MCBS follows respondents for more than two years (the time step length in the FAM simulation), the FAM cost models for the Medicare-eligible population include covariates for the stage of each disease. In the initial stage, a patient has a diagnosis in the current two-year period, but did not have the diagnosis in the previous two-year period. Then, in the maintenance stage, a patient had a diagnosis in the previous two-year period and survives with the diagnosis in the current two-year period (all disease states are absorbing—it is impossible to transition out of a diagnosis). Finally, in the terminal stage, a patient has a diagnosis and dies in the current two-year period. The medical costs models tend to underestimate health care spending reported in the National Healthcare Expenditures Account (NHEA) data, due in part to underreporting of Medical costs in MEPS.

### N.3.3 Transition Models

We denote  $j_{i,0}$  to be the first age at which subject  $i$  is observed and  $j_{i,T_i}$  the last age at which he is observed. Hence, we observe outcomes at ages  $j_i = j_{i,0}, \dots, j_{i,T_i}$ . We first start with discrete outcomes which are absorbing states (e.g. disease diagnostic, mortality, and benefit claiming). We record the hazard as  $h_{i,j_i,m} = 1$  if the individual outcome  $m$  has occurred as of age  $j_i$ . We assume the individual-specific component of the hazard can be decomposed into time-invariant and time-variant parts. The time-invariant part is composed of the effect of observed characteristics,  $x_i$ , that are constant over the entire life course and initial conditions,  $h_{i,j_0,-m}$ , (where  $-m$  denotes outcomes other than the outcome  $m$ ) that are determined before the first age in which each subject is observed.

The time-variant part is the effect of previously diagnosed outcomes,  $h_{i,j_i-1,-m}$ , on the hazard for  $m$ .<sup>103</sup> We assume an index of the form  $z_{j_i,m} = x_i\beta_m + h_{i,j_i-1,-m}\gamma_m + h_{i,j_0,-m}\psi_m$ . Hence,

---

<sup>103</sup>With some abuse of notation,  $j_i - 1$  denotes the previous age at which the subject was observed.

the latent component of the hazard is modeled as

$$h_{i,j_i,m}^* = x_i\beta_m + h_{i,j_i-1,-m}\gamma_m + h_{i,j_0,-m}\psi_m + a_{m,j_i} + \varepsilon_{i,j_i,m}, \quad (17)$$

where  $m = 1, \dots, M$ ,  $j_i = j_{i0}, \dots, j_{i,T_i}$ , and  $i = 1, \dots, N$ . The term  $\varepsilon_{i,j_i,m}$  is a time-variant shock specific to age  $j_i$ . We assume that this last shock is normally distributed and uncorrelated across diseases. We approximate  $a_{m,j_i}$  with an age spline with knots at ages 35, 45, 55, 65, and 75. This simplification is made for computational reasons since the joint estimation with unrestricted age fixed effects for each condition would imply a large number of parameters. The absorbing outcome, conditional on being at risk, is defined as

$$h_{i,j_i,m} = \max\{I(h_{i,j_i,m}^* > 0), h_{i,j_i-1,m}\}.$$

The occurrence of mortality censors observation of other outcomes in a current year.

A number of restrictions are placed on the way feedback is allowed in the model. Our microsimulation model starts the health predictions at age 30, with the information on observed characteristics available at this age. We restrict it to the individuals for whom we have information from the health follow-up. This allows us to account for components that are crucial for predicting health outcomes, such as the body mass index (BMI). In sum, the models predict the probability of being in any of the states in the horizontal axis of Table N.1 at age  $a + 1$  based on the state at age  $a$ , which is described by the vertical axis of the table. The crosses indicate if the estimation of the probability of being in a state at age  $a + 1$  considers the relevant state at age  $a$ . Absorbing states are an exception. For example, heart disease at age  $a$  does not enter in the estimation of transitions for heart disease at age  $a + 1$  because it is an absorbing state: once a person has heart disease, she carries it through the rest of her life. The same is true for chronic or permanent conditions such as hypertension, having a stroke, etc.



Table N.1: Health State Transitions, Age  $a$  as Predictor of Age  $a + 1$

Age $a$	Age $a + 1$													
	Heart Disease	Hyper-tension	Stroke	Lung Disease	Diabetes	Cancer	Disability	Mortality	Smoking	Obesity	Health Insurance	DI Claim	SS Claim	SSI Claim
Heart Disease	×		×				×				×	×	×	×
Hypertension		×					×				×	×	×	×
Stroke			×				×				×	×	×	×
Lung Disease				×			×				×	×	×	×
Diabetes		×			×		×				×	×	×	×
Cancer					×		×				×	×	×	×
Disability						×	×				×	×	×	×
Smoking	×	×	×	×	×		×	×		×	×	×	×	×
BMI	×	×	×	×	×		×	×	×		×	×	×	×
Physical Activ.	×	×	×	×	×		×	×	×		×	×	×	×
Binge Drinking									×					
DI Claim												×	×	×
SS Claim													×	×
SSI Claim														×

Note: This table illustrates how health outcomes at age  $a$  predict health outcomes at age  $a + 1$ . The crosses indicate if we use the age  $a$  outcome to predict the age  $a + 1$  outcome. DI Claim: disability insurance claim; SS Claim: social security claim; DB Claim: disability benefits claim; SSI Claim: supplemental security income claim. The age  $a$  states that do not predict themselves at age  $a + 1$  are absorbing states by construction.

We have five other types of outcomes:

1. Binary outcomes that are not absorbing states, such as starting to smoke. We specify latent indices as in Equation 17 for these outcomes as well but where the lag-dependent outcome also appears as an independent variable. This allows for state-dependence.
2. Ordered outcomes, which are also modeled as in Equation 17 recognizing that the observation rule is a function of unknown thresholds  $\varsigma_m$ . Similar to binary outcomes, we allow for state-dependence by including the lagged outcome on the right-hand side.
3. Censored outcomes, such as capital income, are considered. For capital income, there are a non-negligible number of observations with zero capital income. For these observations, we consider two-part models where the latent variable is specified as in Equation 17 but models probabilities only when censoring does not occur.
4. Continuous outcomes are modeled with linear models. An example continuous outcome is the transitions in  $\log(\text{BMI})$ . We allow for state-dependence by including the lagged outcome on the right-hand side.
5. Categorical models, but without an ordering, are considered. For example, an individual can transition to being unemployed, out of the labor force, or working (either part- or full-time). In situations like this, we utilize a multinomial logit model, including the lagged outcome on the right-hand side.

In total, we have  $M$  outcomes. The parameters  $\theta_1 = \left( \{\beta_m, \gamma_m, \psi_m, \varsigma_m\}_{m=1}^M \right)$ , can be estimated by maximum likelihood. Given the normality distribution assumption on the time-variant unobservable, the joint probability of all time-intervals until failure, right-censoring, or death conditional on the initial conditions,  $h_{i,j_0,-m}$ , is the product of normal univariate probabilities. Since these sequences, conditional on initial conditions, are also independent across diseases, the joint probability over all disease-specific sequences is simply the product

of those probabilities.

For a given subject observed from the initial age,  $j_{i0}$ , to the last age,  $j_{T_i}$ , the probability of the observed health history is (omitting the conditioning on covariates for notational simplicity)

$$l_i^{-0}(\theta; h_{i,j_{i0}}) = \left[ \prod_{m=1}^{M-1} \prod_{j=j_{i1}}^{j_{T_i}} P_{ij,m}(\theta)^{(1-h_{ij-1,m})(1-h_{ij,M})} \right] \times \left[ \prod_{j=j_{i1}}^{j_{T_i}} P_{ij,M}(\theta) \right]$$

We use the  $-0$  superscript to make explicit the conditioning on  $\mathbf{h}_{i,j_{i0}} = (h_{i,j_{i0},0}, \dots, h_{i,j_{i0},M})'$ .

We have limited information on outcomes prior to this age. The likelihood is a product of  $M$  terms with the  $m$ th term containing only  $(\beta_m, \gamma_m, \psi_m, \varsigma_m)$ . This allows the estimation to be done separately for each outcome.

### N.3.3.1 Further Details on Specific Transition Models

This section describes the modeling strategy for particular outcomes.

#### Employment Status

Ultimately, we aim to simulate whether an individual is unemployed, out of the labor force, working part-time, or working full-time at time  $t$ . We treat the estimation of this as a two-stage process. In the first stage, we predict whether the individual is unemployed, out of the labor force, or working for pay using a multinomial logit model. Then, conditional on working for pay, we estimate if the individual is working part- or full-time using a probit model.

#### Relationship Status

We are interested in three relationship statuses: single, cohabiting, and married. In each case, we treat the transition from time  $t$  to time  $t + 1$  as a two-stage process. In the first

stage, we estimate if the individual will remain in his current status. In the second stage, we estimate which of the two other states the individual will transition to, conditional on leaving his current state.

## **Childbearing**

We estimate the number of children born in two-year periods separately for females and males. We model this using an ordered probit with three categories: no new births, one birth, and two births. Based on the PSID data, we found the exclusion of three or more births in a two-year period to be appropriate.

### **N.3.4 FAM simulation**

A simulation of the model starts by loading the entering cohort, generated from the ABC and CARE data. Missing values are imputed with the imputation models described in section [N.3.1](#). To this entering cohort, the model applies the transition models for mortality, health, working status, family structure, wealth, and benefit claiming, estimated from PSID, with Monte Carlo decisions to calculate the new states of the population. The simulated financial outcomes are in 2014 USD.

To match the biennial structure of the PSID data used to estimate the transition models, the simulation proceeds in two-year increments.<sup>104</sup> Once the new states have been determined, the cross-sectional models for medical costs and QALYs are applied. Computation of medical costs includes the people who died to account for end-of-life costs. The simulation ends when all simulated ABC and CARE subjects are deceased. Mortality is forced at age 120 if the simulated mortality model has not already predicted a subject's death.

---

<sup>104</sup>The end of each two-year step is designed to occur on July 1st to allow for easier matching with population forecasts from Social Security Administration (SSA).

Among the ABC and CARE subjects simulated in FAM, the years of completion of the age-30 interview range from 2003 to 2009. FAM’s two-year time step only allows the simulation of even or odd years. For this reason, we ran the simulation twice—once for the ABC and CARE subjects entering in odd years and again for the ABC and CARE subjects entering in even years.

The simulation model takes as inputs assumptions regarding the normal retirement age, future improvements in mortality, and real medical cost growth. The normal retirement age is assumed to be 67 for all ABC and CARE subjects.

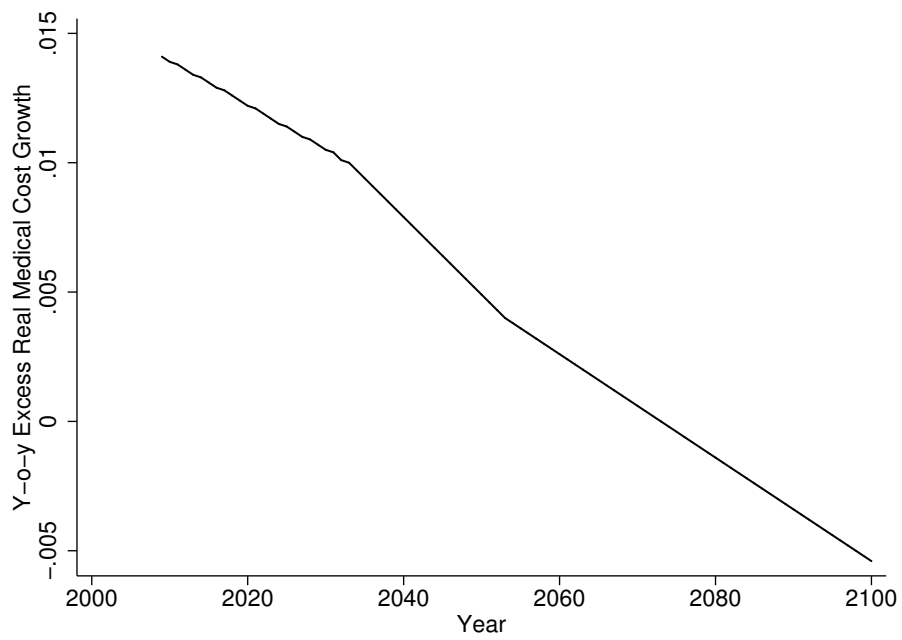
The FAM mortality model is assumed to represent mortality in 2009. The estimated mortality probabilities are reduced in simulated future years to represent improvements in mortality from sources such as medical innovation that are not included in the model. There are different adjustment factors for the populations under and over the age of 65. The mortality reduction factors are taken from the intermediate cost mortality projections in the 2013 Social Security Trustee’s Report. The medical cost growth assumptions are derived from several underlying assumptions about growth in GDP and the labor force.

The real medical cost growth factor in each year is calculated by first finding the minimum of (i) the year-over-year GDP growth plus year-over-year excess medical cost growth or (ii) the Affordable Care Act cap on year-over-year medical cost growth. In order to obtain the medical cost adjustment factor for the current year of the simulation, FAM takes the cumulative product of the yearly growth factors since 2004 and then divides it by the relative growth in the labor force since 2004.<sup>105</sup>

---

<sup>105</sup>The medical cost growth assumptions come from Congressional Budget Office and SSA assumptions. The year-over-year growth assumptions for medical costs are shown in Figure N.1. The 2010-2019 GDP assumptions are based on CBO’s analysis of the President’s Budget, March 2009. GDP assumptions for 2020-2100 are based on the 2008 OASDI Trustee’s Report long-term projection of 2.1% real GDP growth.

Figure N.1: Year-over-Year Excess Real Growth in Medical Costs



Sources: Congressional Budget Office, Social Security Administration.

Note: The year-over-year excess real medical cost growth over GDP is used to model medical cost growth in FAM.

### N.3.5 Bootstrap Sampling Strategy

The data sets used for estimating models in FAM are resampled in order to get estimates of sampling error in simulated outcomes. PSID, HRS, MEPS, and MCBS all have complex survey designs that incorporate stratification and clustering. In each of these data sets, the exact sampling strata and clusters are not publicly available, but approximations are included for the purpose of variance estimation. The bootstrap methods in FAM were developed to implement the sampling strategy of [McCarthy and Snowden \(1985\)](#): randomly select  $n_h - 1$  sampling units with replacement from stratum  $h$ , where  $n_h$  is the number of sampling units in stratum  $h$ .

PSID started with an initial sample of families in 1968. Each generation of children from those families became sample members as they formed their own households. The sample members present in the 1999–2013 PSID waves are used to estimate FAM models. The PSID data used for estimation also includes a sample targeting immigrants that entered in 1997. We randomly select one cluster from each variance estimation stratum in the 1968 and 1997 cohorts, with replacement. Any family (original sample members or their descendants) belonging to that cluster are eligible to be included in the estimation for that bootstrap sample.

The HRS is resampled at the household level. From each variance estimation stratum,  $n_h - 1$  households are randomly selected with replacement, where  $n_h$  is the number of households in the stratum. Observations for each household member are eligible to be included in the estimation for that bootstrap sample.

In MCBS,  $n_h - 1$  individuals are resampled with replacement from each variance estimation stratum, where  $n_h$  is the number of individual beneficiaries in the stratum.

MEPS is resampled by dwelling unit. Each MEPS bootstrap sample randomly selects  $n_h - 1$

dwelling units with replacement from each variance estimation stratum, where  $n_h$  is the number of dwelling units in the stratum. All members of the dwelling unit are eligible to be in the estimation sample.

### **N.3.6 Medical Costs Before Age 30 Interview**

Data on utilization of medical services is sparse before the age 30 interview. There are questions about utilization in the age 12, 15, and 21 interviews along with records of births for female subjects. We combined this with information about demographics, family structure, and parents' utilization of public services to estimate medical costs at each age from 8 to 32. Models were estimated separately for males and females. All imputation and cost models are estimated using MEPS data.

Medical costs for ages 8 to 11 were estimated in three stages using age 12 interview data. First, we impute whether or not a subject spent a week in the hospital for those subjects who are missing this information in their age 12 interview. The imputation model predicts utilization based on race and whether or not the subject was ever diagnosed with asthma between ages 8–11. Next we separate the ABC and CARE subjects into the group that spent a week in the hospital in this age range and the group that did not. For the group that did not spend any time in the hospital, we predicted medical costs as a two stage model. The first stage predicts whether there were any medical costs at all. Then, the second stage predicts the amount of medical costs for those subjects who were predicted to have some costs. We assume the group that spent time in the hospital had some medical costs, so we skip the first stage and go directly to predicting the amount. The cost models use race, asthma diagnosis, whether or not the father was absent from the home, family use of food stamps, and number of siblings as predictors.

Medical costs for ages 12 to 14 follow a strategy similar to the age 8–11 costs. First, we

impute whether or not a subject had any hospitalization for those subjects who do not report this in their age 15 interview. Imputations are based on race and presence or absence of an asthma diagnosis between ages 12–14. Again, we separated the ABC and CARE subjects into a group that had a hospitalization between age 8–11 and a group that did not. A two-stage model was used to predict medical costs for those with no hospitalization. Medical costs for the group that had a hospitalization were estimated directly from a single-stage model. These cost models use race, asthma diagnosis, whether the mother, father, or both parents were absent from the home, family use of food stamps, and number of siblings.

To estimate medical costs for ages 15 to 20, we first impute whether or not the subject spent time in the hospital for those who are missing this information in the age 21 interview. The imputation model was based on race, asthma diagnosis between ages 15–20, and, for females, the birth of any children. The age 21 interview asks about the number of days spent in the hospital. However, it does not record the ages at which these hospital stays occurred. Considering the difficulty of assigning the hospital days to specific ages in the absence of other information, we decided to use only the indicator of whether or not there were any days spent in the hospital. Next, we separated subjects into a group that spent some time in the hospital between ages 15–20 and those who did not. As before, we used the direct model to predict costs for those who had been to the hospital and used a two-stage model for those who had not. The cost models predict costs based on race, asthma diagnosis, any births (female model only), use of food stamps, whether or not the subject was working age, work status, living at college, and living with parents, and marital status.

Unlike the interviews at younger ages, the age 30 interview does not ask about utilization of medical services. To estimate costs for ages 21–31, we skipped the utilization imputation step and moved directly to cost models. We used two-stage cost models. The first stage predicts whether or not there were any costs based on race, asthma diagnosis between ages

21-31, education, use of food stamps, any births (female model only), whether or not the subject was working age, living at college, living with parents, and marital status.

## N.4 Results

If the treatment affects health outcomes, it does so by impacting the incidence rate of health conditions. For an individual, health conditions in turn impact their medical costs and quality of life. It is worth noting that lifetime medical costs depend on the severity of health events, however, they also depend on length of life in two ways: individuals who live longer accumulate medical costs over more years and medical costs increase with age. Because of this, it is important to quantify the benefits of an extra year of life that is healthy. Therefore, we look at the treatment effects on both medical costs and QALYs.

The projections for health outcomes of treatment and control groups differ in the prevalence of chronic conditions. Table N.2 shows the prevalence of six chronic conditions between ages 40 and 65, when chronic conditions and diseases are majorly developed.<sup>106</sup>

Table N.2: Prevalence of Disease by Age 65 (%)

	Female		Female	
	Control	Treatment	Control	Treatment
Cancer	0.198	0.131	0.000	0.036
Diabetes	0.319	0.298	0.129	0.366
Heart Diseases	0.158	0.211	0.110	0.105
Hypertension	0.816	0.813	0.484	0.631
Lunge Disease	0.172	0.228	0.187	0.095
Stroke Disease	0.098	0.062	0.071	0.036

Note: Prevalence of six chronic conditions in the microsimulation exercise for the treatment and control groups by age 65, and by gender.

These differences in prevalence of health conditions impact medical costs and QALYs for

<sup>106</sup>There are also differences in government transfer program participation rates. For example, 14% of males in the treatment group are expected to claim disability insurance by age 60, while the males in the control group are expected to have a higher claiming rate of 25%. The treatment effect for females at age 60 is much lower, showing only a 4% difference.

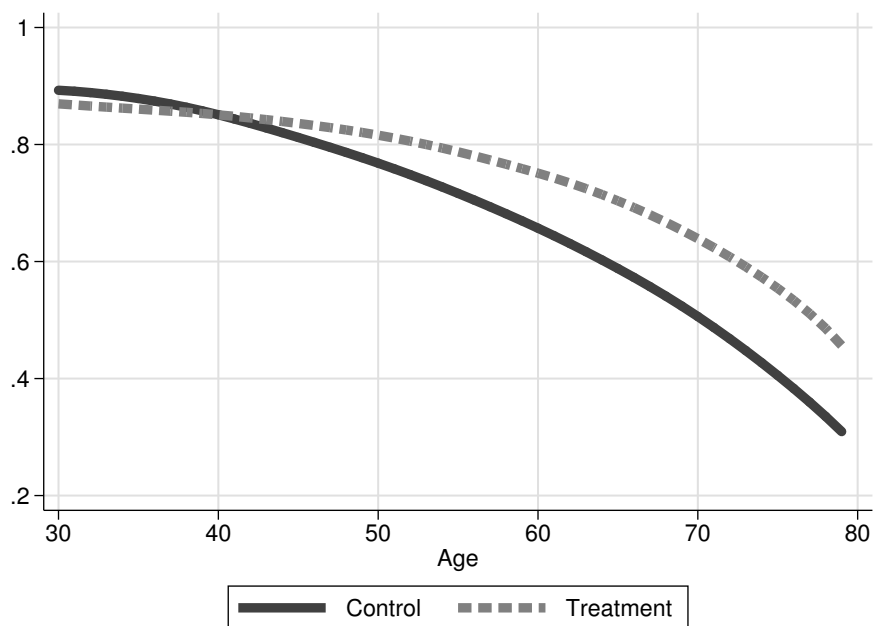
these groups. Figures N.2 and N.3 show the evolution from age 30 of average QALYs and medical costs from all sources over the simulated life-cycle of the treatment and control groups.<sup>107</sup>

---

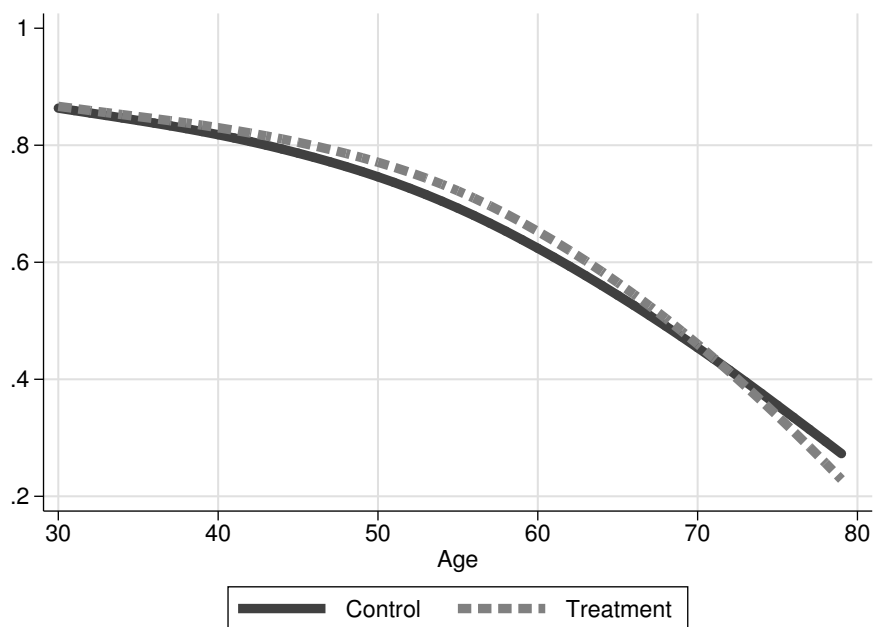
<sup>107</sup>This measure of total costs combines medical spending paid by private health insurance, public insurance, spending paid out-of-pocket, and all other sources.

Figure N.2: Predicted Quality-Adjusted Life Years

(a) Males



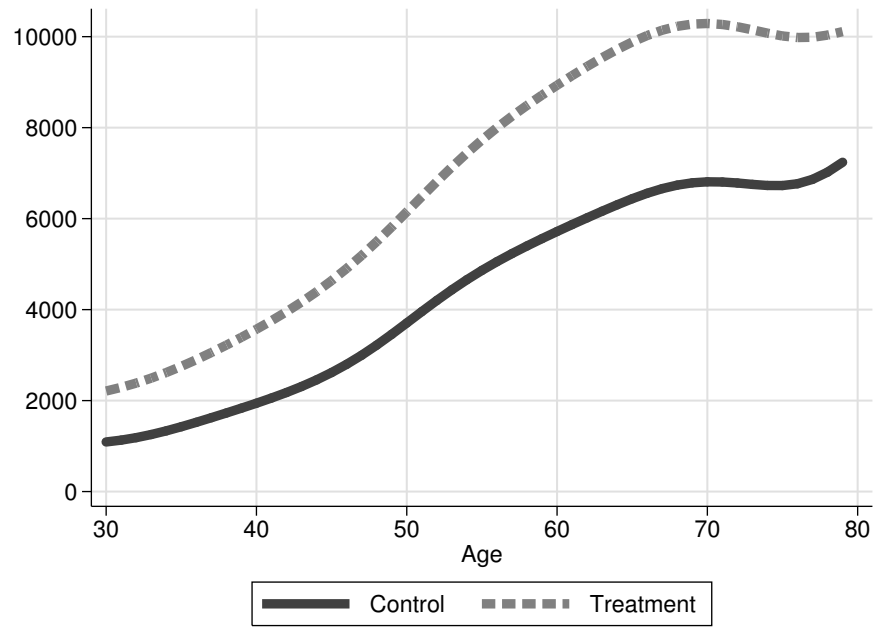
(b) Females



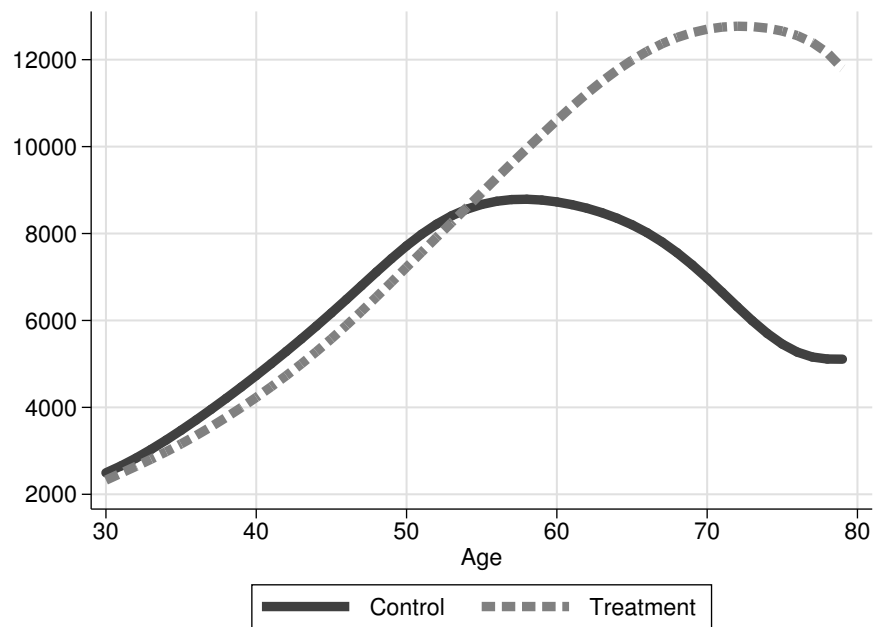
Note: Panel (a) displays the predicted Quality-Adjusted Life Years (QALYs) from age 30 for males, conditional on survival. Panel (b) displays the predicted Quality-Adjusted Life Years (QALYs) from age 30 for females, conditional on survival.

Figure N.3: Predicted Total Medical Costs

(a) Males



(b) Females

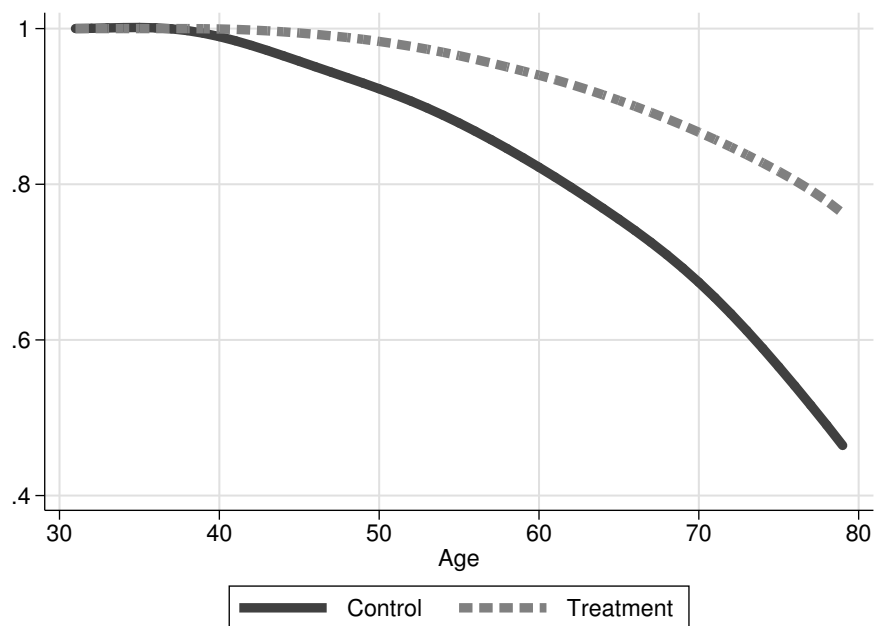


Note: Panel (a) displays the predicted total medical costs for males over the life cycle from age 30 in 2009 USD, conditional on survival. Panel (b) displays the predicted total medical costs for females over the life-cycle from age 30 in 2009 USD, conditional on survival.

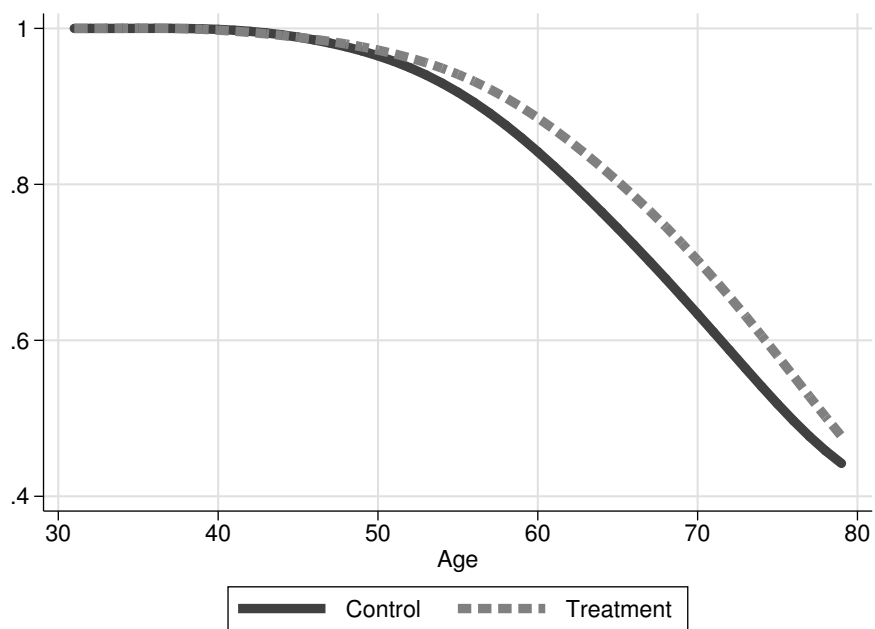
Figure N.2 shows that predicted QALYs are higher over the entire life cycle for the treatment group than for the control group. This difference appears to be larger for males. This is not surprising, given that males show larger treatment effects in prevalence rates of disease than do females. Similarly, projected total medical costs are lower for the treatment group than for the control group over the entire life-cycle, as shown in Figure N.3. Figure N.4 shows estimates of the survival probability for the treatment and control groups in FAM. There is a substantial difference in survival rates in favor of males in the treatment groups when compared with those in the control groups.

Figure N.4: Predicted Survival

(a) Males



(b) Females



Note 1 Panel (a) displays the predicted survival for males after age-30 interview. Panel (b) displays the predicted survival for females after age-30 interview. Individual-level survival probabilities are calculated by estimating age-specific mortality probabilities for each subject using the FAM simulation results. The individual-level survival probabilities are then averaged over the number of subjects in each group who had completed the age 34 interview.

To quantify the impact of treatment on health outcomes, for each simulated cohort, we compute the present discounted value of the difference between treatment and control in the average value of QALYs and average medical costs. We do this for every age. We assume a discount rate of 4% for this analysis, and we include outcome averages from the start of the simulation at age 30 until death for each cohort. We follow the recent literature on benefit-cost analysis of medical interventions and value a QALY of 1 to be worth \$150,000.<sup>108</sup> The raw difference in present discounted value at birth between treatment and control groups amounts to \$37,316 (2014 USD) per individual. Accounting for attrition, this difference falls to \$19,985. Table N.3 shows the present discounted value of health outcomes for treatment and control cohorts in the top panel. The bottom panel of Table N.3 decomposes the treatment effects on health by gender. The difference in health by treatment group is larger for males than for females. To account for a lack of consensus on QALY valuation, we also present the results for a lower valuation of QALYs, following Cutler and Meara (1998), in the second row of each panel of Table N.3.<sup>109</sup> In this case of lower valuation of QALYs, there is still a positive difference in net health outcomes in favor of the treatment groups over the control groups of \$26,374. Accounting for attrition, this difference falls to \$13,323.

Table N.3: Present Discounted Value of Health Net of Medical Costs, by Treatment Status

QALY Valuation	Treatment		Control		Treatment Effect	
QALY valued \$150,000	\$817,505		\$780,189		\$37,316	
QALY valued \$100,000	\$513,615		\$487,241		\$26,374	
	Females	Males	Females	Males	Females	Males
QALY valued \$150,000	\$806,622	\$828,387	\$777,446	\$787,732	\$29,177	\$40,655
QALY valued \$100,000	\$504,058	\$523,172	\$483,063	\$498,731	\$20,995	\$24,441

Note: This table summarizes the present discounted value at birth, at a 4% annual discount rate in 2014 USD, of the health outcomes measured through QALYs, net of medical costs, from age 30 until death.

<sup>108</sup>For a discussion on the limitations of a single value per QALY, see Pinto-Prades et al. (2009) and Mason et al. (2009). Grosse (2008) summarizes the history of QALY valuation.

<sup>109</sup>In Appendix P, we show how the internal rate of return and benefit-cost ratios are impacted by a range of valuations for QALYs.

## O Expanding the Analyzed Outcomes

In this appendix, we expand the analysis in Section 4 by considering an additional set of outcomes. Instead of considering 95, we consider 357 when pooling the sample and 343 and 355 when for males and females, respectively.<sup>110</sup> The difference between this set of outcomes and the set we present in Section 4 is that (i) the outcomes in Section 4 contain outcomes we monetize when computing the cost-benefit analysis; and (ii) it is not entirely clear if treatment should have an effect on the outcomes we add in this appendix. The 95 outcomes in this appendix include the 95 outcomes in Section 4.

To illustrate the difference between the outcomes in this appendix and the outcomes in Section 4, consider two outcomes related to employment. In Section 4, we include employment at age 30. In this appendix, we include both employment at age 30 and job satisfaction at age 30. Job satisfaction is a subjective variable, and it is not as clear what the effect of treatment on outcome is.

Table O.1 presents the results for the pooled sample, and Table O.2 and O.3 present the results for males and females. It is still the case both that (i) we find a majority of positive treatment effects; and (ii) more than 10% of these positive treatment effects are significant at the 10%. It is true that the results weaken if compared to those we consider in Section 4, but this is expected given how we construct the lists of outcomes. Appendix O.1 present the results by categories.

Table O.1: Combining Functions, Pooled Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Pos. TE	60	62	60	65	63	59	59	61
	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.013)</b>	<b>(0.013)</b>	<b>(0.013)</b>
% Pos. TE   10% Significance	28	23	19	16	21	24	21	27
	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.026)</b>	<b>(0.092)</b>	<b>(0.026)</b>	<b>(0.000)</b>	<b>(0.000)</b>	<b>(0.000)</b>

<sup>110</sup>We fail to estimate some treatment effects because in some cases the treatment indicator and the controls entirely determine the outcome. For the 95 in Section 4, we do not have this problem.

Table O.2: Combining Functions, Male Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Pos. TE	52 (0.289)	54 (0.184)	39 (0.987)	46 (0.671)	40 (0.961)	55 (0.158)	57 <b>(0.079)</b>	56 (0.171)
% Pos. TE   10% Significance	15 (0.132)	16 <b>(0.053)</b>	9 (0.539)	8 (0.618)	8 (0.724)	17 <b>(0.026)</b>	14 (0.118)	16 <b>(0.053)</b>

Table O.3: Combining Functions, Female Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
% Pos. TE	67 <b>(0.000)</b>	68 <b>(0.000)</b>	70 <b>(0.000)</b>	72 <b>(0.000)</b>	71 <b>(0.000)</b>	62 <b>(0.013)</b>	61 <b>(0.000)</b>	63 <b>(0.000)</b>
% Pos. TE   10% Significance	31 <b>(0.000)</b>	25 <b>(0.000)</b>	31 <b>(0.000)</b>	25 <b>(0.000)</b>	33 <b>(0.000)</b>	23 <b>(0.000)</b>	18 <b>(0.053)</b>	26 <b>(0.000)</b>

## O.1 Outcomes by Category

Table O.4: Combining Functions by Category, Pooled Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	100 (0.000)	96 (0.000)	100 (0.000)	96 (0.000)	100 (0.000)	96 (0.000)	96 (0.000)	96 (0.000)
Noncognitive Skills	49 (0.539)	54 (0.316)	49 (0.513)	58 (0.184)	60 (0.145)	50 (0.461)	52 (0.395)	54 (0.329)
Mother's Employment, Education, and Income	50 (0.671)	50 (0.237)	25 (0.987)	75 (0.039)	50 (0.671)	50 (0.566)	50 (0.263)	25 (0.987)
Childhood Household Environment	67 (0.105)	80 (0.013)	53 (0.132)	53 (0.474)	60 (0.079)	80 (0.026)	60 (0.224)	93 (0.000)
Adult Household Environment	78 (0.000)	67 (0.066)	56 (0.342)	56 (0.368)	44 (0.684)	78 (0.000)	67 (0.026)	67 (0.026)
Education, Employment, Income	54 (0.395)	54 (0.382)	59 (0.263)	71 (0.066)	57 (0.368)	54 (0.329)	54 (0.342)	61 (0.171)
Crime	33 (0.513)	33 (0.868)	67 (0.513)	33 (0.882)	33 (0.908)	33 (0.461)	33 (0.816)	33 (0.842)
Childhood Health	71 (0.013)	71 (0.053)	57 (0.237)	57 (0.237)	57 (0.224)	71 (0.013)	64 (0.053)	79 (0.000)
Adult Health	63 (0.013)	60 (0.013)	66 (0.000)	61 (0.013)	62 (0.026)	54 (0.237)	51 (0.447)	51 (0.395)
Mental Health	64 (0.158)	68 (0.026)	62 (0.303)	73 (0.013)	64 (0.237)	64 (0.105)	75 (0.000)	68 (0.026)

Table O.5: Combining Functions by Category | 10% Significance, Pooled Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	92 (0.000)	88 (0.000)	48 (0.013)	48 (0.000)	48 (0.013)	88 (0.000)	68 (0.000)	88 (0.000)
Noncognitive Skills	20 (0.053)	18 (0.079)	11 (0.303)	8 (0.618)	11 (0.303)	16 (0.105)	15 (0.118)	21 (0.026)
Mother's Employment, Education, and Income	0 (1.000)	25 (0.368)	25 (0.079)	0 (0.395)	25 (0.053)	0 (1.000)	25 (0.039)	0 (0.461)
Childhood Household Environment	40 (0.000)	7 (0.579)	40 (0.000)	27 (0.039)	47 (0.000)	27 (0.092)	27 (0.079)	40 (0.053)
Adult Household Environment	11 (0.434)	22 (0.039)	22 (0.211)	11 (0.224)	22 (0.197)	11 (0.355)	0 (1.000)	0 (0.645)
Education, Employment, Income	21 (0.026)	18 (0.145)	22 (0.079)	14 (0.132)	29 (0.079)	21 (0.039)	14 (0.250)	21 (0.132)
Crime	33 (0.039)	0 (0.303)	33 (0.250)	0 (1.000)	0 (1.000)	0 (0.316)	0 (1.000)	0 (0.303)
Childhood Health	43 (0.000)	36 (0.013)	36 (0.066)	21 (0.132)	43 (0.000)	36 (0.000)	36 (0.013)	36 (0.000)
Adult Health	23 (0.000)	14 (0.211)	15 (0.171)	15 (0.118)	18 (0.066)	17 (0.250)	13 (0.250)	16 (0.079)
Mental Health	27 (0.079)	27 (0.132)	14 (0.289)	20 (0.211)	18 (0.224)	25 (0.132)	29 (0.092)	34 (0.000)

Table O.6: Combining Functions by Category, Male Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	96 <b>(0.000)</b>	84 <b>(0.000)</b>	64 <b>(0.303)</b>	60 <b>(0.382)</b>	52 <b>(0.474)</b>	96 <b>(0.000)</b>	84 <b>(0.000)</b>	84 <b>(0.000)</b>
Noncognitive Skills	42 <b>(0.908)</b>	45 <b>(0.645)</b>	28 <b>(1.000)</b>	35 <b>(0.934)</b>	34 <b>(0.987)</b>	48 <b>(0.579)</b>	50 <b>(0.461)</b>	50 <b>(0.539)</b>
Mother's Employment, Education, and Income	50 <b>(0.513)</b>	50 <b>(0.658)</b>	50 <b>(0.368)</b>	100 <b>(0.000)</b>	50 <b>(0.368)</b>	25 <b>(0.961)</b>	50 <b>(0.605)</b>	50 <b>(0.632)</b>
Childhood Household Environment	60 <b>(0.329)</b>	87 <b>(0.000)</b>	47 <b>(0.605)</b>	53 <b>(0.421)</b>	47 <b>(0.579)</b>	71 <b>(0.184)</b>	87 <b>(0.000)</b>	80 <b>(0.039)</b>
Adult Household Environment	56 <b>(0.395)</b>	44 <b>(0.724)</b>	38 <b>(0.618)</b>	33 <b>(0.803)</b>	22 <b>(1.000)</b>	56 <b>(0.276)</b>	56 <b>(0.237)</b>	44 <b>(0.592)</b>
Education, Employment, Income	32 <b>(0.987)</b>	36 <b>(0.934)</b>	32 <b>(0.974)</b>	36 <b>(0.934)</b>	36 <b>(0.947)</b>	43 <b>(0.750)</b>	32 <b>(0.974)</b>	36 <b>(0.947)</b>
Crime	33 <b>(0.882)</b>	33 <b>(0.724)</b>	33 <b>(0.513)</b>	0 <b>(0.987)</b>	33 <b>(0.987)</b>	33 <b>(0.500)</b>	33 <b>(0.789)</b>	33 <b>(0.921)</b>
Childhood Health	79 <b>(0.000)</b>	79 <b>(0.039)</b>	77 <b>(0.000)</b>	69 <b>(0.132)</b>	77 <b>(0.000)</b>	86 <b>(0.026)</b>	79 <b>(0.000)</b>	86 <b>(0.000)</b>
Adult Health	68 <b>(0.000)</b>	62 <b>(0.013)</b>	58 <b>(0.237)</b>	62 <b>(0.066)</b>	58 <b>(0.237)</b>	59 <b>(0.039)</b>	62 <b>(0.000)</b>	59 <b>(0.039)</b>
Mental Health	41 <b>(0.750)</b>	52 <b>(0.487)</b>	17 <b>(1.000)</b>	37 <b>(0.789)</b>	22 <b>(1.000)</b>	46 <b>(0.645)</b>	56 <b>(0.408)</b>	52 <b>(0.553)</b>

Table O.7: Combining Functions by Category | 10% Significance, Male Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	48 <b>(0.039)</b>	48 <b>(0.013)</b>	24 <b>(0.237)</b>	24 <b>(0.158)</b>	24 <b>(0.197)</b>	60 <b>(0.000)</b>	44 <b>(0.000)</b>	56 <b>(0.000)</b>
Noncognitive Skills	9 (0.487)	8 (0.684)	5 (0.908)	4 (0.987)	5 (0.895)	10 (0.342)	9 (0.447)	12 (0.316)
Mother's Employment, Education, and Income	0 (0.987)	0 (0.987)	25 (0.250)	25 (0.237)	25 (0.237)	0 (0.974)	0 (0.974)	4 (0.974)
Childhood Household Environment	7 (0.658)	13 (0.434)	7 (0.697)	13 (0.316)	7 (0.618)	14 (0.224)	20 (0.250)	27 (0.197)
Adult Household Environment	0 (1.000)	0 (1.000)	12 (0.197)	0 (0.697)	11 (0.408)	0 (0.618)	0 (1.000)	0 (1.000)
Education, Employment, Income	11 (0.408)	7 (0.632)	0 (0.882)	4 (1.000)	0 (1.000)	11 (0.421)	4 (0.908)	7 (0.592)
Crime	0 (1.000)	0 (1.000)	0 (0.987)	0 (0.987)	0 (0.987)	0 (0.276)	0 (1.000)	0 (0.316)
Childhood Health	29 (0.184)	29 (0.145)	46 <b>(0.000)</b>	31 <b>(0.066)</b>	31 <b>(0.079)</b>	21 (0.276)	14 (0.276)	21 (0.263)
Adult Health	27 <b>(0.000)</b>	31 <b>(0.000)</b>	12 (0.368)	14 (0.224)	10 (0.461)	31 <b>(0.000)</b>	27 <b>(0.000)</b>	26 <b>(0.000)</b>
Mental Health	0 (1.000)	7 (0.474)	0 (1.000)	0 (1.000)	0 (1.000)	2 (1.000)	2 (1.000)	2 (1.000)

Table O.8: Combining Functions by Category, Female Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	100 (0.000)	96 (0.000)	100 (0.000)	100 (0.000)	100 (0.000)	100 (0.000)	92 (0.000)	100 (0.000)
Noncognitive Skills	70 (0.000)	70 (0.000)	74 (0.000)	74 (0.000)	77 (0.000)	57 (0.237)	61 (0.105)	66 (0.013)
Mother's Employment, Education, and Income	50 (0.842)	25 (0.974)	50 (0.724)	25 (0.750)	50 (0.645)	50 (0.224)	25 (0.803)	50 (0.250)
Childhood Household Environment	73 (0.039)	80 (0.026)	60 (0.066)	67 (0.053)	60 (0.132)	64 (0.263)	64 (0.145)	79 (0.053)
Adult Household Environment	78 (0.000)	67 (0.092)	89 (0.000)	78 (0.053)	89 (0.000)	89 (0.000)	56 (0.303)	56 (0.289)
Education, Employment, Income	71 (0.013)	89 (0.000)	74 (0.000)	78 (0.000)	74 (0.000)	71 (0.026)	75 (0.000)	71 (0.000)
Crime	100 (0.000)	100 (0.000)	100 (0.000)	100 (0.000)	100 (0.000)	100 (0.000)	100 (0.000)	67 (0.421)
Childhood Health	64 (0.145)	50 (0.447)	57 (0.382)	50 (0.553)	50 (0.566)	64 (0.145)	43 (0.737)	50 (0.474)
Adult Health	45 (0.750)	44 (0.882)	50 (0.500)	52 (0.329)	51 (0.474)	38 (0.987)	34 (1.000)	33 (1.000)
Mental Health	73 (0.013)	82 (0.000)	73 (0.079)	87 (0.000)	76 (0.000)	77 (0.000)	84 (0.000)	79 (0.000)

Table O.9: Combining Functions by Category | 10% Significance, Female Sample

Category	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Cognitive Skills	92 (0.000)	64 (0.000)	92 (0.000)	92 (0.000)	88 (0.000)	88 (0.000)	28 (0.184)	80 (0.000)
Noncognitive Skills	24 (0.039)	17 (0.092)	23 (0.066)	15 (0.211)	27 (0.039)	15 (0.211)	11 (0.395)	16 (0.158)
Mother's Employment, Education, and Income	25 (0.158)	25 (0.039)	25 (0.211)	25 (0.132)	0 (0.632)	25 (0.197)	25 (0.276)	25 (0.250)
Childhood Household Environment	27 (0.079)	7 (0.724)	47 (0.000)	47 (0.000)	47 (0.000)	0 (0.908)	7 (0.408)	7 (0.395)
Adult Household Environment	22 (0.118)	22 (0.118)	11 (0.395)	22 (0.211)	22 (0.224)	0 (0.816)	0 (1.000)	11 (0.421)
Education, Employment, Income	36 (0.013)	29 (0.039)	44 (0.013)	26 (0.105)	56 (0.000)	29 (0.013)	21 (0.092)	32 (0.053)
Crime	100 (0.000)	33 (0.145)	50 (0.316)	0 (0.224)	0 (0.447)	67 (0.026)	33 (0.092)	33 (0.145)
Childhood Health	29 (0.079)	36 (0.013)	0 (0.605)	0 (1.000)	0 (0.671)	29 (0.118)	29 (0.079)	43 (0.000)
Adult Health	14 (0.145)	7 (0.829)	19 (0.105)	11 (0.303)	19 (0.105)	11 (0.342)	10 (0.553)	13 (0.237)
Mental Health	45 (0.000)	52 (0.000)	39 (0.000)	38 (0.000)	44 (0.000)	36 (0.039)	41 (0.013)	45 (0.026)

## P Sensitivity Analysis

This appendix evaluates how the estimates in the main paper vary as we alter certain sample selections from all of our data sources and other parameters. We analyze sensitivity due to (i) assumptions on the values of the discount rate; (ii) assumptions on the deadweight loss to society coming from taxes raised to fund public programs; and (iii) the magnitude of the components contributing to the benefit-cost ratio and internal rate of return.

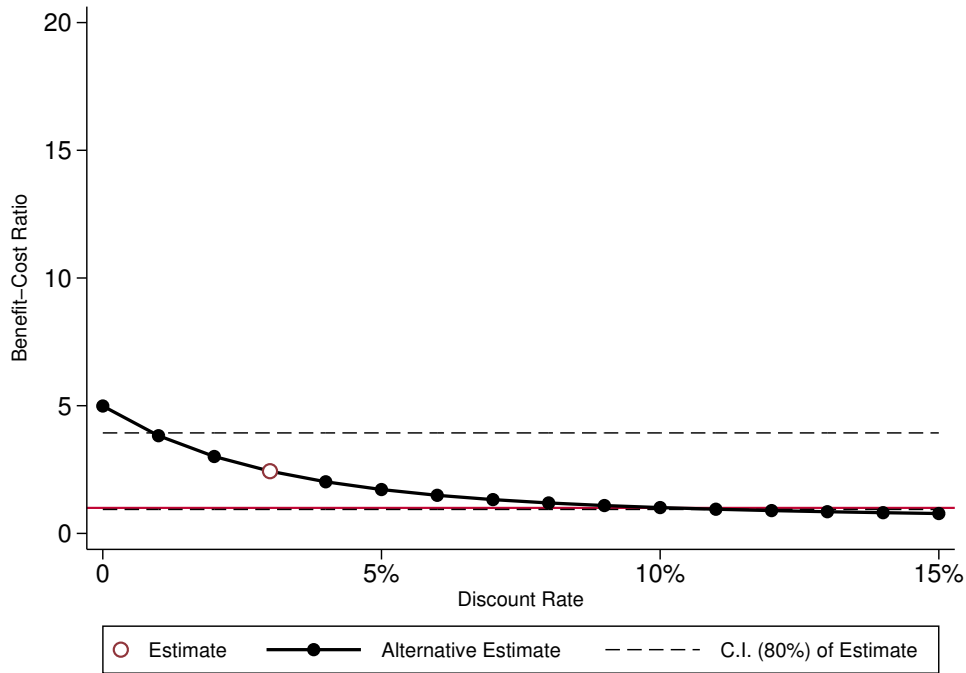
### P.1 Varying the Discount Rate

Below we examine how the benefit-cost ratio is impacted by our choice of the discount rate.

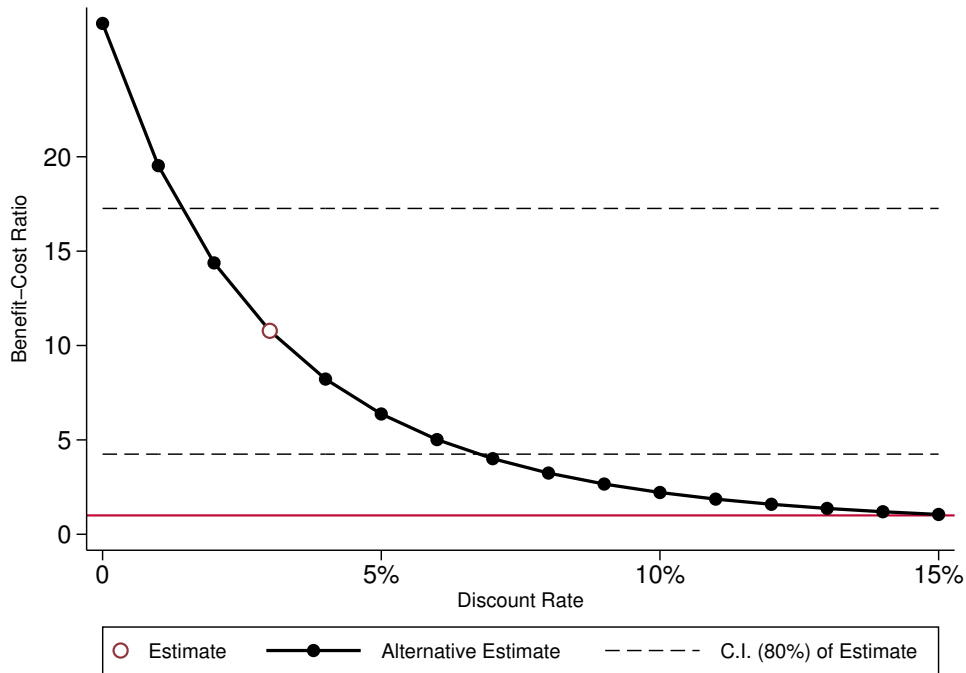
Figure P.1 displays how the benefit-cost ratio changes as we adjust the rate at which we discount the cash flows. We find that for males, the benefits of the programs exceed the costs for discount rates as high as 12%. We also find that the benefit-cost ratios for discount rates of 2–12% remain within the 80% confidence interval of our actual point estimate. The case is different for females, for whom the ratio falls below 1 at a discount rate of 6%. The benefit-cost ratio for females remain within the 80% confidence interval of our estimate for discount rates of 2–15%. Although the alternate estimates generally remain within the 80% confidence intervals for both males and females, the slope of the curves in Figure P.1 indicate that our estimates are sensitive to our choice of the discount rate, especially for males.

Figure P.1: Benefit-Cost Ratio vs. Discount Rate

(a) Females



(b) Males



Note: These graphs display how the benefit-cost ratios change for females and males as we vary the rate at which we discount to obtain the present value. The red line indicates a benefit-cost ratio of 1. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the discount rate. The estimates presented in the paper assume that the marginal cost of welfare is \$0.50 for every dollar of tax revenue. The estimates <sup>236</sup> means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

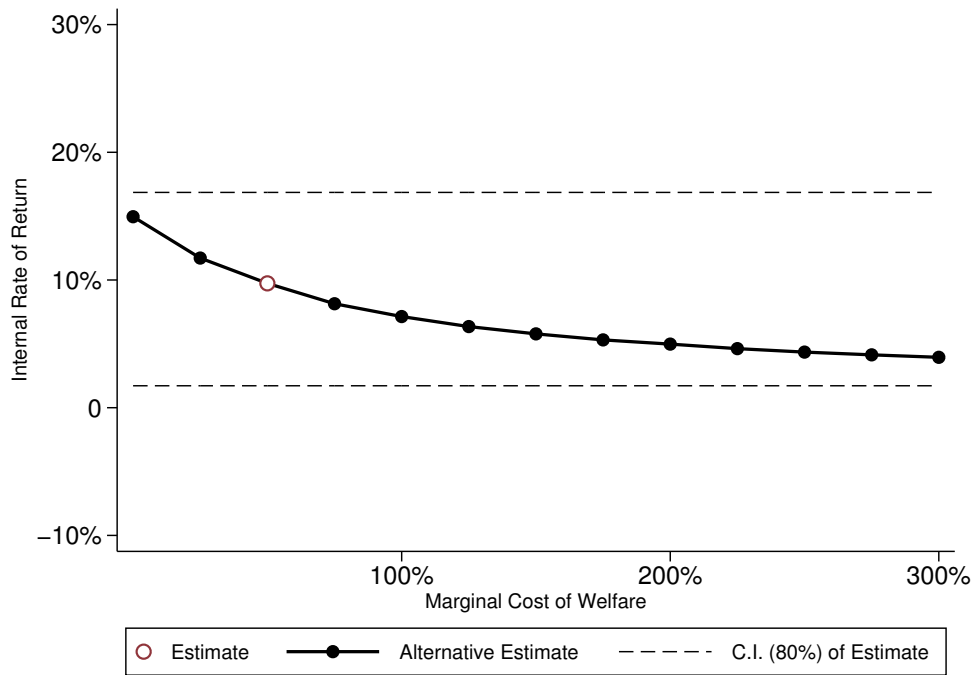
## P.2 Varying Deadweight Loss

Below we examine how our estimates of the internal rate of return (IRR) and benefit-cost ratios move with respect to changes in the marginal cost of welfare.

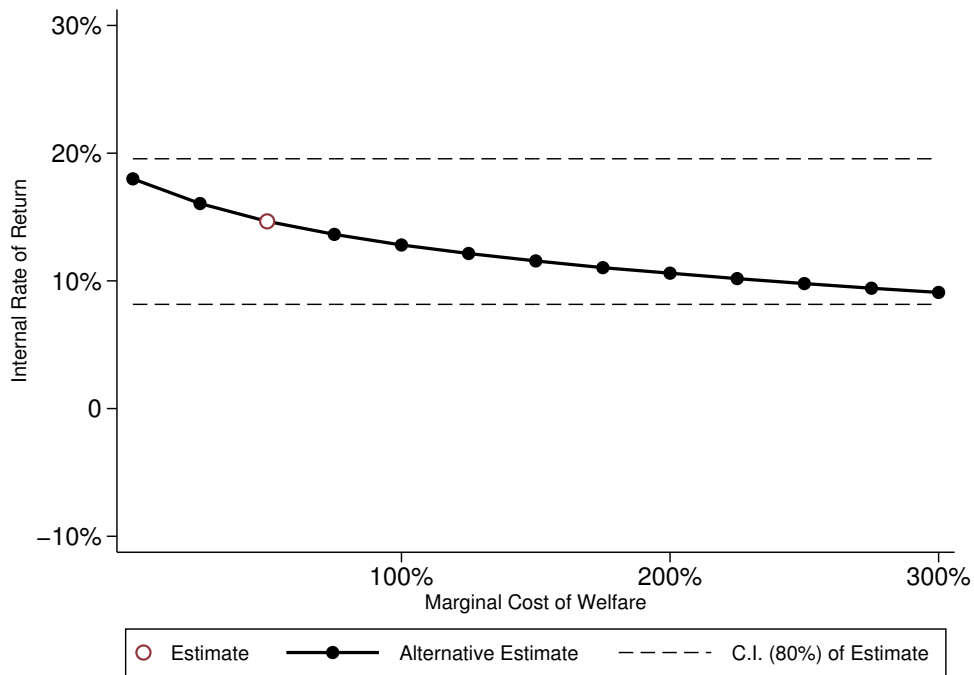
Figure P.2 shows how the IRR changes as we adjust the marginal cost of welfare. For both males and females, we find the IRRs to be most sensitive at the lower marginal costs. The IRRs for both sexes steadily decline as we increase the marginal cost of welfare to \$3 for every dollar of tax revenue. This is likely due to the fact that both females and males in treatment live longer, and are expected to receive more Medicare and Medicaid benefits in their later life. Also, we treat the costs of implementing ABC and CARE as a public cost. Thus, the steady increase in deadweight loss results in a steady decline in the IRR. Nonetheless, we see that the IRRs remain within the 80% confidence interval of our original estimate for both females and males and are not particularly sensitive to changes within the neighborhood of our assumed marginal cost of welfare (note each point on the  $x$ -axis represents a \$0.25 increment in the cost of welfare per dollar of tax revenue).

Figure P.2: Internal Rate of Return vs. Deadweight Loss

(a) Females



(b) Males

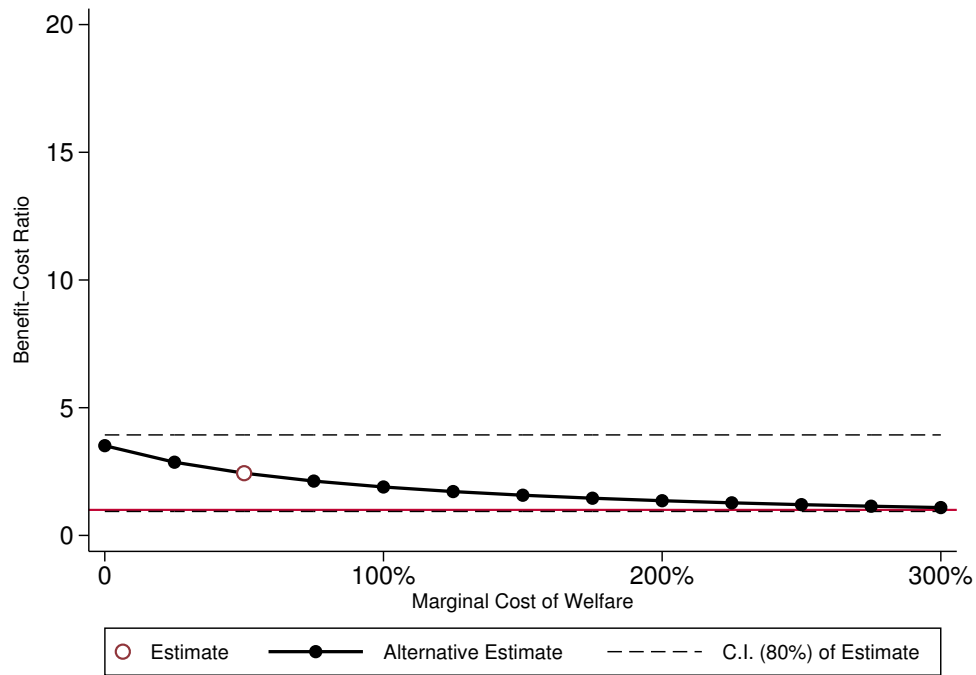


Note: These graphs display how the internal rate of return changes for females and males as we vary the marginal cost of welfare. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the marginal cost of welfare. The estimates presented in the paper assume that the marginal cost of welfare is \$0.50 for every dollar of tax revenue. The estimates are means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

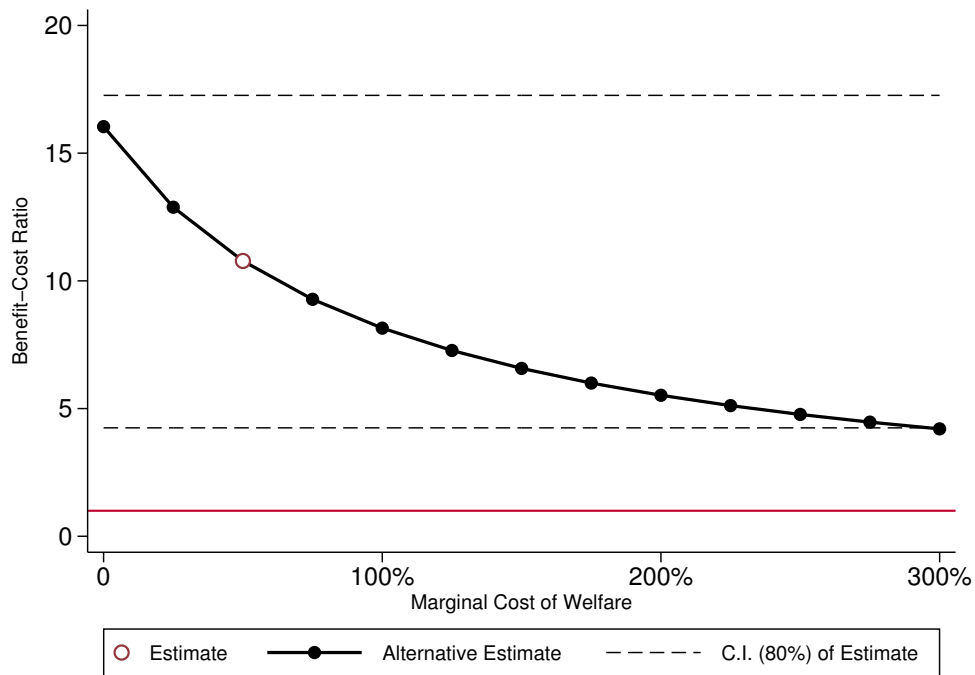
Figure P.3 illustrates how the benefit-cost ratio changes as we vary the marginal cost of welfare. For females we see that the costs continue to exceed the benefits (both are discounted at a rate of 4%) even when the marginal cost of welfare is assumed to equal \$1 for every dollar of tax revenue. For males, the benefits exceed the costs even as we increase the deadweight loss to \$3 for every dollar of tax revenue. The resilience of the benefit-cost ratio of males is attributed to the large impacts ABC and CARE had on criminal activity, which bears a large cost to society. As we increase the marginal cost of welfare, we are in fact increasing the stream of benefits associated with the reduction in criminal costs of males. At the same time, however, we are also increasing the later-life public medical costs associated with Medicare and Medicaid, as well as the program costs.

Figure P.3: Benefit-Cost Ratio vs. Deadweight Loss

(a) Females



(b) Males



Note: These graphs display how the benefit-cost ratio changes for females and males as we vary the marginal cost of welfare. The red line indicates a benefit-cost ratio of 1. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the marginal cost of welfare. The estimates presented in the paper assume that the marginal cost of welfare is \$0.50 for every dollar of tax revenue. The estimates are means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

### P.3 Varying Component Magnitudes

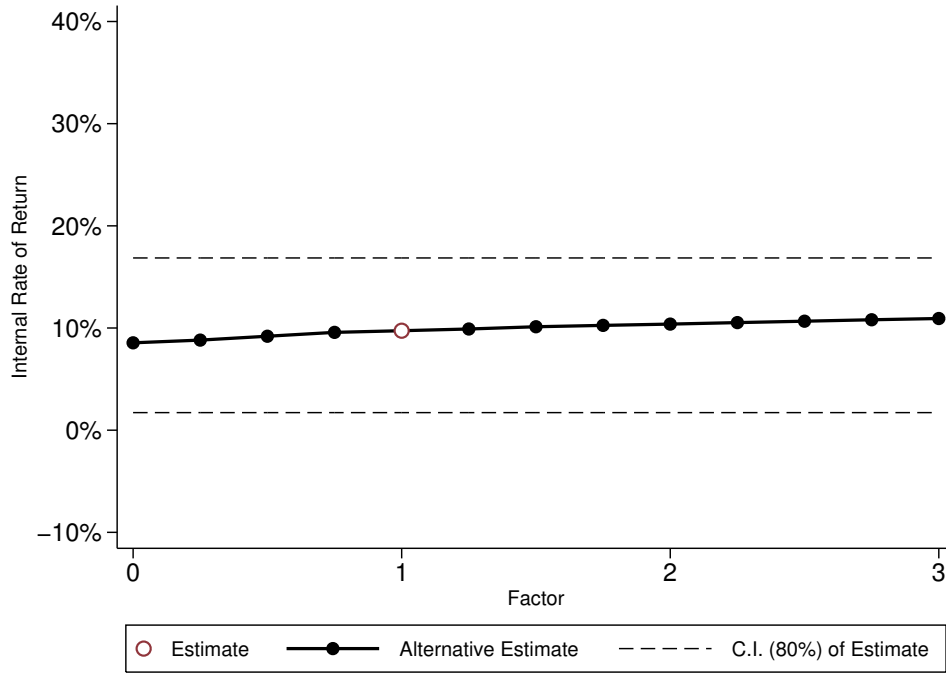
Below we explore how the internal rate of return (IRR) and benefit-cost ratio change as we increase and decrease the value of each component of the benefit and cost streams. This entails multiplying each component by factors ranging from 0 to 3. A factor of 0 is equivalent to removing the component entirely from our analysis. In the particular case of QALYs, the multiplicative factors correspond to different valuations of a year of perfect health, e.g., a QALY equal to 1. For instance, as our current estimates assume a QALY of 1 to be worth \$150,000, a factor of 0.5 corresponds to a year of perfect health being worth \$75,000, and a factor of 3 corresponds to a year of perfect health being worth \$450,000 (all values are in 2014 USD).

Figure P.4 displays how the IRR for females changes as we multiply each component by a factor between 0 and 3. We find that the IRR is stable across different levels of public-transfer income, QALYs, health costs, and criminal costs. Parental income has the biggest effect on the IRR, and is one of the only components for which the alternative IRR falls outside of the 80% confidence interval of the original estimate. This occurs when we double the parental income component in the flow of benefits. By tripling the component, we obtain an IRR exceeding 100%, which we do not display in the chart above. This sensitivity is due to both the magnitude of the treatment effect on parental income, as well as how the treatment effect took place earlier in the ABC and CARE subjects' lives. The sensitivity of the IRR to the program costs is also a result of the timing of the costs in the subjects lives. As females in the treatment groups attained higher levels of education than females in the control groups, we observe that the IRR decreases as we multiply the expenditure on education by increasingly large factors. On the other hand, this translates to additional labor income for females, which we observe to have a positive effect on the IRR. However, the IRR appears to be more sensitive to the costs of education relative to labor income, as schooling costs are borne at

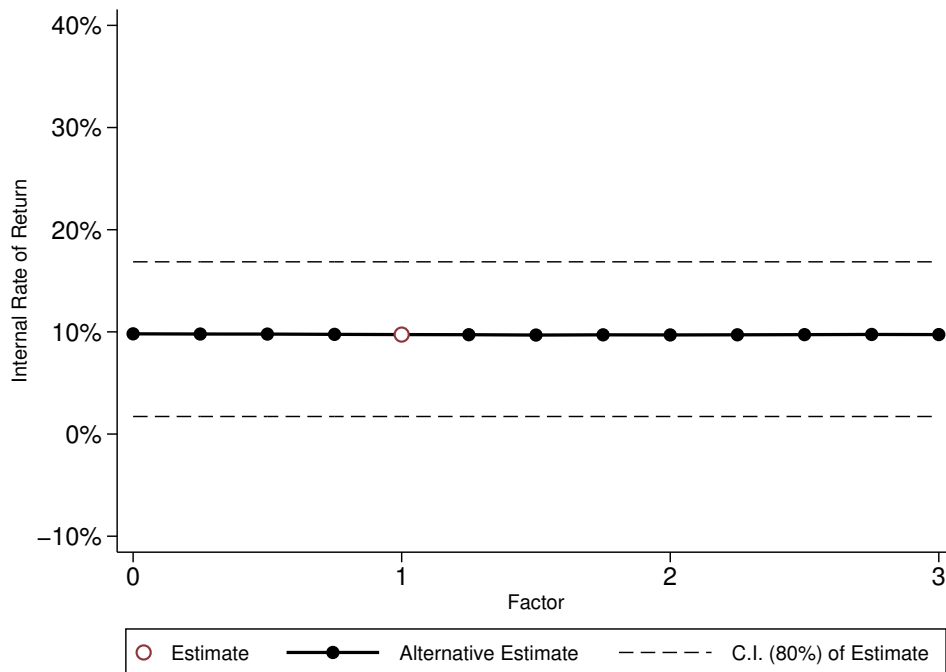
earlier stages of each subject's life.

Figure P.4: Internal Rate of Return vs. Components, Females

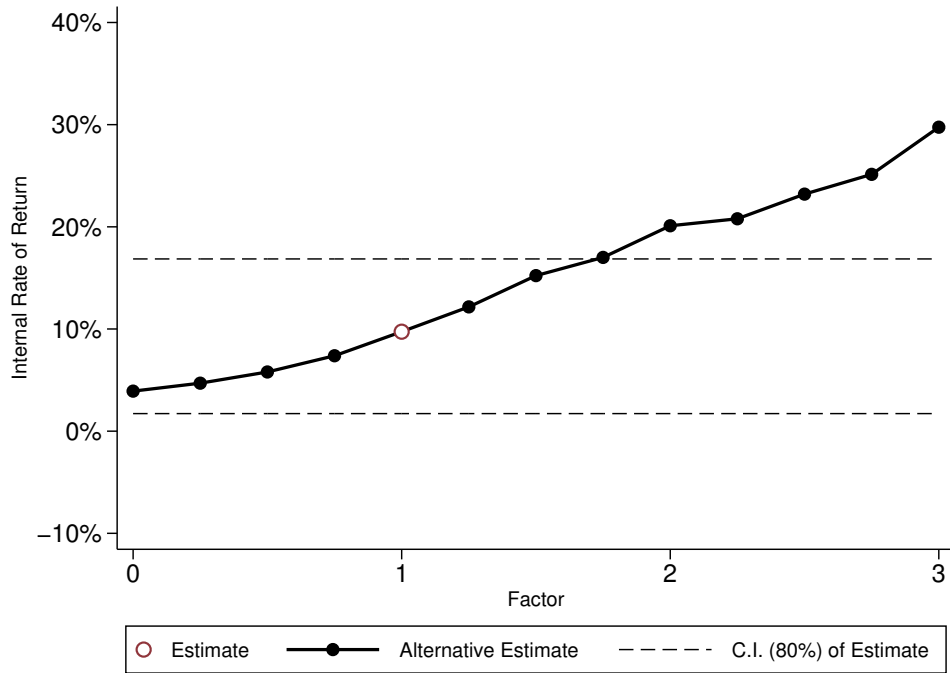
(a) Labor Income



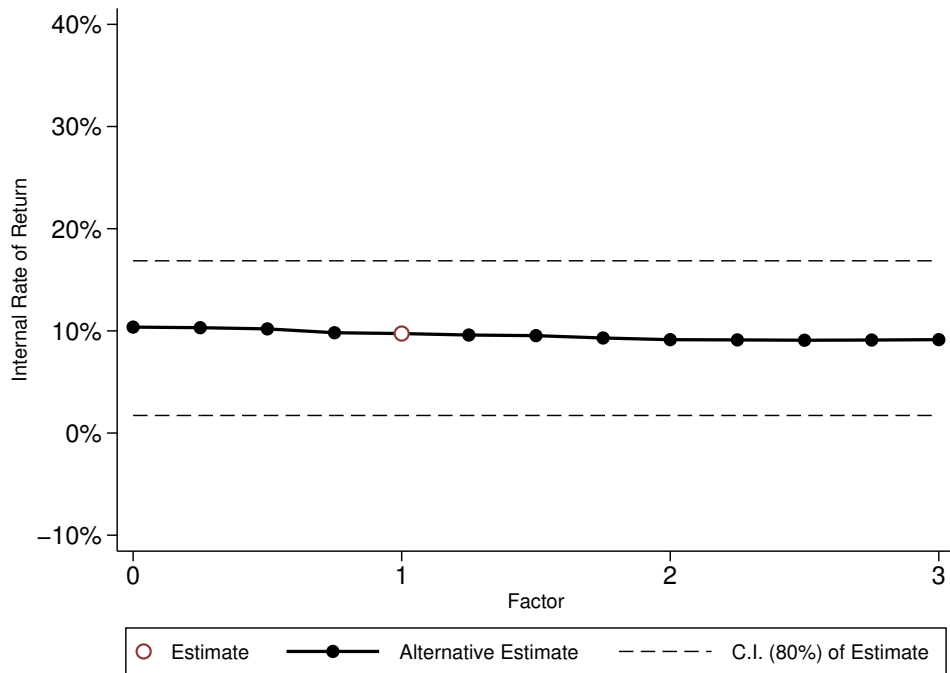
(b) Public-Transfer Income



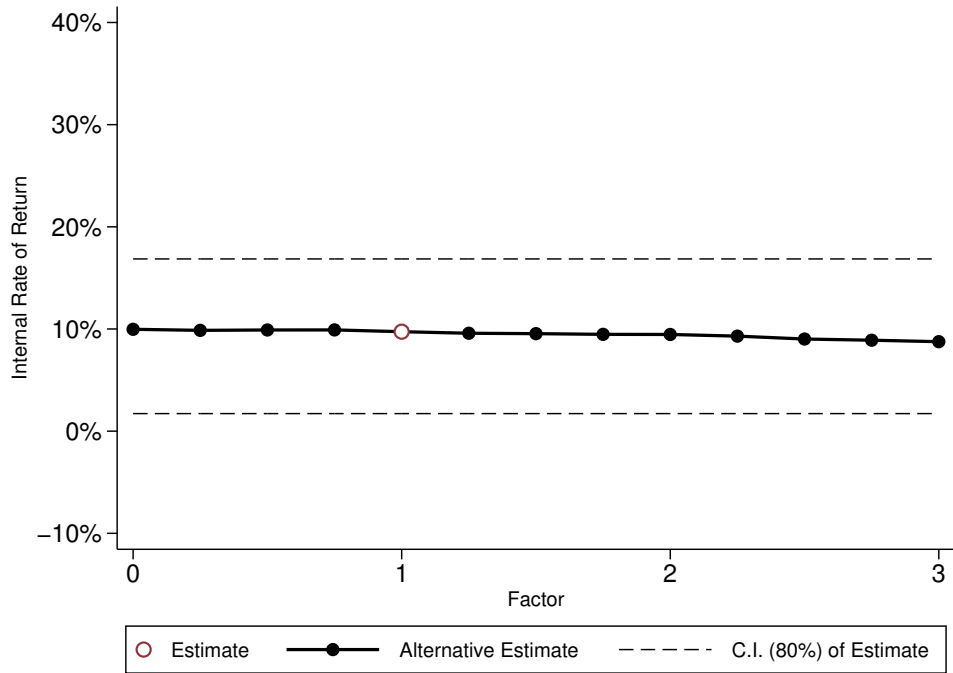
(c) Parental Income



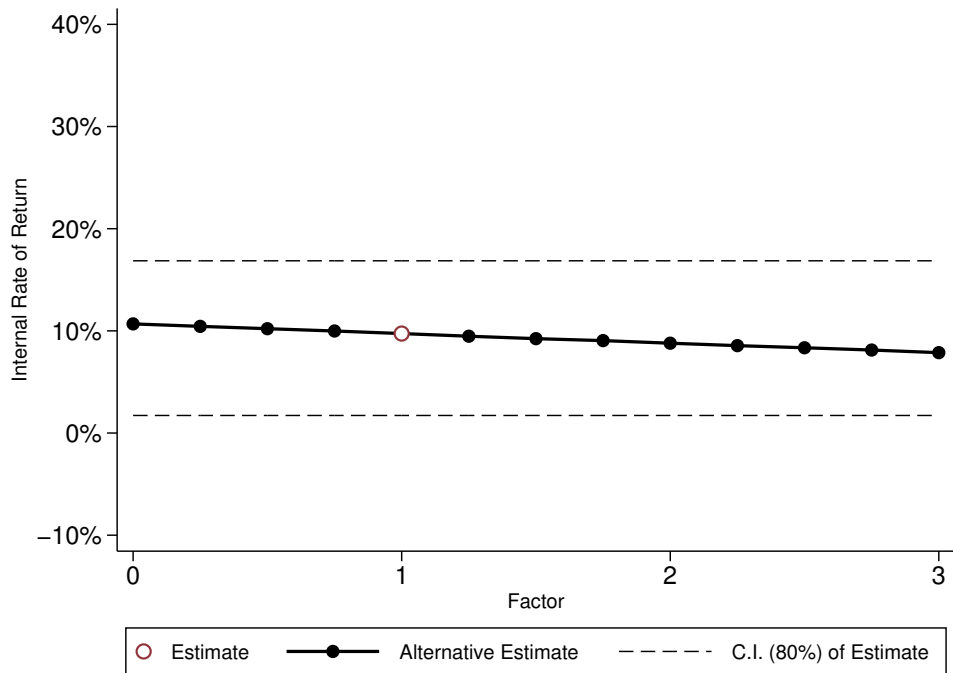
(d) Quality-Adjusted Life Years



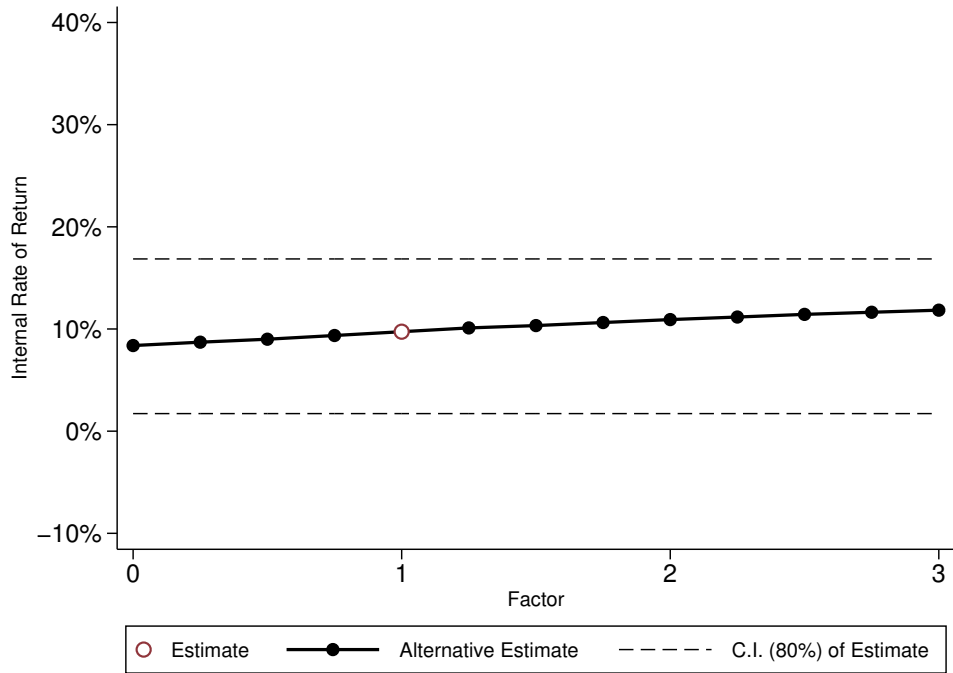
(a) Health Costs



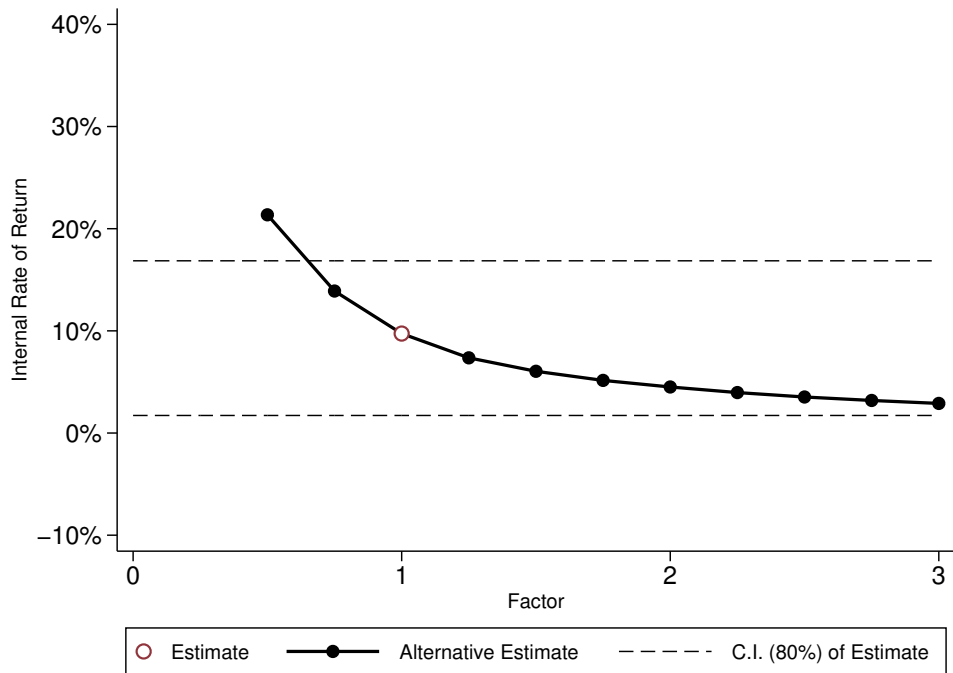
(b) Education Costs



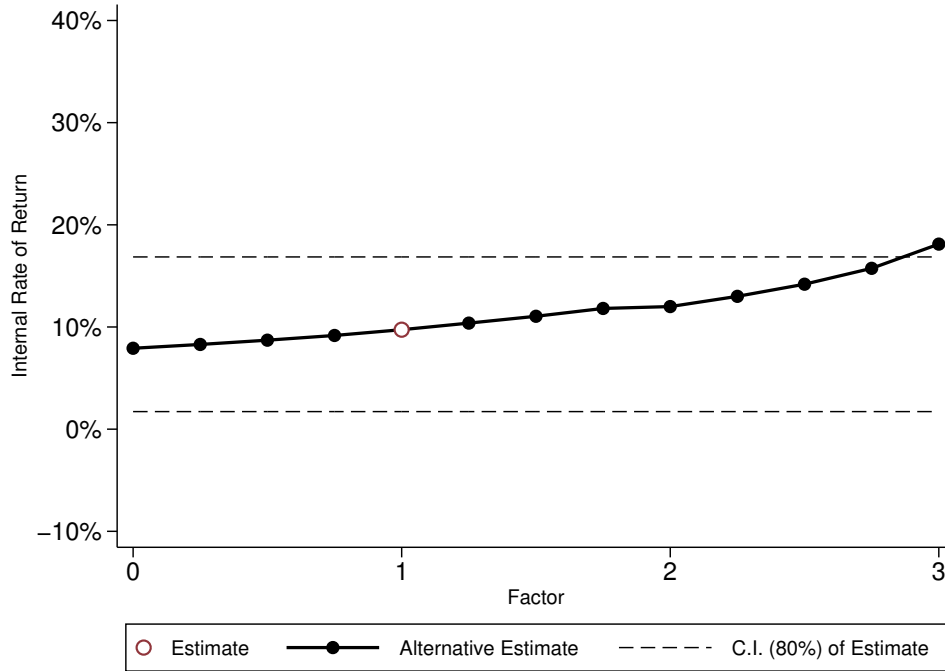
(a) Crime Costs



(b) Program Costs



(a) Control Substitution Costs



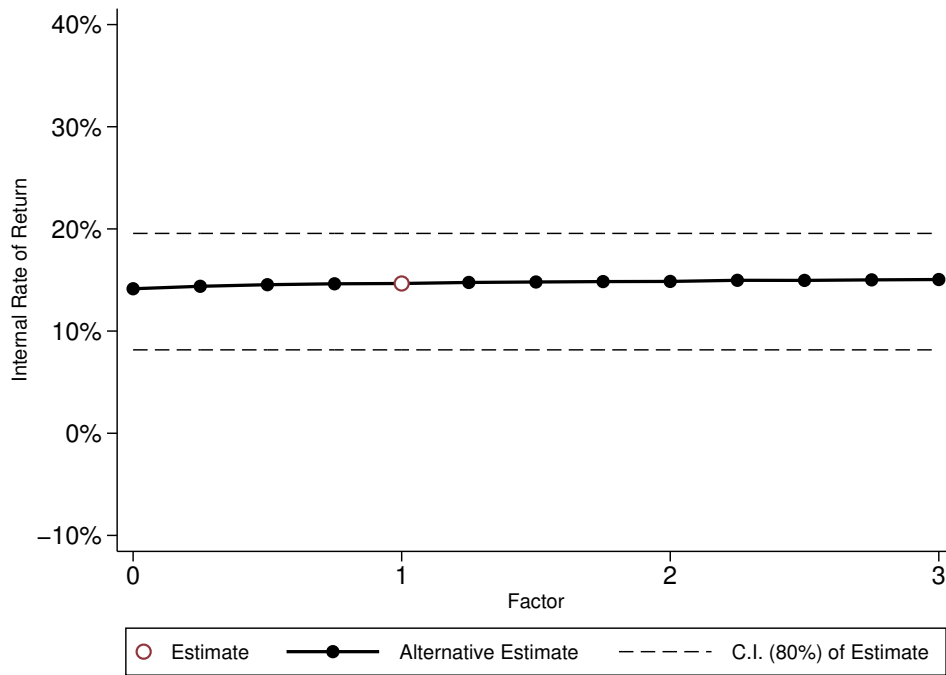
Note: These graphs display how the internal rate of return changes for females as we multiply each component by a factor from 0 to 3. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the magnitude of each component. The estimates presented in the paper are equal to the IRRs presented above when the multiplicative factor is equal to 1. The estimates are means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

Figure P.1 displays how the IRR for males changes as we vary the magnitude of each component of the benefits and costs. Our findings for males are similar to those for females: the IRR is insensitive to changes in public-transfer income, QALYs, and health costs. As the parental income and the program cost components for the female subsample is the same as those of the male subsample, we also observe that the IRR for males is sensitive to changes in both of these components. The IRR for males is similarly sensitive to increasing the weight of the cost of control substitution. Compared to females, however, as the weight increases, the internal rate of return increases at a much faster rate. This is a product of the different populations of males and females who were enrolled in alternative preschools. The IRR for the male subsample is a little less sensitive to changes in labor income and education costs

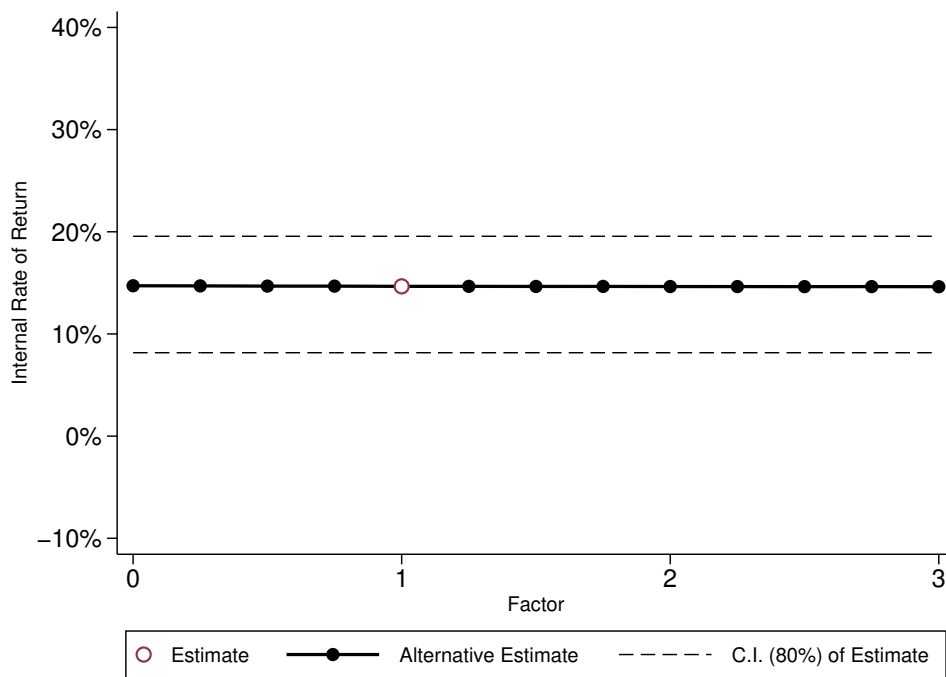
than for the female subsample, but we still observe in Figure P.1 that the IRR for males rises as we multiply the benefits stemming from those components by increasingly large factors. Finally, the IRR steadily increases for males as we increase the magnitude of the criminal cost component. This is not surprising because the reduction in the costs of crimes is the largest benefit of ABC and CARE for males.

Figure P.1: Internal Rate of Return vs. Components, Males

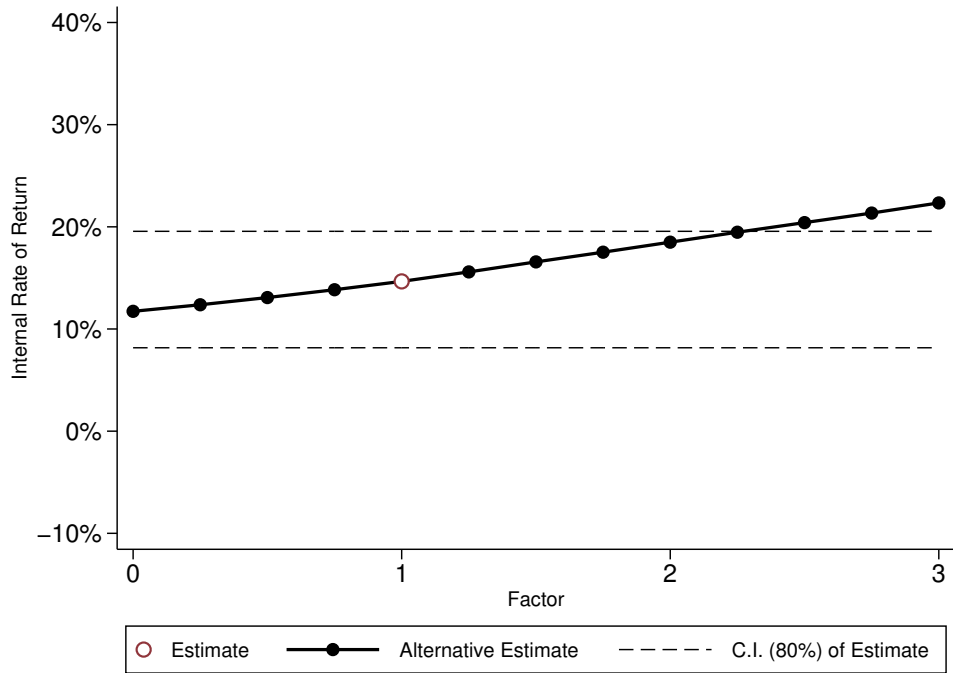
(b) Labor Income



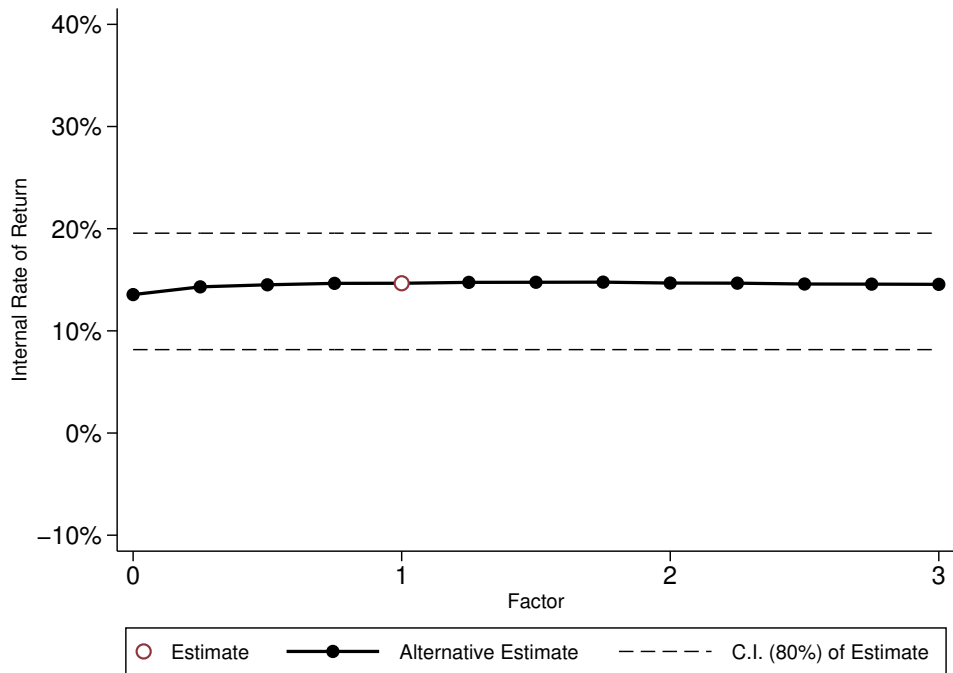
(c) Public-Transfer Income



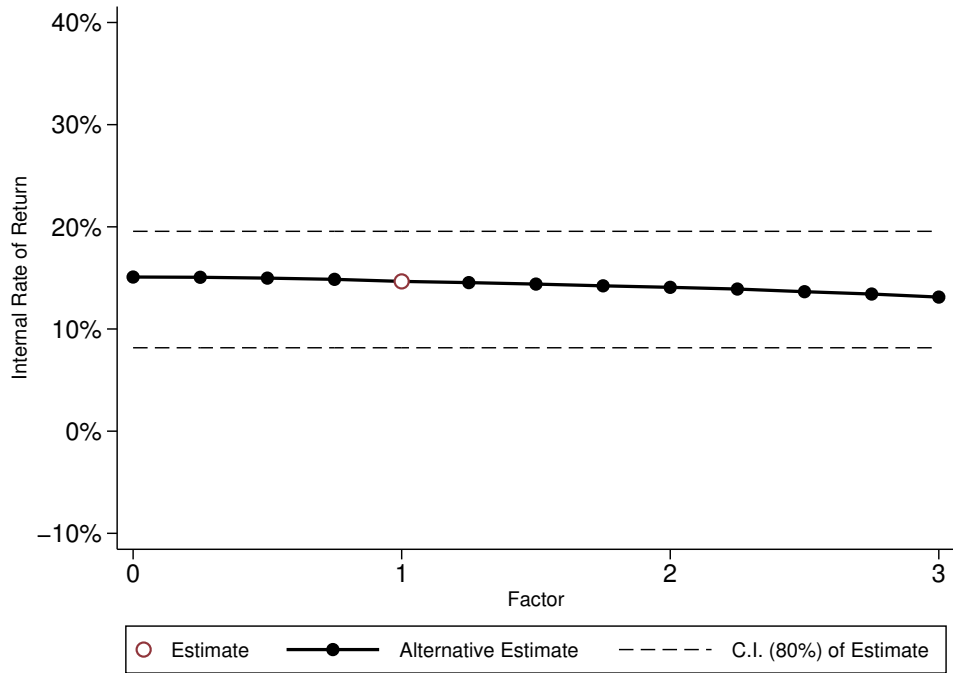
(d) Parental Income



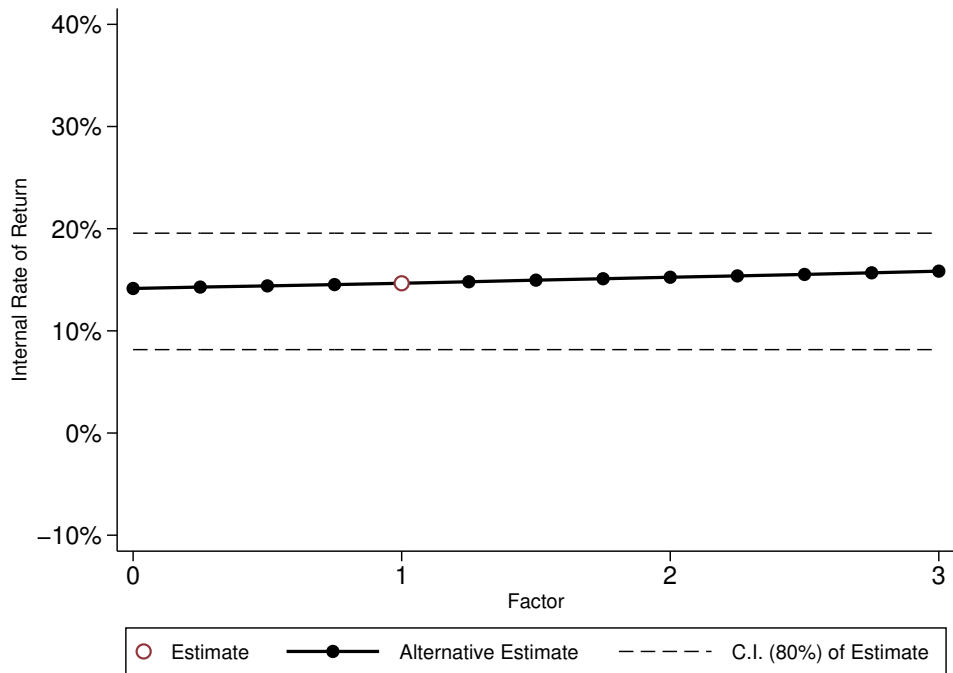
(e) Quality-Adjusted Life Years



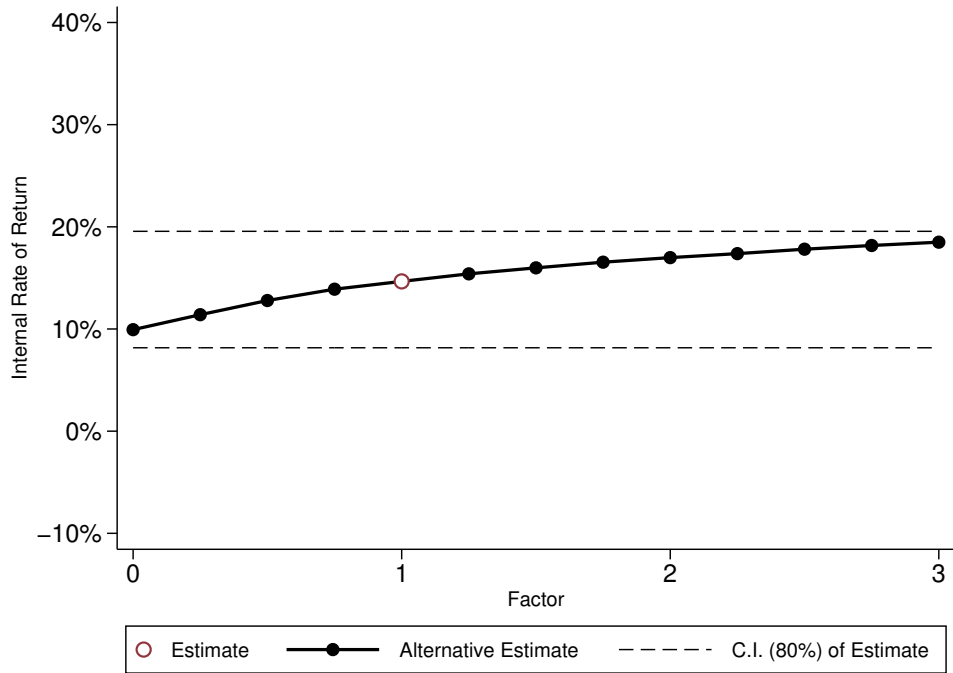
(a) Health Costs



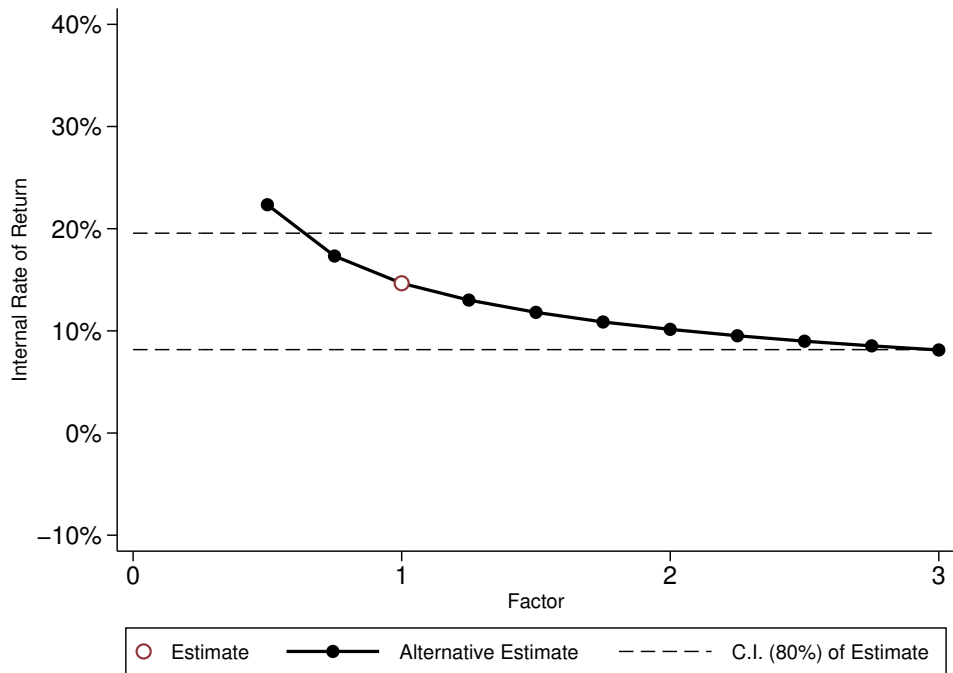
(b) Education Costs



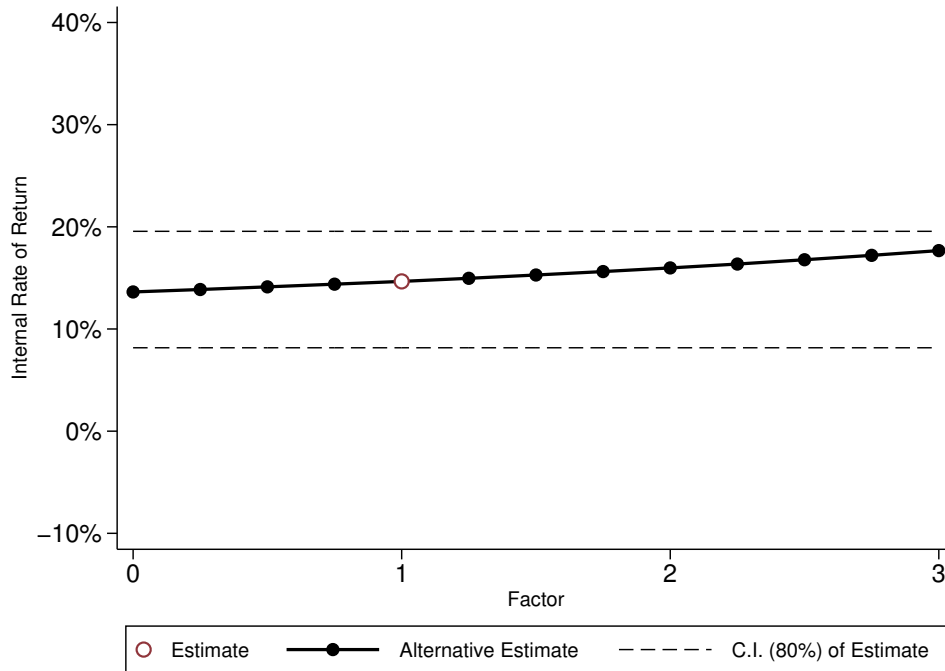
(a) Crime Costs



(b) Program Costs



(a) Control Substitution Costs

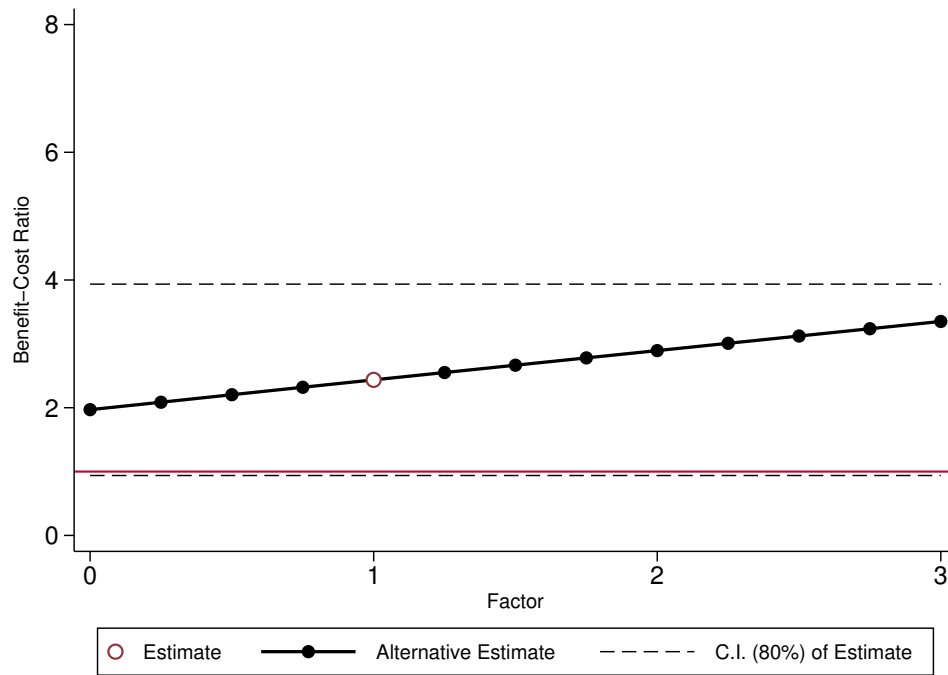


Note: These graphs display how the internal rate of return changes for males as we multiply each component by a factor from 0 to 3. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the magnitude of each component. The estimates presented in the paper are equal to the IRRs presented above when the multiplicative factor is equal to 1. The estimates are means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

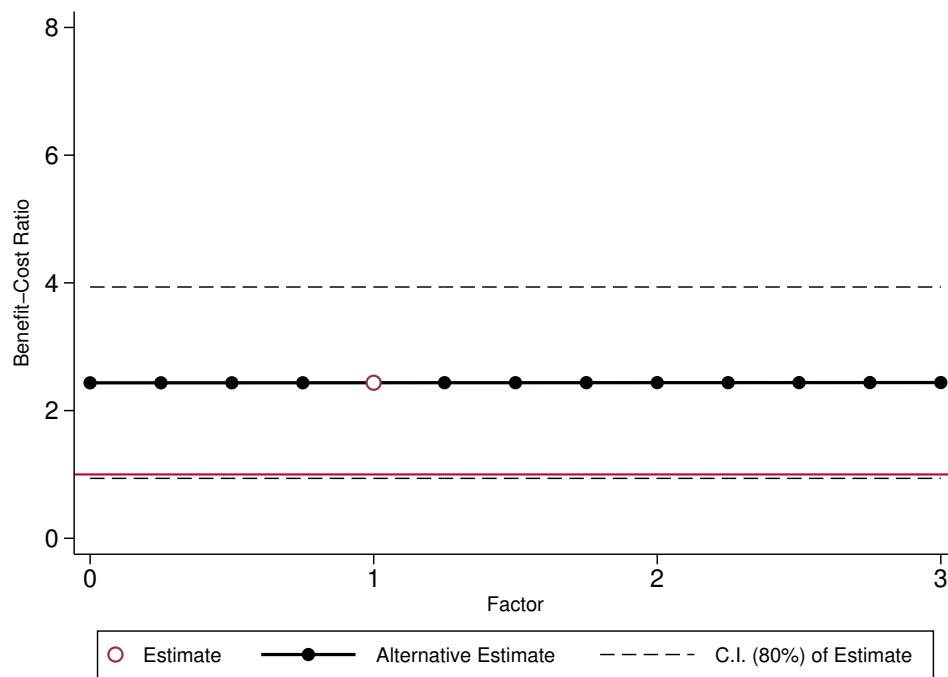
Figure P.-2 shows how the benefit-cost ratio changes for females as we multiply each component of the benefits and costs by a factor between 0 and 3. We observe that the ratio is generally insensitive to changes in all the components, except parental income, labor income, education costs, and program costs. The sensitivity to program costs is due to the fact that it is the denominator of the benefit-cost ratio. In the case of the other three components, when discounted, they exhibit the three largest present values. The sensitivity of the benefit-cost ratio to changes in these components therefore indicates the magnitude of those components relative to the rest, with parental income having the largest magnitude in terms of discounted treatment effect, followed by labor income, and then education costs.

Figure P.-2: Benefit-Cost Ratio vs. Components, Females

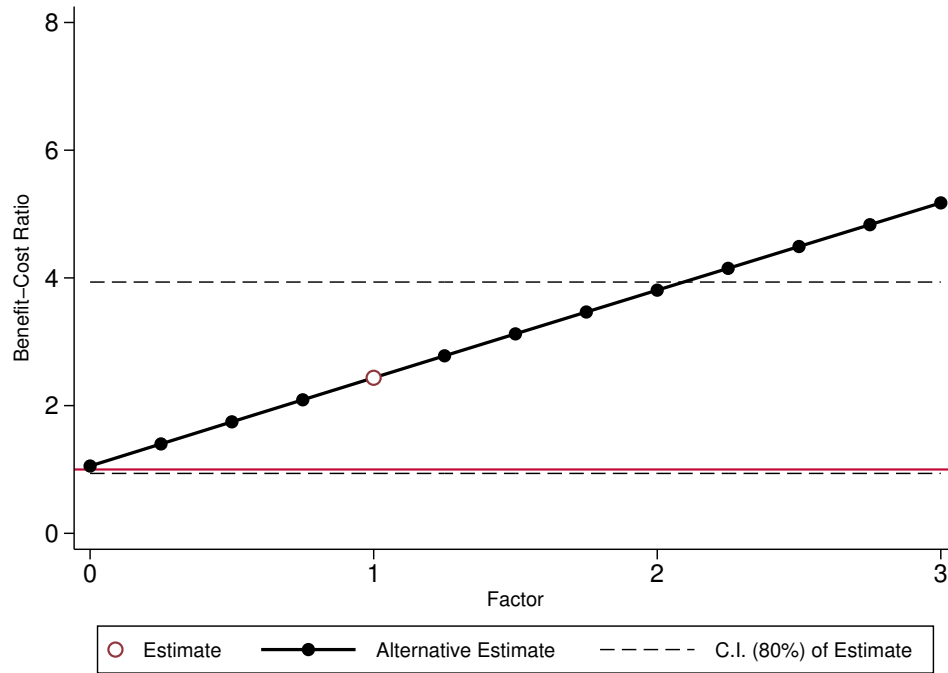
(b) Labor Income



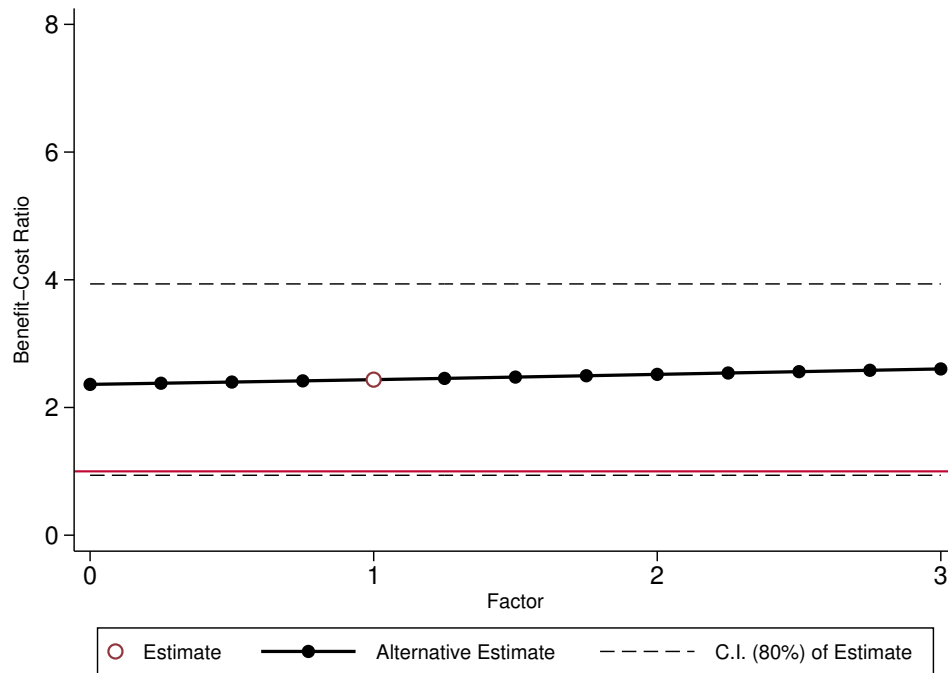
(c) Public-Transfer Income



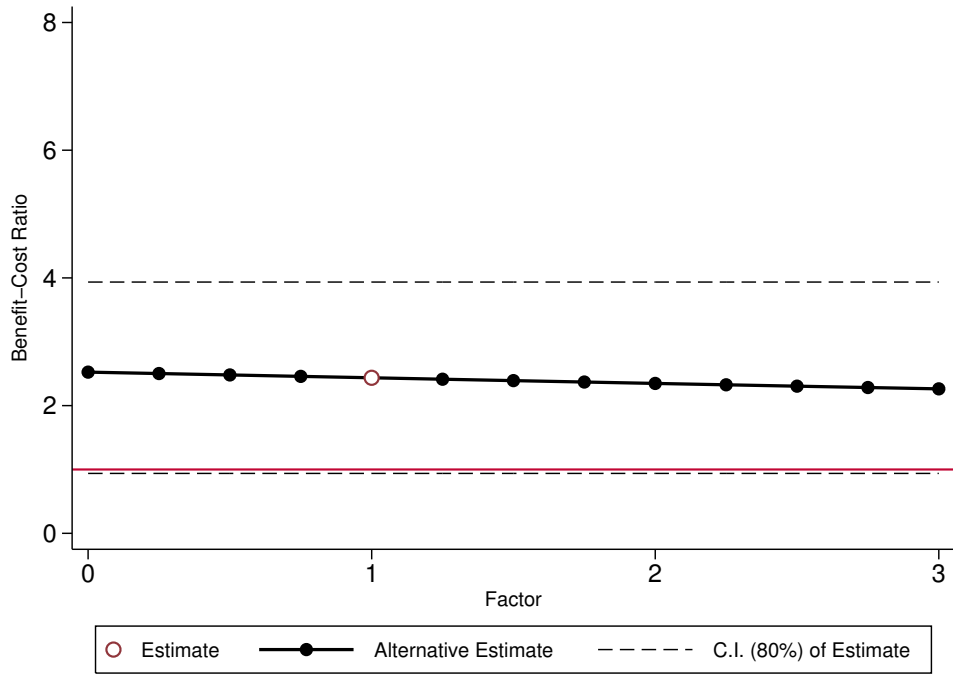
(d) Parental Income



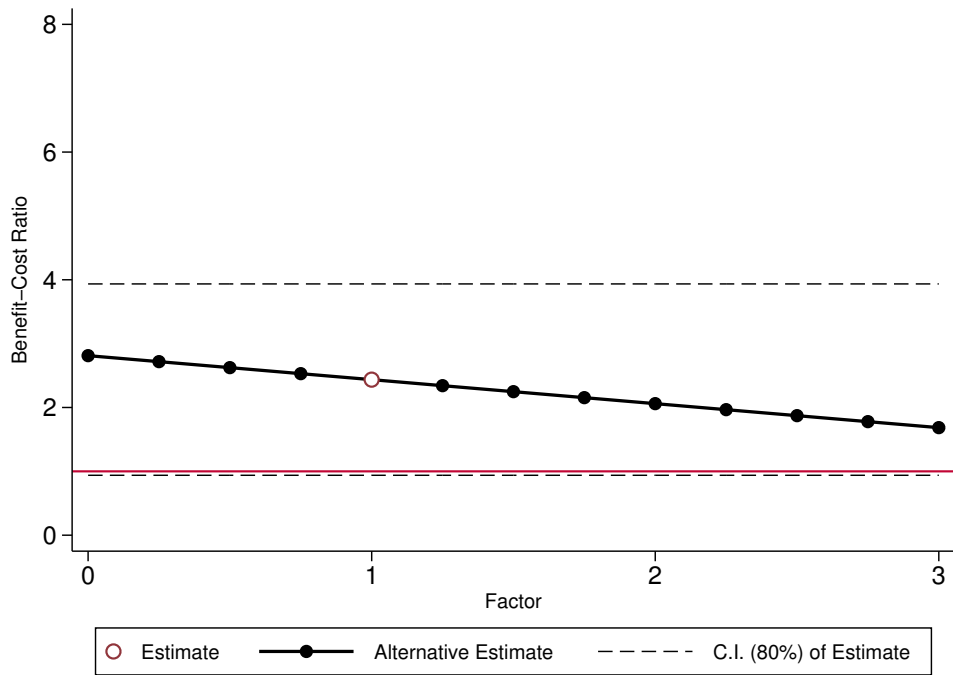
(e) Quality-Adjusted Life Years



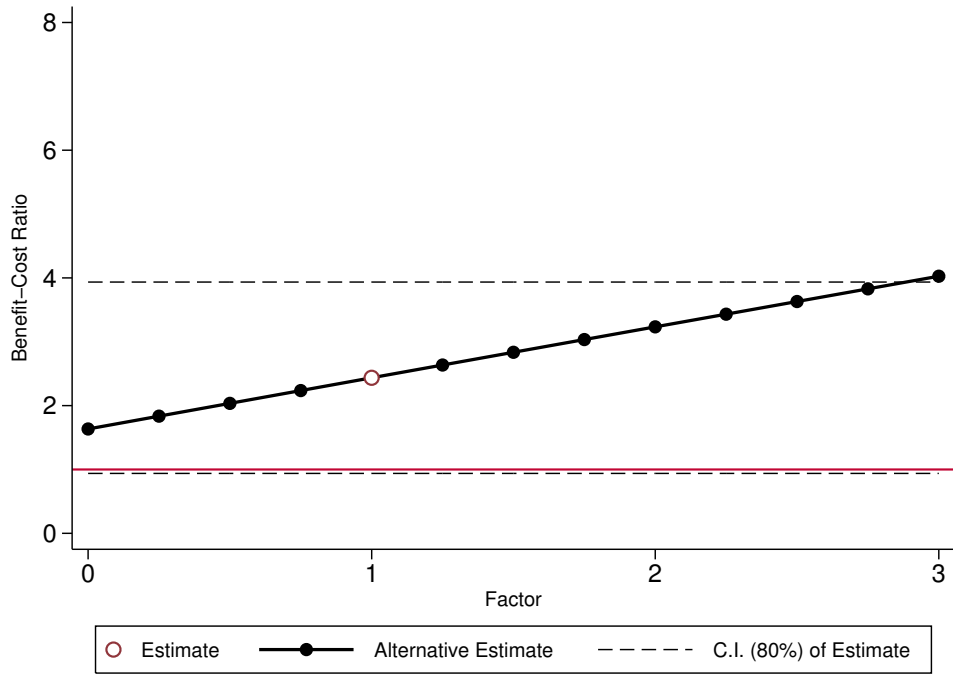
(a) Health Costs



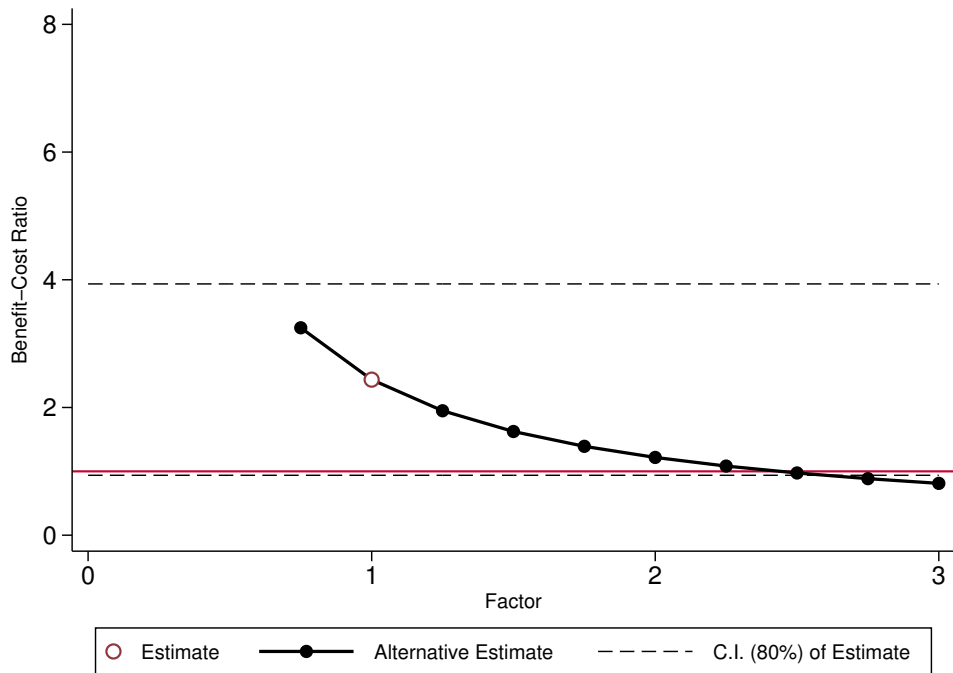
(b) Education Costs



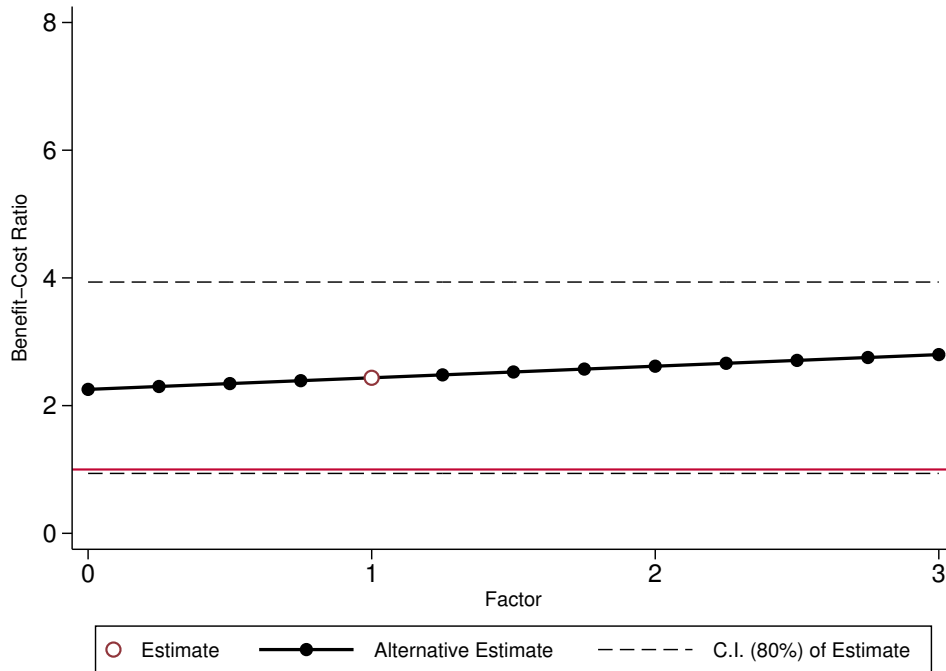
(a) Crime Costs



(b) Program Costs



(a) Control Substitution Costs

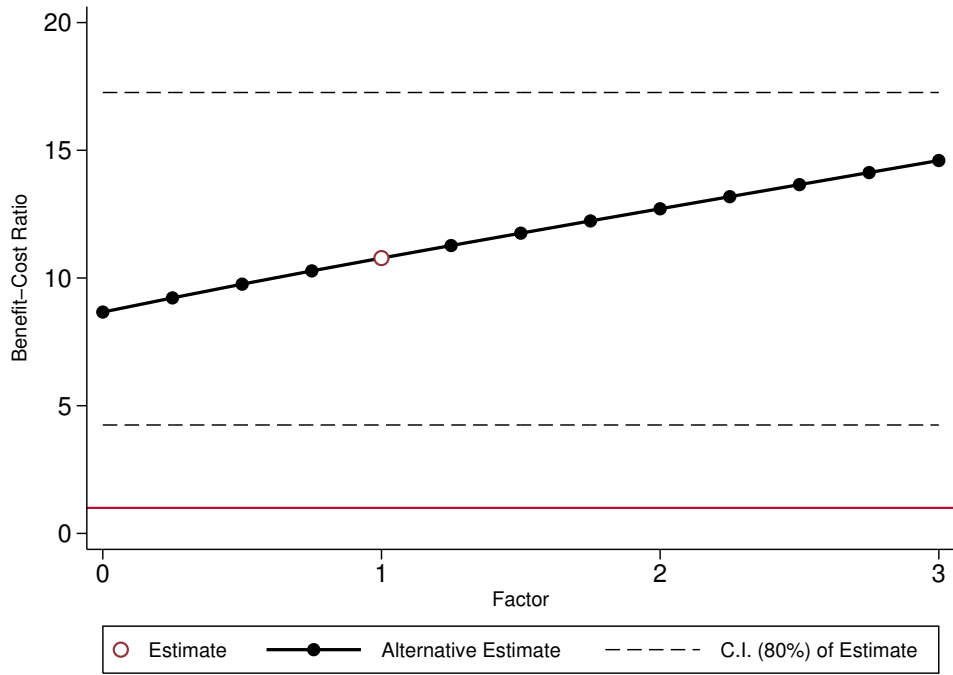


Note: These graphs display how the benefit-cost ratio changes for females as we multiply each component by a factor from 0 to 3. The red line indicates a benefit-cost ratio of 1. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the magnitude of each component. The benefit-cost ratio presented in the paper is equal to those presented above when the multiplicative factor is equal to 1. The estimates are means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

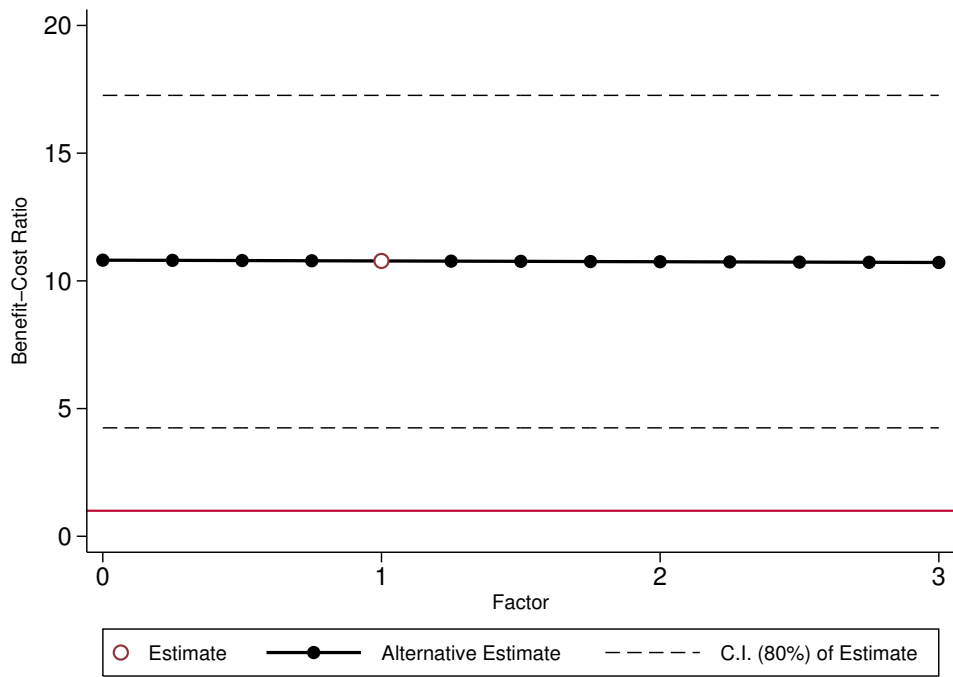
Figure P.-5 displays how the benefit-cost ratio for males varies as we multiply each component of the benefits and costs by a factor between 0 and 3. We observe that the ratio is insensitive to the scaling of public-transfer income, health costs, education costs, and control substitution costs. Barring program costs, the components that vary the benefit-cost ratio the most are crime costs, parental income, labor income, health expenditure, and QALYs, in that order. This is simply a result of the relevant magnitude of each of those components in the benefits stream after discounting.

Figure P.-5: Benefit-Cost Ratio vs. Components, Males

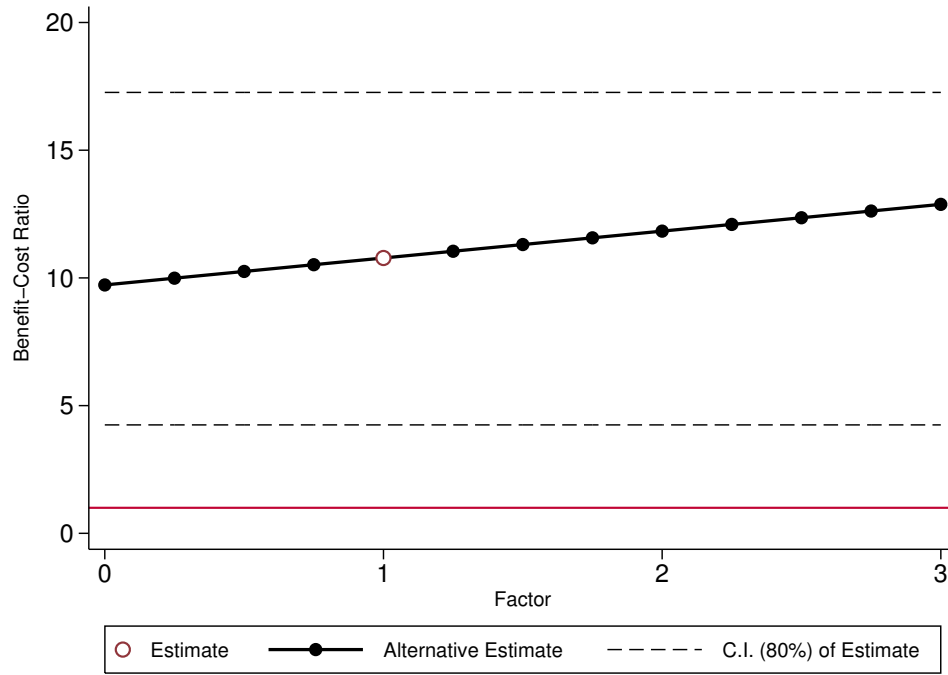
(b) Labor Income



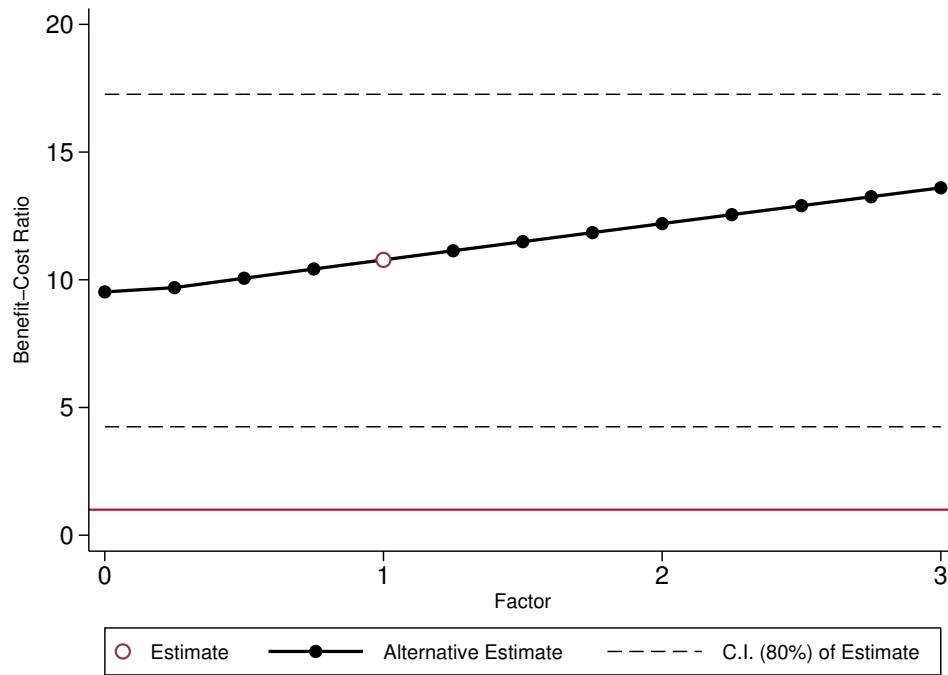
(c) Public-Transfer Income



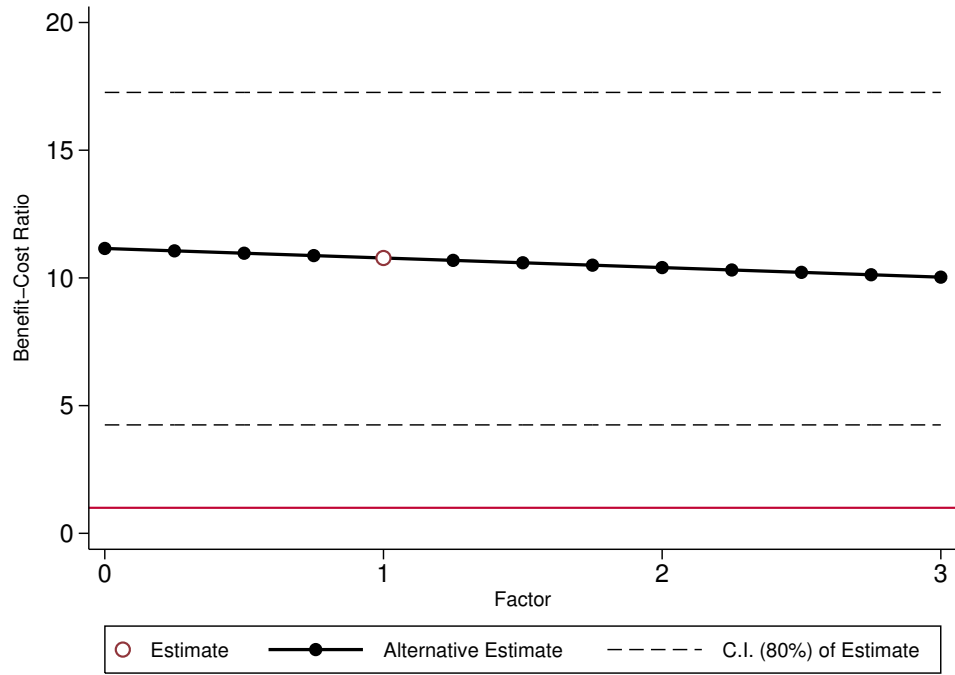
(d) Parental Income



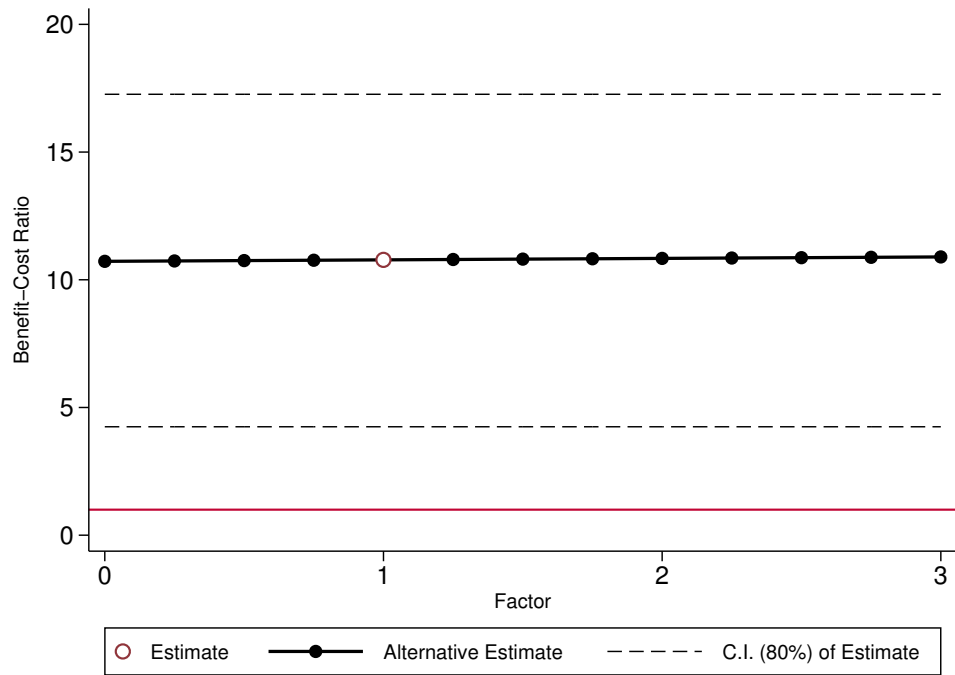
(e) Quality-Adjusted Life Years



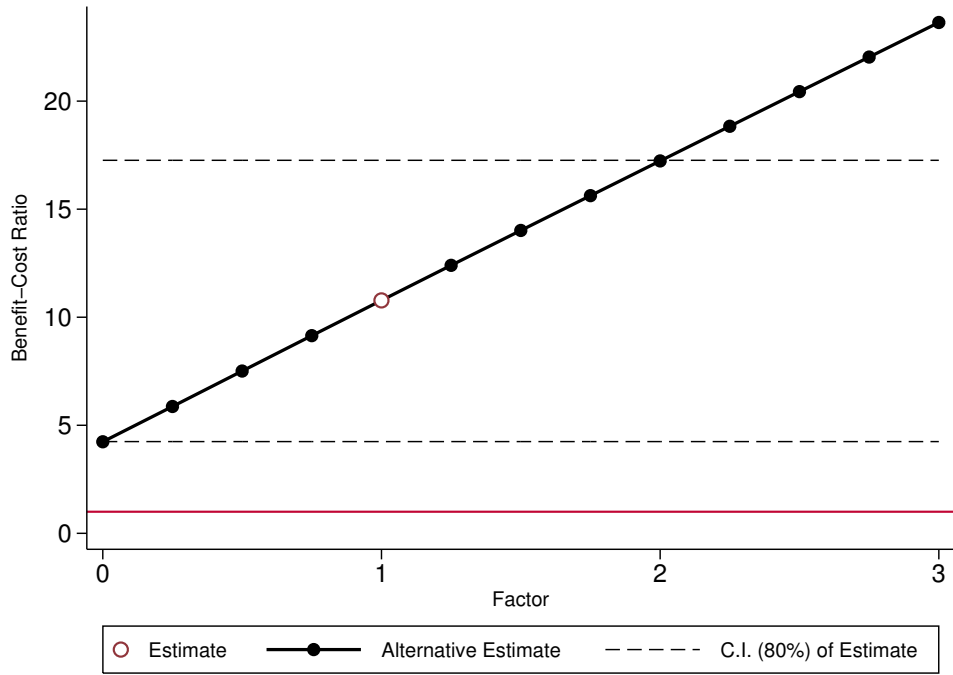
(a) Health Costs



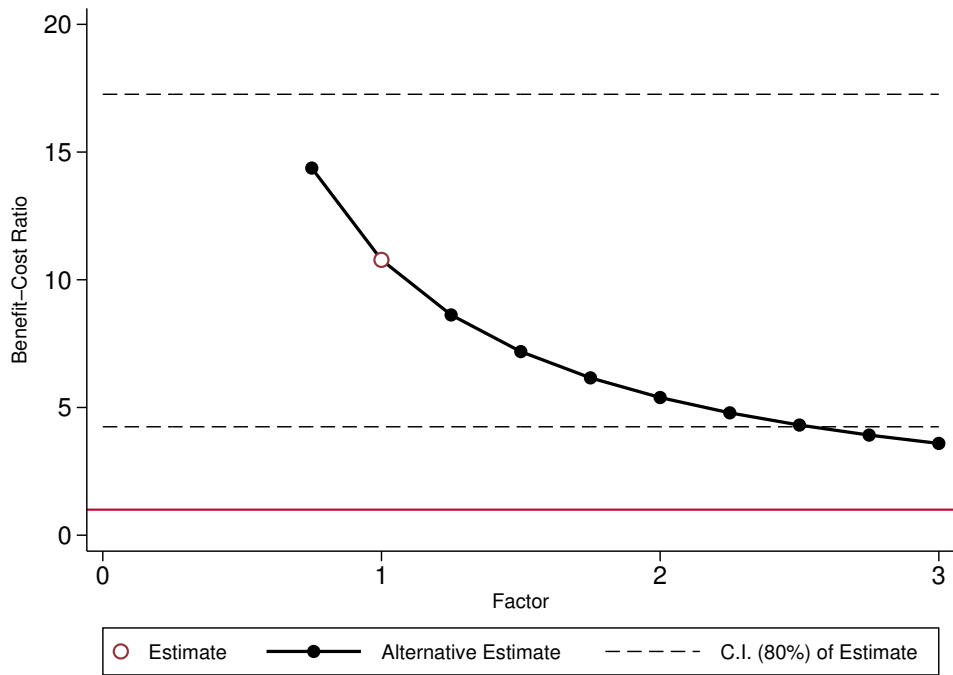
(b) Education Costs



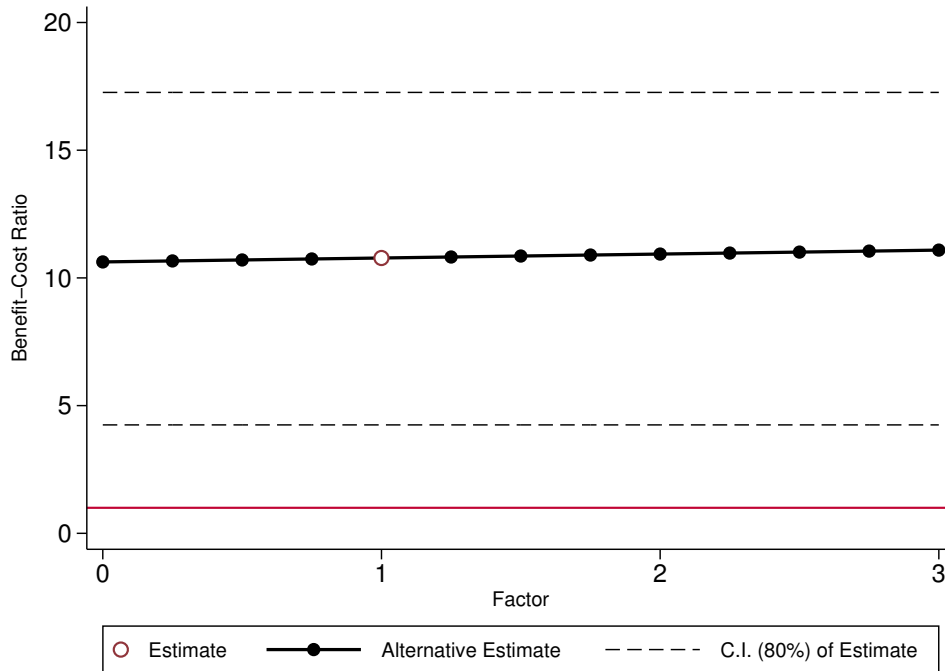
(a) Crime Costs



(b) Program Costs



(a) Control Substitution Costs



Note: These graphs display how the benefit-cost ratio changes for males as we multiply each component by a factor from 0 to 3. The red line indicates a benefit-cost ratio of 1. The hollow circle represents our actual estimates, whereas the solid dots represent the alternative estimates we obtain by varying the magnitude of each component. The benefit-cost ratio presented in the paper is equal to those presented above when the multiplicative factor is equal to 1. The estimates are means of the empirical bootstrap distribution. The 80% confidence intervals are obtained by taking the 10<sup>th</sup> and 90<sup>th</sup> quantiles of the bootstrap distribution.

Overall, we find the benefit-cost ratio to be stable across changes in the data, as well as changes in our assumptions regarding the discount rate and the marginal cost of welfare. The IRR tends to be more sensitive to changes in the data and assumptions.

## Q Back-of-the-Envelope Cost Benefit Analysis of ABC

This section presents a back-of-the-envelope cost-benefit analysis in a similar spirit to that of [Chetty et al. \(2010\)](#) and [Kline and Walters \(2015\)](#). For simplicity, we leave CARE out of this exercise. The implementation is straightforward and it is based on the argument of [Chetty et al. \(2010\)](#) that a standard deviation in a kindergarten IQ score generates an increase of 13.1% in earnings at age 27.

We perform two exercises. First, one similar to that of [Kline and Walters \(2015\)](#). We calculate the treatment effect of ABC on a kindergarten IQ test—Wechsler Preschool and Primary Scale of Intelligence at Age 5,  $\widehat{\Delta}$ , and let the (discounted to age 5) life-cycle earnings in 2014 USD dollars be \$460,793.34. We take this number straight from [Kline and Walters \(2015\)](#). Based on this, we calculate the total benefits induced by the program as:

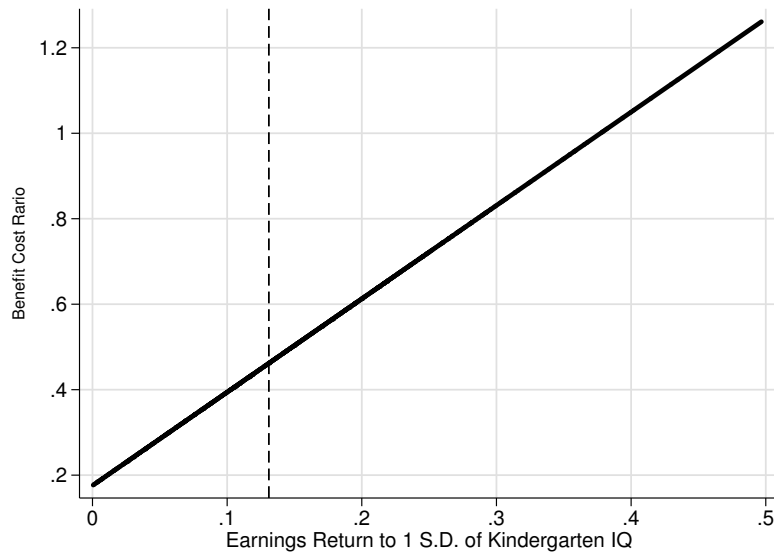
$$B = 460,793.34 \left( \widehat{\Delta} R_{IQ} N_{\text{Treated}} - N_{\text{Control}} \right), \quad (18)$$

where  $N_{\text{Treated}}$  is the number of subjects in the treatment groups and  $N_{\text{Control}}$  is the number of subjects in the control groups.  $R_{IQ}$  is the earnings return to a standard deviation increase in a kindergarten IQ test—13.1% in the case of ([Chetty et al., 2010](#)). Then, we let the total cost of the program (discounted to age 5) be \$5,403,133, as documented in Appendix I. We denote this by  $C$ .

For the program to break-even with respect to costs and benefits, a much higher return to the kindergarten IQ test is needed than the one reported in ([Chetty et al., 2010](#)), 13.1% vs. > 40%. Figure [Q.1](#) plots the relationship between the  $B/C$  and various values of  $R_{IQ}$ .

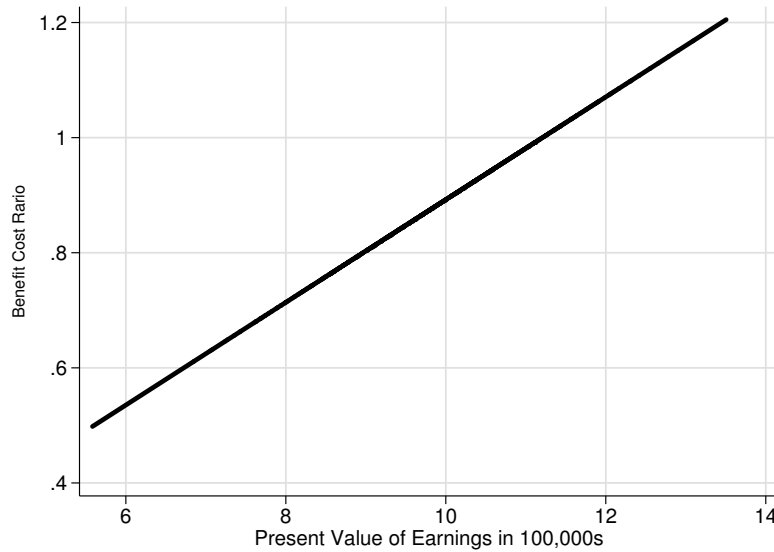
Next, we perform a similar calculation. Instead of letting  $R_{IQ}$  vary, we fix it to 13.1% and let the net present value of earnings over the life cycle vary. For the program to break even, this value needs to be beyond a million dollars (Figure [Q.2](#)). [Kline and Walters \(2015\)](#) argue that this value is \$359,064.50 for disadvantaged populations in the U.S.

Figure Q.1: Benefit-to-cost Ratio and the Earnings Returns to a Kindergarten IQ Test



Note: Benefit-to-cost ratio as a function of diverse earnings-returns to a standard deviation of a kindergarten IQ test.

Figure Q.2: Benefit-to-cost Ratio and the Discounted Net Present Value of Earnings



Note: Benefit-to-cost ratio as a function of diverse net present value of earnings.

## Appendix References

- Health affairs: Records of the Office of the Vice Chancellor. Subgroup 2: Centers, institutes and programs. Series 2: Child Development Institute. Day care center funding, 1978; 1981.
- Ad Hoc Committee of Professionals in Child Care Services, North Carolina (1974, July). Letter to Governor Holshouser.
- Administrative Branch, Office of Day Care Services (1982, December). Day care cost study. Technical report, Administrative Branch, Office of Day Care Services, Raleigh, NC.
- Angrist, J. D., G. W. Imbens, and D. Rubin (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association* 91(434), 444–455.
- Barnett, W. S. and L. N. Masse (2002). A benefit-cost analysis of the Abecedarian Early Childhood Intervention. Technical report, National Institute for Early Education Research, New Brunswick, NJ.
- Baumeister, A. A. and V. R. Bacharach (2000). Early generic educational intervention has no enduring effect on intelligence and does not prevent mental retardation: The Infant Health and Development Program. *Intelligence* 28(3), 161–192.
- Belfield, C. R., M. Nores, W. S. Barnett, and L. J. Schweinhart (2006). The High/Scope Perry Preschool Program: Cost-benefit analysis using data from the age-40 followup. *Journal of Human Resources* 41(1), 162–190.
- Besharov, D. J., P. Germanis, C. A. Higney, and D. M. Call (2011). The Abecedarian Project. In *Assessing the Evaluations of Early Childhood Education Programs*, Chapter 2. College Park, MD: Welfare Reform Academy.
- Bloom, H. S. (1984, April). Accounting for no-shows in experimental evaluation designs. *Evaluation Review* 8(2), 225–246.
- Browning, E. K. (1987, March). On the marginal welfare cost of taxation. *American Economic Review* 77(1), 11–23.
- Bryant, D. M., C. T. Ramey, J. J. Sparling, and B. H. Wasik (1987). The carolina approach to responsive education a model for day care. *Topics in Early Childhood Special Education* 7(1), 48–60.
- Buckles, K. S. and D. M. Hungerman (2013). Season of birth and later outcomes: Old questions, new answers. *Review of Economics and Statistics* 95(3), 711–724.
- Burchinal, M. R., F. A. Campbell, D. M. Bryant, B. H. Wasik, and C. T. Ramey (1997, October). Early intervention and mediating processes in cognitive performance of children of low-income African American families. *Child Development* 68(5), 935–954.
- Burchinal, M. R., M. Lee, and C. T. Ramey (1989). Type of day-care and preschool intellectual development in disadvantaged children. *Child Development* 60(1), 128–137.

- Campbell, F. A., G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. P. Pungello, and Y. Pan (2014). Early childhood investments substantially boost adult health. *Science* 343(6178), 1478–1485.
- Campbell, F. A. and C. T. Ramey (1989, April). Preschool vs. school-age intervention for disadvantaged children: Where should we put our efforts? Report, North Carolina University, Frank Porter Graham Center, Chapel Hill. Paper presented at the Biennial Meeting of the Society for Research in Child Development (Kansas City, MO, April 27–29, 1989).
- Campbell, F. A. and C. T. Ramey (1994, April). Effects of early intervention on intellectual and academic achievement: A follow-up study of children from low-income families. *Child Development* 65(2), 684–698.
- Cawley, J. (2004, Spring). The impact of obesity on wages. *Journal of Human Resources* 39(2), 451–474.
- Chetty, R., J. N. Friedman, N. Hilger, E. Saez, S. W. Diagne, and D. Yagan (2010). How does your kindergarten classroom affect your earnings? evidence from Project STAR. Working Paper 16381, NBER.
- Cohen, M. A. (1988). Pain, suffering, and jury awards: A study of the cost of crime to victims. *Law and Society Review* 22(3), 537–556.
- Cohen, M. A. and R. Bowles (2010). Estimating costs of crime. In A. R. Piquero and D. Weisburd (Eds.), *Handbook of Quantitative Criminology*, Chapter 8, pp. 143–161. New York: Springer.
- Cohen, M. A., R. T. Rust, S. Steen, and S. T. Tidd (2004, February). Willingness-to-pay for crime control programs. *Criminology* 42(1), 89–110.
- Community Planning Services (1973, August). Durham’s share in day care: A task force study. Summary report, Community Planning Services, Durham, NC.
- Courtemanche, C., J. C. Pinkston, and J. Stewart (2015, December). Adjusting body mass for measurement error with invalid validation data. *Economics & Human Biology* 19, 275–293.
- Cutler, D. M. and E. Meara (1998). The medical costs of the young and old: A forty year perspective. In D. A. Wise (Ed.), *Frontiers in the Economics of Aging*, pp. 215–242. Chicago: University of Chicago Press.
- Davis, D. E. (1977). *My friends and me*. Circle Pines, MN: American Guidance Service.
- Department of Health, Education, and Welfare (1968). Federal interagency day care requirements, pursuant to Sec. 522 (D) of the Economic Opportunity Act. Technical report, U.S. Department of Labor, Washington, DC.

- Dhanidina, L. and W. S. Griffith (1975). Costs and benefits of delayed high school completion. *Adult Education Quarterly* 25(4), 217–230.
- Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care* 35(11), 1095–1108.
- Dunn, L. M., L. T. Chun, D. C. Crowell, L. G. Dunn, L. G. Avery, and E. R. Yachel (1976). *Peabody early education kit*. Circle Pines, MN: American Guidance Service.
- Encyclopedia of Children’s Health (2016). Bayley scales of infant development. Website, <http://www.healthofchildren.com/B/Bayley-Scales-of-Infant-Development.html> (Accessed 4/19/2016).
- Feagans, L. V. (1996). *Children’s Talk in Communities and Classrooms*. Cambridge, MA: Blackwell.
- Feldstein, M. (1999, November). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics* 81(4), 674–680.
- Finkelstein, N. W. (1982). Aggression: Is it stimulated by day care? *Young Children* 37(6), 3–9.
- Frank Porter Graham Child Development Center (1979). The Frank Porter Graham Child Development Center progress report 1979. Technical report, Child Development Institute, University of North Carolina at Chapel Hill, Chapel Hill, NC.
- Freeman, R. B. (1974). Occupational training in proprietary schools and technical institutes. *The Review of Economics and Statistics* 56(3), 310–318.
- Goldman, D. P., B. Shang, J. Bhattacharya, A. M. Garber, M. Hurd, D. N. Joyce, Geoffrey F. Lakdawalla, C. Panis, and P. G. Shekelle (2005). Consequences of health trends and medical innovation for the future elderly. *Health Affairs* 24, W5–R17.
- Goldman, D. P., P. G. Shekelle, J. Bhattacharya, M. Hurd, and G. F. Joyce (2004). Health status and medical treatment of the future elderly. Technical report, RAND Corporation.
- Goldman, D. P., B. Tysinger, and D. Ermini Leaf (2015). The Future America Model: Technical Documentation. Technical report, Working Document.
- Greenberg, P. and B. Epstein (1973). *Bridges to reading*. Morristown, NJ: General Learning Corporation.
- Gross, R. T., D. Spiker, and C. W. Haynes (1997). *Helping Low Birth Weight, Premature Babies: The Infant Health and Development Program*. Stanford, CA: Stanford University Press.
- Grosse, S. D. (2008). Assessing cost-effectiveness in healthcare: History of the \$50,000 per QALY threshold. *Expert Review of Pharmacoeconomics & Outcomes Research* 8(2), 165–178.

- Haskins, R. (1985, June). Public school aggression among children with varying day-care experience. *Child Development* 56(3), 689–703.
- Haskins, R., A. M. Collier, C. T. Ramey, and P. O. Hirschbiel (1978). The effect of mild illness on habituation in the first year of life. *Journal of Pediatric Psychology* 3(3), 150–155.
- Heckman, J. J. (1979, January). Sample selection bias as a specification error. *Econometrica* 47(1), 153–162.
- Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Q. Yavitz (2010, February). The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics* 94(1–2), 114–128.
- Heckman, J. J. and J. A. Smith (1998). Evaluating the welfare state. In S. Strom (Ed.), *Econometrics and Economic Theory in the Twentieth Century: The Ragnar Frisch Centennial Symposium*, pp. 241–318. New York: Cambridge University Press.
- Heckman, J. J. and S. Urzúa (2010, May). Comparing IV with structural models: What simple IV can and cannot identify. *Journal of Econometrics* 156(1), 27–37.
- Heckman, J. J., S. Urzúa, and E. J. Vytlačil (2006). Understanding instrumental variables in models with essential heterogeneity. *Review of Economics and Statistics* 88(3), 389–432.
- Henderson, F. W., A. M. Collier, M. A. Sanyal, J. M. Watkins, D. L. Fairclough, W. A. Clyde, Jr., and F. W. Denny (1982, June). A longitudinal study of respiratory viruses and bacteria in the etiology of acute otitis media with effusion. *New England Journal of Medicine* 306(23), 1377–1383.
- Hoernack, S. A. and W. C. Weiler (1975). Cost-related tuition policies and university enrollments. *Journal of Human Resources* 10(3), 332–360.
- Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association* 47(260), 663–685.
- Johnson, J. H. (2003). The Changing Face of Poverty in North Carolina. *Popular Government* 68(3), 14–24.
- Karnes, M. B. (1973). *GOAL program: Mathematical Concepts*. Springfield, MA: Milton-Bradley.
- Klein, C. and J. Sanders (1982). Status report on programs and projects (as of January 1, 1982). Report, North Carolina University, Frank Porter Graham Center, Chapel Hill.
- Kline, P. and C. Walters (2015). Evaluating public programs with close substitutes: The case of Head Start. Working Paper 21658, National Bureau of Economic Research.
- Kuperman, S. (2014-2015). Interviews of Frances Campbell, Carrie Bynum, Phyllis Royster, Gael McGinness, Joseph Sparling, Albert Collier, Barbara Wasik, Lynne Vernon-Feagans, Tom Richey, Margaret Burchinal, Thelma Harms, and Richard Clifford. University of Chicago’s Center for the Economics of Human Development, Chicago.

- Kuperman, S. and C. Cheng (2014). Interviews of Frances Campbell, Carrie Bynum, Phyllis Royster, Gael McGinness, Joseph Sparling, Albert Collier, Barbara Wasik, Lynne Vernon-Feagans, Tom Richey, Margaret Burchinal, Thelma Harms, Richard clifford, Ron Haskins, and Susann Hutaff Haskins. University of Chicago's Center for the Economics of Human Development, Chicago.
- Kuperman, S. and A. Hojman (2015a). Interview with mary hodgers.
- Kuperman, S. and A. Hojman (2015b). Interview with Richard Clifford and Sue Russell.
- Lakdawalla, D. N., D. P. Goldman, and B. Shang (2004). The Health and Cost Consequences of Obesity Among the Future Elderly. *Health Affairs (Project Hope)* 24, W5R30–41.
- Lehmann, E. L. and J. P. Romano (2005). *Testing Statistical Hypotheses* (3 ed.). New York: Springer-Verlag.
- Lochner, L. (2011a, January). Non-production benefits of education: Crime, health, and good citizenship. Working Paper 16722, National Bureau of Economic Research.
- Lochner, L. (2011b). Nonproduction benefits of education: Crime, health, and good citizenship. In E. A. Hanushek, S. Machin, and L. Wößmann (Eds.), *Handbook of the Economics of Education*, Volume 4, Chapter 2, pp. 183–282. Amsterdam: Elsevier B. V.
- Mason, H., M. Jones-Lee, and C. Donaldson (2009). Modelling the Monetary Value of a QALY: A New Approach Based on U.K. Data. *Health Economics* 18(8), 933–950.
- McCarthy, P. and C. Snowden (1985). The bootstrap and finite population sampling. In *Vital and Health Statistics*, pp. 2–95. Washington, DC: U.S. Government Printing Office.
- McCollister, K. E., M. T. French, and H. Fang (2010). The cost of crime to society: New crime-specific estimates for policy and program evaluation. *Drug and Alcohol Dependence* 108(1–2), 98–109.
- McGinness, G. D. (1982). The language of the poverty child: Implications from center-based intervention and evaluation programs. In L. Feagans and D. Farran (Eds.), *The Language of Children Reared in Poverty: Implications for Evaluation and Intervention*, pp. 219–240. New York: Academic Press.
- McGinness, G. D. and C. T. Ramey (1981). Developing sociolinguistic competence in children. *Canadian Journal of Early Childhood Education* 1(2), 22–43.
- Miller, T. R., M. A. Cohen, and B. Wiersema (1996). *Victim Costs and Consequences: A New Look*. Washington, DC: U.S. Dept. of Justice, Office of Justice Programs, National Institute of Justice.
- Moore, S. C. (2006). The Value of Reducing Fear: An Analysis Using the European Social Survey. *Applied Economics* 38(1), 115–117.

- Moore, S. C., S. T. Mayne, B. I. Graubard, A. Schatzkin, D. Albanes, C. Schairer, R. N. Hoover, and M. F. Leitzmann (2008). Past Body Mass Index and Risk of Mortality Among Women. *International Journal of Obesity* 32(5), 730–739.
- Moore, S. C. and J. Shepard (2006). The Cost of Fear: Shadow Pricing the Intangible Costs of Crime. *Applied Economics* 38(3), 293–300.
- Noll, S. and J. Trent (Eds.) (2004). *Mental Retardation in America: A Historical Reader (The History of Disability)*. New York: NYU Press.
- North Carolina General Assembly (1971). Chapter 803, House bill 100. In *North Carolina General Assembly 1971 Session*, North Carolina. An Act to protect children through licensing of day-care facilities and other limited regulation.
- North Carolina State Department of Social Services (1972). Licensed day care facilities in North Carolina as of february 1, 1971. Raleigh, NC.
- O'Brien, C. and J. Sanders (1974). The Carolina Abecedarian Project. Frank Porter Graham Child Development Center.
- Pinto-Prades, J. L., G. Loomes, and R. Brey (2009). Trying to Estimate a Monetary Value for the QALY. *Journal of Health Economics* 28(3), 553–562.
- Prados, M. J., J. Zissimopoulos, and D. P. Goldman (2015). How Much Can an Education Intervention Reduce Lifecycle Health Disparities?
- Ramey, C. T., D. M. Bryant, J. J. Sparling, and B. H. Wasik (1985). Project CARE: A comparison of two early intervention strategies to prevent retarded development. *Topics in Early Childhood Special Education* 5(2), 12–25.
- Ramey, C. T. and F. A. Campbell (1979, February). Compensatory education for disadvantaged children. *The School Review* 87(2), 171–189.
- Ramey, C. T. and F. A. Campbell (1984). Preventive education for high-risk children: Cognitive consequences of the Carolina Abecedarian Project. *American Journal of Mental Deficiency* 88(5), 515–523.
- Ramey, C. T. and F. A. Campbell (1991). Poverty, early childhood education and academic competence: The Abecedarian experiment. In A. C. Houston (Ed.), *Children in Poverty: Child Development and Public Policy*, Chapter 8, pp. 190–221. New York: Cambridge University Press.
- Ramey, C. T., F. A. Campbell, M. Burchinal, M. L. Skinner, D. M. Gardner, and S. L. Ramey (2000). Persistent effects of early childhood education on high-risk children and their mothers. *Applied Developmental Science* 4(1), 2–14.
- Ramey, C. T., A. M. Collier, J. J. Sparling, F. A. Loda, F. A. Campbell, D. A. Ingram, and N. W. Finkelstein (1976). The Carolina Abecedarian Project: A longitudinal and multidisciplinary approach to the prevention of developmental retardation. In T. Tjossem

- (Ed.), *Intervention Strategies for High-Risk Infants and Young Children*, pp. 629–655. Baltimore, MD: University Park Press.
- Ramey, C. T. and R. Haskins (1981). The modification of intelligence through early experience. *Intelligence* 5(1), 5–19.
- Ramey, C. T., M. C. Holmberg, J. H. Sparling, and A. M. Collier (1977). An introduction to the Carolina Abecedarian Project. In B. M. Caldwell and D. J. Stedman (Eds.), *Infant Education: A Guide for Helping Handicapped Children in the First Three Years*, Chapter 7, pp. 101–121. New York: Walker and Company.
- Ramey, C. T., G. D. McGinness, L. Cross, A. M. Collier, and S. Barrie-Blackley (1982). The Abecedarian approach to social competence: Cognitive and linguistic intervention for disadvantaged preschoolers. In K. M. Borman (Ed.), *The Social Life of Children in a Changing Society*, Chapter 7, pp. 145–174. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ramey, C. T., P. Mills, F. A. Campbell, and C. O'Brien (1975, July). Infants' home environments: A comparison of high-risk families and families from the general population. *American Journal of Mental Deficiency* 80(1), 40–42.
- Ramey, C. T. and B. J. Smith (1977, January). Assessing the intellectual consequences of early intervention with high-risk infants. *American Journal of Mental Deficiency* 81(4), 318–324.
- Ramey, C. T., J. J. Sparling, and S. L. Ramey (2012). *Abecedarian: The Ideas, the Approach, and the Findings* (1 ed.). Los Altos, CA: Sociometrics Corporation.
- Robins, P. (1988). Federal support for child care: Current policies and a proposed new system. *Focus* 11(2), 1–9.
- Sanders, J. and D. B. Stokes (1979, March). Status report on programs and projects as of March 1, 1979. Report, North Carolina University, Frank Porter Graham Center, Chapel Hill.
- Sanyal, M. A., F. W. Henderson, E. C. Stempel, A. M. Collier, and F. W. Denny (1980). Effect of upper respiratory tract infection on Eustachian tube ventilatory function in the preschool child. *Journal of Pediatrics* 97(1), 11–15.
- Shaw, J. W., J. A. Johnson, and S. J. Coons (2005). Us valuation of the EQ-5D health states: Development and testing of the D1 valuation model. *Medical Care* 43(3), 203–220.
- Snyder, T. D. and S. A. Dillow (2012). *Digest of Education Statistics 2011*. National Center for Education Statistics.
- Sparling, J. and I. Lewis (1979). *Learning Games for the First Three Years : A Guide to Parent/Child Play*. New York: Walker and Company.
- Sparling, J. and I. Lewis (1984). *Learning Games for Threes and Fours*. New York: Walker and Company.

- Sparling, J. J. (1974). Synthesizing educational objectives for infant curricula. Annual Meeting of the American Educational Research Association.
- Spitz, H. H. (1992). Does the Carolina Abecedarian Early Intervention Project prevent sociocultural mental retardation? *Intelligence* 16(2), 225–237.
- Thaler, R. (1978, January). A Note on the Value of Crime Control: Evidence from the Property Market. *Journal of Urban Economics* 5(1), 137–145.
- Tong, L., M. R. Spitz, J. J. Fueger, and C. I. Amos (1996). Lung Carcinoma in Former Smokers. *Cancer* 78(5), 1004–1010.
- U.S. Department of Justice (1988). Report to the Nation on Crime and Justice.
- Vila, L. E. (2000). The Non-Monetary Benefits of Education. *European Journal of Education*, 21–32.
- Wallach, M. A. and L. Wallach (1976). *Teaching All Children to Read*. Chicago: University of Chicago Press.
- Wasik, B. H., C. Ramey, D. M. Bryant, and J. J. Sparling (1990, December). A longitudinal study of two early intervention strategies: Project CARE. *Child Development* 61(6), 1682–1696.
- Zissimopoulos, J., E. Crimmins, and P. St. Clair (2014). The Value of Delaying Alzheimer’s Disease Onset. In *Forum for Health Economics and Policy*.