Understanding the Multi-Faceted Life-Cycle Benefits of an Influential Early Childhood Program

Jorge Luis García, James J. Heckman
University of Chicago
and
Duncan Ermini Leaf, María José Prados
University of Southern California

Tel Aviv University Seminar — Tel Aviv, Israel
November 22nd, 2017

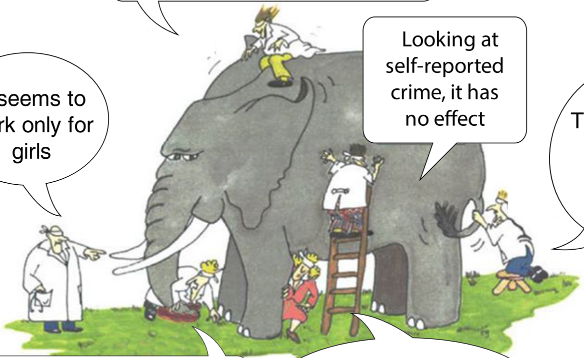
Using Economics and Econometrics to Interpret and Enhance The Evidence from Social Experiments Designed to Evaluate The Social Efficiency of Government Programs

- Use non-experimental data and structural econometric methods to supplement and expand what can be learned from the treatment effects directly generated by experiments.

- **No conflict between "causal" models and structural models.**
- Good causal models are structural models
- No particular form of data is privileged
- Use all available data and available knowledge to interpret information from social experiments

- Unaided, randomization is rarely, if ever, the “gold standard” of policy evaluation.
- Use multiple methodologies to examine all aspects of multifaceted social experiments.

Consider A Multi-Faceted Social Program: Is ABC/CARE Effective?



ABC/CARE is Effective

It supports childcare and educational attainment for mothers and is economically efficient

It works for boys and girls, but in different dimensions and is economically efficient

It reduces crime and is economically efficient

It's expensive, but has a very high rate of return

Program permits long-term follow-up - In addition, program widely replicated and motivates a variety of new initiatives

Control group substitution is informative on differential impact of low quality child-care on boys and benefits of home care

Different methodological approaches agree
The rate of return is 13.7% - it survives extensive sensitivity analysis

Framework for Benefit/Cost Rate-of-Return Analysis

- An economic evaluation of a prototypical and currently influential early childhood program that builds lasting skills
- Estimate benefit/cost ratio and rate of return of program accounting for its multiple benefits

Does this program compete favorably with other investment projects?

Empirical Challenges:

(A) Multiplicity of outcomes over the life cycle.

(B) Need metrics for evaluating multiple outcomes .

(C) Monetizing non-market benefits and costs associated with crime, health, special educational benefits, etc.

(D) Multiple hypothesis testing: avoid "cherry picking" of hypotheses.

(E) Extrapolating multiple benefits over the life cycles of agents. (Data only go through age 35, although that's much longer than most early childhood interventions.)

(F) Control group substitution and treatment group substitution.

Contributions:

(i) Economically interpretable treatment effects

(ii) Using structural economic theory (policy invariant parameters) testing policy invariance and using policy invariant empirical relationships to make forecasts.

(iii) Show danger of casual projection and extrapolation methods to estimate benefit/cost ratios currently in wide use in applied economics.

(iv) Robustness

  (a) to estimation error
  (b) to forecast error
  (c) functional form specifications
  (d) to assumptions made about non-market costs and benefits

(v) Understanding the opportunities and limitations created by *substitution bias* in randomized trials:

- What causal analysts call "threats to validity" offer opportunities for learning more deeply about the program being evaluated.

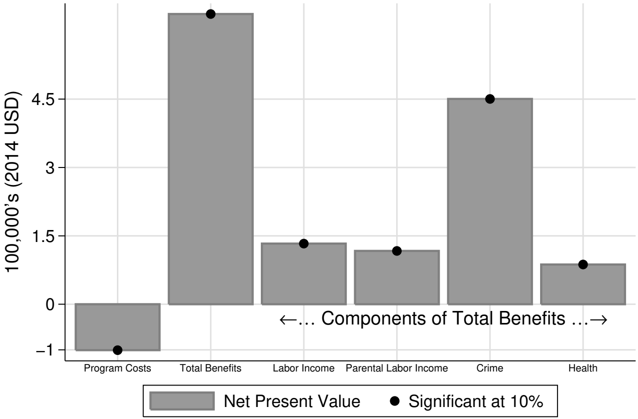
Findings: An Overview

- Program has substantial benefits on
  (a) Health, the quality of life and healthy behaviors
  (b) Earnings of participants
  (c) Crime
  (d) Education
  (e) Earnings, employment and education of **mothers** through providing childcare

Findings: An Overview

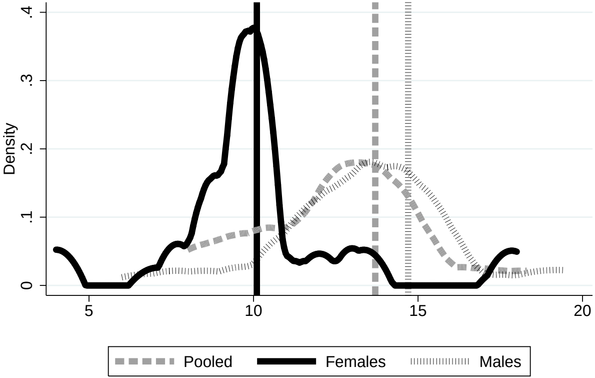
- The overall rate of return: statistically significant 13% per annum with a benefit/cost ratio of 6.3, even after accounting for the welfare costs of taxation to finance the intervention.
- Range of estimates from extensive sensitivity analysis:
  - Estimates of overall B/C ratio range: (1.52, 17.40)
  - Estimates of overall IRR range: (8.0%, 18.3%)

Figure 1: Median Net Present Value of Main Components of the Benefit/Cost Analysis Over the Life Cycle per Program Participant, Treatment vs. Next Bests



Per−annum Rate of Return: Males and Females 13.7% (s.e. 3%).
Benefit−cost Ratio: Males and Females 7.3 (s.e. 1.8).

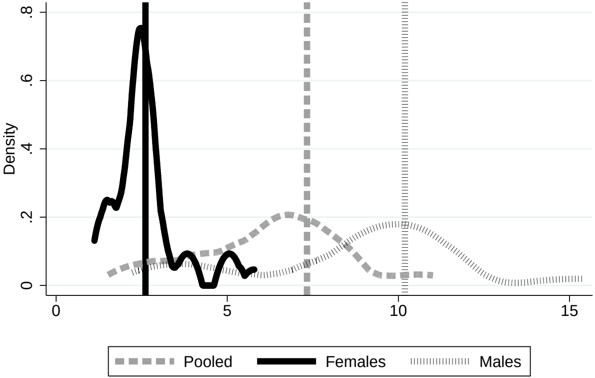
Figure 2: Range of Estimates for Benefit/Cost Ratio



The vertical line represents the baseline estimate.
Average. Pooled: 12.64. Females: 10.07. Males: 13.34.
Median. Pooled: 13.00. Females: 10.00. Males: 13.70.

Note: This figure shows the empirical distribution of all the estimates of the benefit/cost ratio that we provide in the paper, and vary as a function of a wide set of assumptions about the data generating process and the parameters we take from other literature.

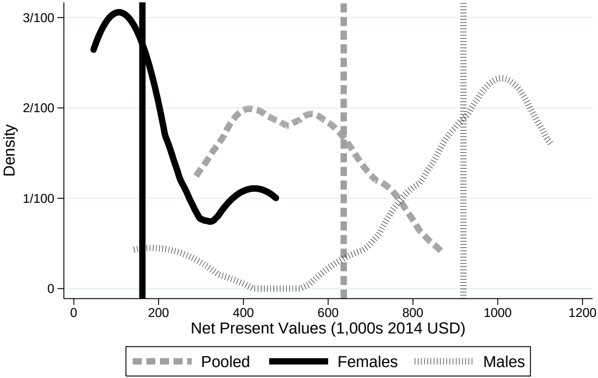
Figure 3: Range of Estimates for Internal Rate of Return



The vertical line represents the baseline estimate.
Average. Pooled: 6.90. Females: 2.63. Males: 9.67.
Median. Pooled: 6.75. Females: 2.49. Males: 9.85.

Note: This figure shows the empirical distribution of all the estimates of the internal rate of return that we provide in the paper, and vary as a function of a wide set of assumptions about the data generating process and the parameters we take from other literature.

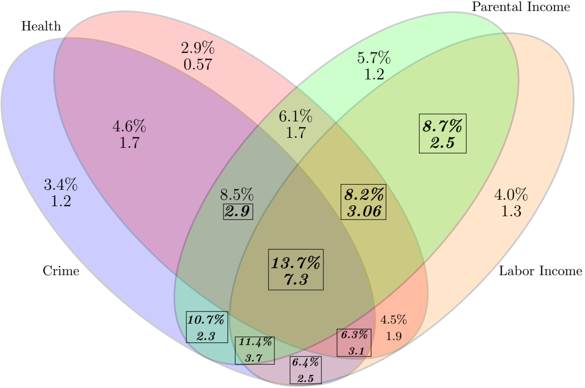
Figure 4: Range of Estimates for Net Present Value



The vertical line represents the baseline estimate.
Average. Pooled: 517,608. Females: 190,627. Males: 872,593.
Median. Pooled: 509,466. Females: 129,260. Males: 966,735.

Note: This figure shows the empirical distribution of all the estimates of the net present value that we calculate. They vary as a function of a wide set of assumptions about the data generating process and the parameters we take from other literature.

Figure 5: Benefit/Cost Ratio and Internal Rate of Return when Accounting for Different Combinations of the Main Benefits



Note: This figure presents all possible combinations of accounting for the benefits from the four major categories in our analysis. The non-overlapping areas present estimates accounting for a single category as the benefit. When two categories overlap, these are the benefits that we account for. And so on for the rest of the plot. The costs remain constant across all calculations and are the same as in Figure 13. Inference is based on non-parametric, one-sided p-values from the empirical bootstrap distribution. We bold point estimates significant at the 10% level.

Findings differ by gender:

(a) More beneficial treatment effects across a variety of domains for girls.
(b) More substantial *monetized* benefits for boys.
(c) Substantial benefits for health and Quality of Adjusted Life Years (QALY, primarily for men)
(d) Childcare for mothers: facilitates work, promotes earnings and educational attainment
(e) Crime (for men)
(f) Earnings gains (both)
(g) Education (primarily for women)
(h) Lasting effects on IQ through age 21 (stronger and persistent for women)

Program Analyzed

- ABC/CARE Experiment
  (a) Two very similar programs launched in the early 1970s that have long-term follow-ups through age 34
  (b) Starts early (age 8 weeks)/intensive (8 hours a day) through age 5
  (c) A second stage through age 8 that gave home visits but that had little effect on participant outcomes
  (d) Focus on the 8 weeks–age 5 segment

Link to Program Details

Link to ABC/CARE Tables

No Second Stage Effects

Link to A.6

Goals of the Program

- Goals: to remediate lifetime disadvantage by fostering the early-life skills of at-risk populations.
- Full-day curriculum: emphasized active learning experiences, dramatic play, pre-academic skills (simple concepts of order, ranking, organization), and language skills.
- For ages 3 through 5, as the cohorts approached public school entry, classroom experiences were increasingly structured towards the development of academic skills and "socio-linguistic and communicative competence."
- Access to health screenings (but not costs of medication and costs of medical procedures)

Subsidized Childcare

- For both ABC and CARE, centers were open to families from 7:45 a.m. to 5:30 p.m., 5 days per week, 50 weeks per year.
- Subjects offered free transportation to and from the center.
- Subsidized childcare:
  - (i) Promotes wage growth of women through work experience
  - (ii) Promotes their educational attainment

Program Relevant Today

- Many programs use ABC/CARE as prototype
  - Infant Health and Development Program (IHDP)—eight different cities around the U.S. (Spiker et al., 1997);
  - Early Head Start and Head Start in the U.S. (Schneider and McDonald, 2007);
  - John's Hopkins Cerebral Palsy Study in the U.S. (Sparling, 2010);
  - CLIO study in the U.S. (Sparling, 2010);
  - Massachusetts Family Child Care Study (Collins et al., 2010);
  - Healthy Child Manitoba Evaluation (Healthy Child Manitoba, 2015);
  - Abecedarian Approach within an Innovative Implementation Framework (Jensen and Nielsen, 2016);
  - Building a Bridge into Preschool in Remote Northern Territory Communities in Australia (Scull et al., 2015).
  - Educare programs are based on ABC/CARE (Educare, 2014; Yazejian and Bryant, 2012).
  - More than 20 Educare programs have been launched all over the U.S.

Link to Data Availability for ABC and CARE

Link to Tables of Full Experimental Data

ABC First-Phase Randomization Compromises and Attrition

- Treatment group (58); Control group (56)
- Data on non-compliers available up to age 8
- **Asymptotic theory valid (see Campbell et al., 2014)**
- Analyses accounting for non-compliance in estimating early treatment effects indicate insensitivity to their exclusion or inclusion
- Account for attrition using inverse probability weighting:
  - Use conditional independence (Horvitz and Thompson, 1952)

Virtually all people offered treatment accepted it.

- All people offered the program agreed to participate
- Non-compliance not an issue for treatment arm
- Attrition related to health (primarily child death) and parental relocation

Defining Treatment Effects: Framework and Notation

Panel Data Structure

$$\overbrace{\underbrace{[0, \ldots, \bar{a}\,]}_{\text{Treatment Period}} \quad \underbrace{(\bar{a}, \ldots, a^{*}]}_{\text{Follow-up}}}^{\text{Our Sample}} \quad \underbrace{(a^{*}, \ldots, A]}_{\text{Remainder of Life Cycle}}$$

- Individuals eligible to participate in the program if $B \in \mathcal{B}_0$.
- $\mathcal{B}_0$: set of scores on the high risk index (HRI).

Link to Appendix:
Determinants of High Risk Index

- $W = 1$: parents wish to participate in the program.
- $R = 1$ indicates that a child is randomized to be able to participate in the program.
- $R \in \{0, 1\}$.
- D: participates in the program ($D \in \{0, 1\}$).
- $D = RW$

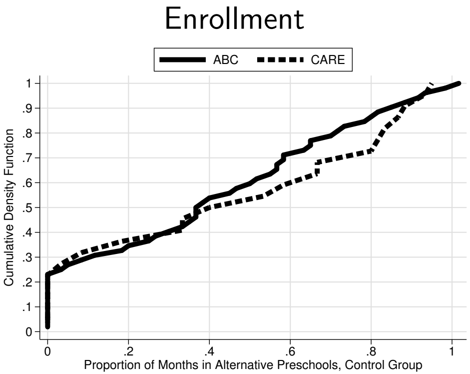
- All eligible families ($B \in \mathcal{B}_0$) given the option to participate ($R = 1$) wish to participate in the program ($D = 1$).
- Full compliance with the randomization: $R = 1 \Rightarrow W = 1$.
- *Ex ante* parents perceive that ABC/CARE was superior to other childcare alternatives.
- *Can safely interpret the treatment effects generated by the experiment as average treatment effects for the population for which $B \in \mathcal{B}_0$ and not just treatment effects for the treated (***TOT***).*

- Y_a^1: outcome vector with treatment.
- Y_a^0: outcome vector without treatment.

Exposure to Treatment

- **Control-group substitution** or **substitution bias** (control group takes treatment)

Figure 5: Months in Alternative Preschools (Control Group)



"Threat to Validity"?

- A problem to be avoided?
- Or a source of information?

Control Group Substitution: Simplifications

- Sample sizes too small to make credible estimates of detailed control exposures.
- Simplify analysis: create two categories of control status.
- "H": child in home care throughout the entire length of the program.
- "C": child in an alternative preschool anytime.

$Y^0_{a,H}$: **Control subject received home care exclusively**

$Y^0_{a,C}$: **Control subject received some alternative childcare**.

- $V \in \{0, 1\}$.
- $V = 1$: participation.
- $V = 0$: staying at home.
- Control outcome (Y_a^0):

$$Y_a^0 := (1 - V)\, Y_{a,H}^0 + (V)\, Y_{a,C}^0. \tag{1}$$

Parameters Estimated

- Effect of program compared to the next best alternative as perceived by the parents:

$$\mathbf{\Delta}_a := \mathbb{E}\left[Y_a^1 - Y_a^0 | W = 1\right] := \mathbb{E}\left[Y_a^1 - Y_a^0 | B \in \mathcal{B}_0\right] \qquad (2)$$

- LATE parameter identified by experiment.

The Experimental Randomization Does Not Directly Identify Other Parameters of Interest

- Effectiveness of the program with respect to a counterfactual world in which the child stays at home full time:

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

- Effectiveness of a program relative to attendance in an alternative preschool for those who would choose an alternative:

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

THE UNIVERSITY OF CHICAGO

- **For this randomization, need non-experimental methods to identify these parameters.**

How Best to Summarize and Interpret Multiple Treatment Effects?

Four Approaches

(1) Step-down multiple hypothesis testing by blocks of interpretable outcomes

(2) Form factors over blocks of outcomes; test if treatment factors $>$ control factors

(3) Combining functions within and across blocks

(4) Benefit/Cost and Rate of Return analyses

(1) Step-down Block Analysis

- Males more likely to be employed
- Receive higher labor income.
- Have fewer felony arrests
- Better health
- Females are more likely to graduate from high school and achieve more years of schooling.
- Positive effects on their employment.
- Not as strong effects on female labor income as on male labor income.
- Fewer felony arrests.

Link to Treatment Effects

Link to Appendix:
Treatment Effects Accounting Correcting the p-values Using Step-down (Romano-Wolf; Romano-Shaikh)

(2) Factors by Block

- Agree with method (1)

(3) Combining Functions

- Block of "similar" N_l outcomes: indexed by set $Q_l = \{1, \ldots, N_l\}$
- $j \in Q_l$ outcome; within block l
- Associated mean treatment effect:

$$\Delta_{j,a} := \mathbb{E}(Y^1_{j,a} - Y^0_{j,a} | B \in \mathcal{B}_0) \tag{5}$$

- Assume outcomes can be ordered so that $\Delta_{j,a} > 0$ is beneficial (all but 5% are so ordered).
- Summarize estimated beneficial effects of the program on outcomes within block l:

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

- **Proportion of beneficial outcomes in block l: C_l/N_l.**

THE UNIVERSITY OF CHICAGO

- $\mathcal{L}$: number of blocks
- Assume the validity of asymptotic approximations
- C_l/N_l should be centered around $1/2$ under the null hypothesis of no treatment effects for all $j \in Q_l, l \in \mathcal{L}$
- Bootstrap to obtain p-values for the null hypothesis of no treatment effect for each block and over all blocks

- **Alternative: Count "significant" results**.
- For a 10% significance level, on average 10% of all outcomes should be "significant" at the 10% level even if there is no treatment effect of the program.

- Combining counts across all blocks enables us to avoid
  (i) Arbitrarily picking outcomes that have statistically significant effects—"cherry picking"
  (ii) Arbitrarily selecting blocks of outcomes to correct the p-values when accounting for multiple hypothesis testing.

Estimated Combining Functions

Pronounced Gender Differences

Figure 6: Percentage of Outcomes with a Positive Treatment Effect

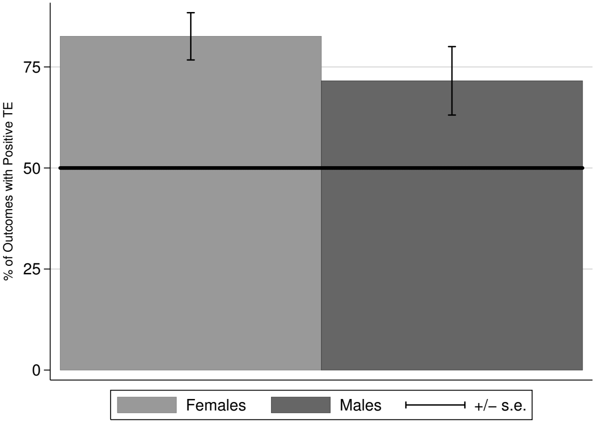


Figure 7: Percentage of Outcomes with Positive and Significant Treatment Effects, at 10% Level (Treatment vs. Next Best)

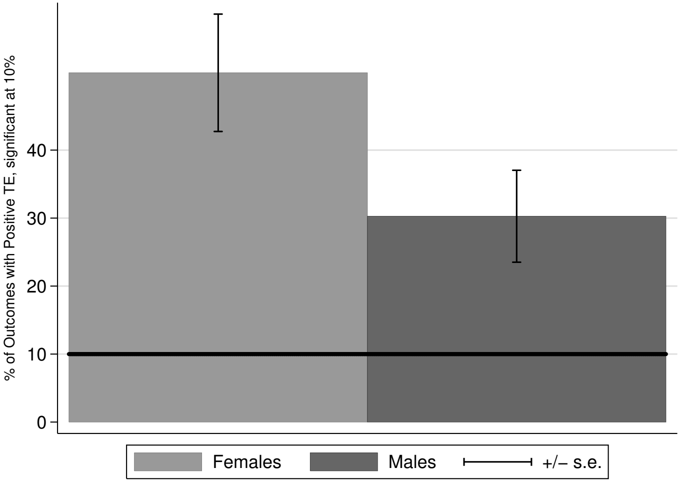
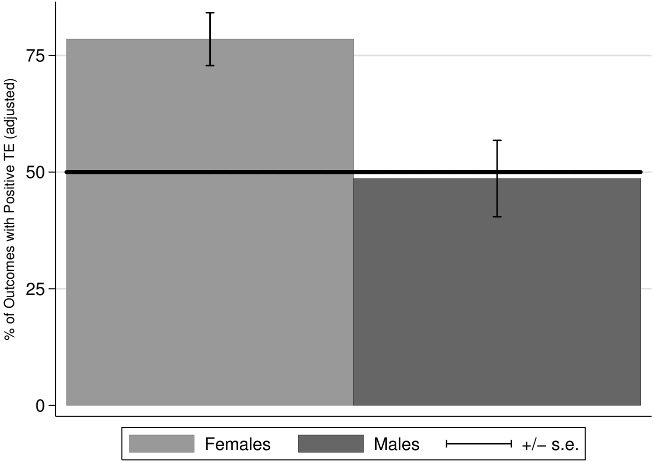


Figure 8: Percentage of Outcomes with a Positive Treatment Effect Fixing Control Group to Stay at Home



UNIVERSITY OF CHICAGO

Estimated Combining Functions by Outcome Categories

Gender Differences Vary by Category

Main findings:

- Control males do better than control females
- This weakens and reverses for treatment group
- Weak or non-existent benefits for treatment vs stay-at-home for boys
- Strong treatment effects for treatment vs low quality formal childcare
- Explains the analysis of Baker et al. (2005) and Kottelenberg and Lehrer (2014): Low quality childcare can harm children; boys especially vulnerable

Predicting and Monetizing Life-cycle Costs and Benefits

The Forecasting Problem

- Data on control- and treatment-group members through age a^*.
- Post-a^* treatment effects required to construct counterfactual life-cycle profiles.
- Making valid predictions of out-of-sample treatment effects does not require making valid predictions of separate out-of-sample treatment and control profiles.
- Only valid predictions of treatment effects is required, not components of them.

Using Auxiliary Data Sources to Predict Out-of-Sample Outcomes

Empirical Bayes Methods Applied to Combined Multiple Samples

Structural Forecasting

- No available direct source of information on future outcomes for experimental group
- Predict post-sample life cycle benefits using synthetic cohorts.
- Groups comparable to experimental baselines that are not exposed to treatment.
- **Big Issue: Are relationships between experimental outcomes and later life variables from non- experimental samples like those that would have arisen from the experimental sample had it been followed up over its lifetime?**

Example

- Education-income relationship in the non-experimental sample different from what is in experiment because of different sources of exogenous variation in education in experiment.
- Experiment gives an ability-constant relationship.
- Non-experimental data: need to control for unobservables to align with experimental data.
- Need to align samples (experimental and synthetic cohorts): (common support; common stochastic relationships – experimental and synthetic cohorts).

- **Requirements for Structural Forecasting (Sufficient Conditions)**

  (i) Common variables

  (ii) Common supports

  (iii) Common stochastic processes: relationship between markers in the data (e.g., test scores) and outcomes (earnings) different (experiment induces exogenous variation in the marker-outcome relationship, not present in non-experimental data.

THE UNIVERSITY OF CHICAGO

Using structural economic models to assess the validity of our forecasting equations distinguishes our analysis from large body of work on "surrogate markers" in biostatistics that seeks to forecast long-run outcomes from short-run markers (e.g., Prentice, 1989; Zheng and Little, 2003)

- How to forecast treatment effects for non-treated sample.
- Start from and extend the analysis of Heckman et al. (2013).
- **Key Assumption: The effect of treatment on outcomes operates through its effects on inputs creating outputs for a stable production function.**
- Structural relationships estimated in the non-experimental sample are the basis for principled extrapolation.

Link to Data Sources for Cost/Benefit Analysis

Analytical Framework

- Causal model for treatment ($d = 1$) and control ($d = 0$)
- Outcome j, in sample k, at age a, for measure j at age a in sample $k \in \{e, n\}$
- e: membership in the experimental sample
- n: membership in the auxiliary sample:

$$Y^d_{k,j,a} = \phi^d_{k,j,a}(X^d_{k,a}, B_k) + \varepsilon^d_{k,j,a}, \quad j \in \mathcal{J}_a. \tag{7}$$

- $\phi^{d}_{k,j,a}(\cdot,\cdot)$: **policy-invariant structural production function**.
- $E(\varepsilon^{d}_{k,j,a}) = 0$
- $X^{d}_{k,a}$ includes variables caused by treatment.
- Y^{d}_{k}: all outcomes $a \in \{1,\ldots,A\}$ for $k \in \{e,n\}$, when treatment status is fixed to d.
- X^{d}_{k}: vector of all causal predictors of Y^{d}_{k} at all ages.
- Background variables B may have different distributions in the two samples.

- Joint distribution conditional on $B_k = b$: $F_{Y_k^d, X_k^d | B_k = b}(\cdot, \cdot)$.
- Can use D_e and R_e interchangeably because of full compliance.
- Standard **Quandt (1972) switching regression** model:

$$\begin{aligned} Y_{k,j,a} &= (1 - D_k)\, Y^0_{k,j,a} + (D_k)\, Y^1_{k,j,a}, \\ & \qquad j \in \mathcal{J}_a, a \in \{1, \ldots, A\}, \quad k \in \{e, n\} \\ X_{k,a} &= (1 - D_k)\, X^0_{k,a} + (D_k)\, X^1_{k,a}. \end{aligned} \tag{8}$$

Age, Period, and Cohort Effects

- $Y^d_{j,k,a,c,t}$: outcome j for sample k at age a for birth cohort c at time t when treatment is fixed to d.

Assumption A–1

Alignment of Cohort and Time Effects

For experimental sample cohort c_e and auxiliary sample cohort c_n:

$$Y^d_{e,a,c_e,t_e} = Y^d_{n,a,c_n,t_n} \tag{9}$$

for $d \in \{0,1\}$, $a \geq a^*$, where t_e, t_n are the years for which cohorts c_e, c_n are observed, where $t_e = t_n + c_e - c_n$, and t_n is the year that the age a outcome is observed for cohort n $(t_n = a + c_n)$. $\square$

Support Conditions

Assumption A–2

Support Conditions

For $a \in \{1, \ldots, A\}$, *the support of* $\left(Y^d_{e,a}, X^d_{e,a}, B_e\right)$ *in the experimental sample is contained in the support of* $\left(Y^d_{n,t}, X^d_{n,t}, B_e\right)$ *in the auxiliary sample:*

$$supp(Y_{e,a}, X^d_{e,a}, B_e) \subseteq supp(Y_{n,a}, X^d_{n,a}, B_n), \quad d \in \{0, 1\}. \quad \square \tag{10}$$

- Testable for ages $a \leq a^*$.
- Condition satisfied in our samples.

Link to Appendix:
Common Support between Experimental and Non-Experimental Samples

Conditions for Valid Out-of-Sample Predictions

I. Strong Sufficient Condition

Condition C–1

Equality of Distributions Across the Experimental and Auxiliary Samples

$$F_{Y_e^d, X_e^d | B_e = b}(\cdot, \cdot) = F_{Y_n^d, X_n^d | B_n = b}(\cdot, \cdot), \quad d \in \{0, 1\} \tag{11}$$

for $Y_e^d, X_e^d | B_e = b$ and $Y_n^d, X_n^d | B_n = b$ contained the support of the experimental sample supp $\left(Y_e^d, X_e^d, B\right)$.

- Assumption C-1 standard assumption in "surrogate marker" literature (see Prentice, 1989; Little 2004)
- Ignores the fundamental distinction between experimental variation and non-experimental sample variation
- Assumes common stochastic relationships in experimental and auxiliary samples
- Does not investigate the mechanisms that make C-1 plausible

II. Weaker

Condition C–2

Equality in Conditional Expectations Across the Experimental and Auxiliary Samples

$$\mathbb{E}\left[Y_e^d | X_e^d = x, B_e = b\right] = \mathbb{E}\left[Y_n^d | X_n^d = x, B_n = b\right], \quad d \in \{0,1\} \tag{12}$$

for $d \in \{0,1\}$ *over supp* $\left(Y_{e,a}^d, X_{e,a}^d, B_e\right)$.

THE UNIVERSITY OF CHICAGO

III. Necessary and Sufficient (For Means)

Condition C–3

Equality in Mean Treatment Effects Across the Experimental and Auxiliary Samples

$$\mathbb{E}\left[Y_e^1 - Y_e^0 | B_e = b\right] = \mathbb{E}\left[Y_n^1 - Y_n^0 | B_n = b\right] \tag{13}$$

over supp $\left(Y_{e,a}^d, B_e\right)$.

- Minimal requirement for valid age by age forecasts (necessary and sufficient)
- Testable (for $a \leq a^*$)
- Not testable(for $a > a^*$) without further assumptions.

THE UNIVERSITY OF CHICAGO

- Could simply invoke Condition C–2 or C–3 and be done.
- **Our approach:** empirically examine Condition C–2 and test (when possible) assumptions that justify it.

Exogeneity of Regressors: A Convenience, Not a Necessity

Sufficient Condition for Exogeneity in Both Samples

Assumption A–3

Exogeneity

For all $a, a'' \in \{1, \ldots, A\}$ *and for* $d, d' \in \{0, 1\}$,

$$\varepsilon^{d}_{k,j,a} \perp\!\!\!\perp X^{d'}_{k,a''} | B_k = b \tag{14}$$

for all b *in the support of* B_k, $k \in \{e, n\}$, *for all outcomes* $j \in \mathcal{J}_a$, *where "*$M \perp\!\!\!\perp N|Q$*" denotes independence of* M *and* N *given* Q. $\square$

- Exogeneity facilitates the use of economic theory to generate and interpret treatment effects, to test the validity of our synthetic control groups, and to find auxiliary sample counterparts to treatments and controls.
- **Not essential to our approach.**
- Recognizing and responding to the distinction between the data generation processes in the experiment and in auxiliary samples is a contribution of this paper.
- We relax and test it using a variety of econometric specifications.

Link to Appendix:

Tests of Exogeneity and Impact on Forecasts of Using Different Models

Link to Appendix:

Exploring the Impact of Using Different Prediction Models

Structural Invariance

Assumption A–4

Structural Invariance

For all $x, b \in supp(X^d_{e,a}, B_e), k \in \{e, n\}$

$$\begin{aligned} \phi^0_{k,j,a}(x, b) &= \phi^1_{k,j,a}(x, b) \\ &=: \phi_{j,a}(x, b), \end{aligned} \tag{15}$$

$\phi^d_{k,j,a}(x)$ *is the function generating the causal effect of setting* $X^d_{k,a} = x$ *holding* $\varepsilon^d_{k,j,a}$ *fixed for* $a \in \{1, \ldots, A\}$ *for any outcome* $j \in \mathcal{J}_a$. $\square$

- (Frisch, 1938, Marschak, 1953, Hurwicz, 1962)

Messages of A–4:

(i) Structural functions evaluated at the same arguments have identical values for treatment and control groups in the experimental sample.

(ii) Structural relationships are identical in the experimental and auxiliary samples.

(iii) Requires:

  (a) Common inputs
  (b) Common supports
  (c) Common stochastic structure

Testable Implications

Alignment of Conditional (on X) Treatment and Control Groups

$$\mathbb{E}\left[Y^d_{e,j,a}|X^d_{e,a} = x, B_e = b, D = d\right] \tag{16}$$
$$= \mathbb{E}\left[Y_{n,j,a}|X_{n,a} = x, B_n = b\right] \text{ for } d \in \{0,1\}$$

- Relationship (16) testable for $a \leq a^*$, when $Y_{k,j,a}$ is observed in both the experimental and auxiliary samples, $k \in \{e, n\}$.

Alignment of Predicted and Within-Sample Treatment Effects (Across Experimental and Auxiliary Samples)

$$\mathbb{E}\left[Y^1_{e,j,a}|X^1_{e,a} = x^1, B_e = b, D = 1\right] - \mathbb{E}\left[Y^0_{e,j,a}|X^0_{e,a} = x^0, B_e = b, D = 0\right] \tag{17}$$

$$= \mathbb{E}\left[Y_{n,j,a}|X_{n,a} = x^1, B_n = b\right] - \mathbb{E}\left[Y_{n,j,a}|X_{n,a} = x^0, B_n = b\right].$$

Within Experiment:
Adjusted Treatment and Controls Equated

$$\mathbb{E}\left[Y^1_{e,j,a} | X^1_{e,a} = x, B_e = b, D = 1\right] = \mathbb{E}\left[Y^0_{e,j,a} | X^0_{e,a} = x, B_e = b, D = 0\right] \tag{18}$$

- We test and do not reject all three hypotheses (16)-18 singly and jointly within the experimental sample age range ($a \leq a^*$).

Tests

- In the experimental sample all of the parents of children with characteristics $B \in \mathcal{B}_0$ agree to participate in the program.
- Because the auxiliary samples have no treatment group members, we can evaluate our procedure by comparing the labor incomes of individuals in the auxiliary samples for whom $B \in \mathcal{B}_0$ to the labor incomes of individuals in our constructed synthetic **control** group.

Figure 9: Labor Income Profile, Disadvantaged Individuals Synthetic Control Group in the Auxiliary Samples, Females

(a)

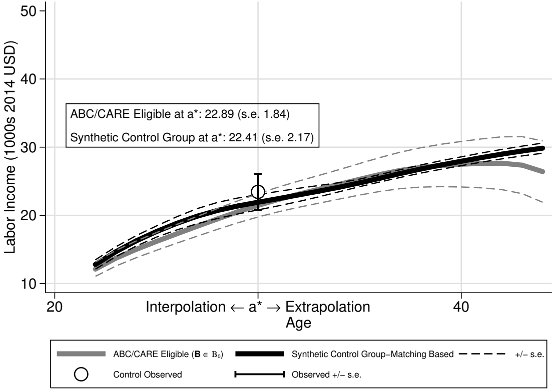
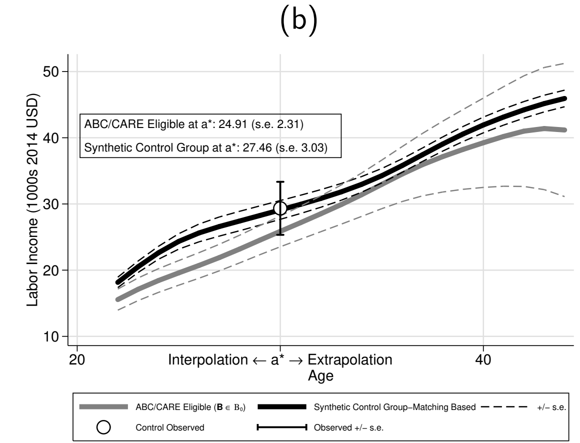


Figure 9: Labor Income Profile, Disadvantaged Individuals Synthetic Control Group in the Auxiliary Samples, Males



Link to Appendix:

Tests of Implications

Theorem 1

Valid Out-of-Sample Predictions

Under Assumptions A–1-A–4, Conditions C–1 through C–3 hold for any value of $\left(X_{k,a}^{d}, B_{k}\right)$.

This is an immediate consequence of the cited assumptions. $\square$

Implementation: Five-Step Procedure

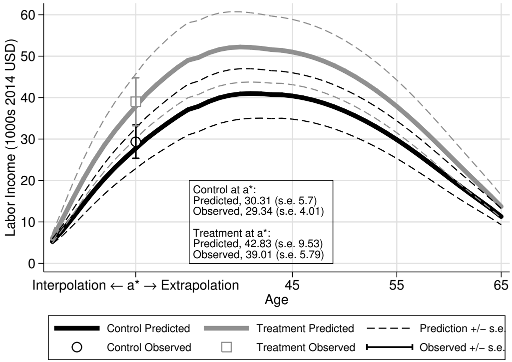
- Step 1: Use experimental sample to conduct structural mediation analyses relating the vector of outcomes at age a for person i ($Y^d_{i,a}$) for $a \leq a^*$ to predictor variables (and interactions) that are affected by treatment ($X^d_{i,a}$), as well as background variables (B_i).
- Step 2: Construct counterpart predictions of treatment and control outcomes using the auxiliary samples.
- Step 3: Use the estimated dynamic relationships fit on the constructed samples to predict the post-a^* outcomes.
- Step 4: Explore a variety of alternative assumptions on the data generation process and use these to present alternative forecasts as a form of sensitivity analysis.
- Step 5: Bootstrap or sub-sample bootstrap all empirical procedures to account for estimation error and model search uncertainty (for pre-testing, use bootstrap subsampling)

Link to Appendix:

Inference

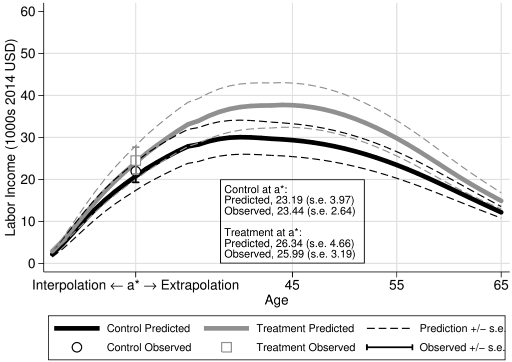
Example

Figure 10: Predicted Labor Income Profiles for ABC/CARE Participants, Males



Note: This figure displays the predicted life-cycle labor income profiles for ABC/CARE males by treatment status, based on the method proposed in this section. We combine data from the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey of Youth 1979 (NLSY79), and the Children of the National Longitudinal Survey of Youth 1979 (CNLSY79). We highlight the *observed* labor income at a^* (age 30) for the ABC/CARE control- and treatment-group participants.

Figure 10: Predicted Labor Income Profiles for ABC/CARE Participants, Females



Note: This figure displays the analogous figure for females. Our predictions go up to age 67, age of assumed retirement. Standard errors are based on the empirical bootstrap distribution. We combine data from the Panel Study of Income Dynamics (PSID), the National Longitudinal Survey of Youth 1979 (NLSY79), and the Children of the National Longitudinal Survey of Youth 1979 (CNLSY79). We highlight the *observed* labor income at a^* (age 30) for the ABC/CARE control- and treatment-group participants.

Link to Appendix:
Estimation Procedure and Data Combination Estimator in the GMM Framework

Alternative Intuitive Approach: Matching

- Under exogeneity of X, matching constructs valid control and treatment groups nonparametrically.
- Potential danger: endogeneity in the auxiliary samples.
- Results very comparable with estimates from comparable approaches.
- Implicitly a joint test of exogeneity and robustness to functional form.

Link to Appendix:
Using Matching to Construct Virtual Treatment and Comparison Groups

Health

- Future America Model (FAM): Dynamic Competing Risks Model
- Projects health outcomes from the subjects' mid-30s up to their projected death (Goldman et al., 2015)

Figure 11: Abecedarian Project, Health Effects at Age 34 (Males)

	Treatment Mean	Control Mean	Treatment p-value
Systolic Blood Pressure	125.79	143.33	0.018
Diastolic Blood Pressure	78.53	92.00	0.024
Pre-Hypertension	0.68	0.78	0.235
Hypertension	0.10	0.44	0.011
HDL Cholesterol	53.21	42.00	0.067
Cholesterol/HDL-C	3.89	4.69	0.057
Abdominal Obesity	065	0.87	0.136
Metabolic Syndrome	0.00	0.25	0.009

Source: Campbell, Conti, Heckman, Moon, Pinto, Pungello, and Pan (2014).

Five Steps to Estimate Competing Risks Model

(1) Estimate age-by-age health state transition probabilities using the Panel Study of Income Dynamics (PSID);
(2) Match these transition probabilities to the ABC/CARE subjects based on observed characteristics;
(3) Estimate quality-adjusted life year (QALY) models using the Medical Expenditure Panel Survey (MEPS) and the PSID;
(4) Estimate medical cost models using the MEPS and the Medicare Current Beneficiary Survey (MCBS), allowing estimates to differ by health state and observed characteristics;
(5) Predict the medical expenditure and QALYs that correspond to the simulated individual health trajectories
(6) First Order Markov
(7) For sensitivity analysis, we also fit Second Order Markov models.

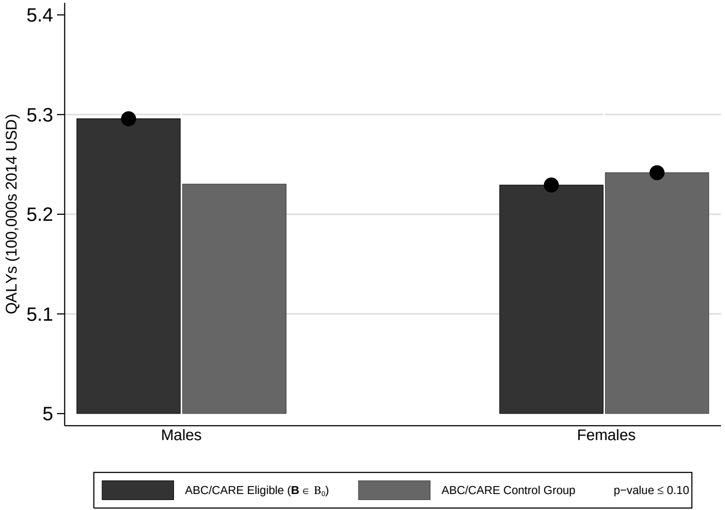
Table 1: Health State Transitions, Age a as Predictor of Age $a+1$

Age a	Age $a+1$						
	Heart Disease	Hyper-tension	Stroke	Lung Disease	Diabetes	Cancer	Disability
Heart Disease			×				×
Hypertension	×		×				×
Stroke							×
Lung Disease							×
Diabetes	×	×	×				×
Cancer			×				×
Disability							×
Smoking	×	×	×	×	×	×	×
BMI	×	×	×	×	×	×	×
Physical Activ.	×	×	×	×	×	×	×
Binge Drinking							
DI Claim							
SS Claim							
SSI Claim							

Table 1: Health State Transitions, Age a as Predictor of Age $a+1$, Cont.

Age a	Age $a+1$						
	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							×

Figure 12: Quality Adjusted Life Years for Control Group, Predictions and Comparison to PSID



Crime

Link to Appendix

Quantifying the Benefits in Crime Reduction

1. Count arrests and sentences
2. Construct predictions
3. Estimate number of victims from the crimes
4. Find total costs of crimes

Extensive Sensitivity Analyses for Different Assumptions About Costs of Crime (E.g., Value of Life)

Education

- Follow-up data on educational attainment were collected through age 30.
- Using auxiliary data sources, education up to this age is an accurate predictor of lifetime educational attainment.

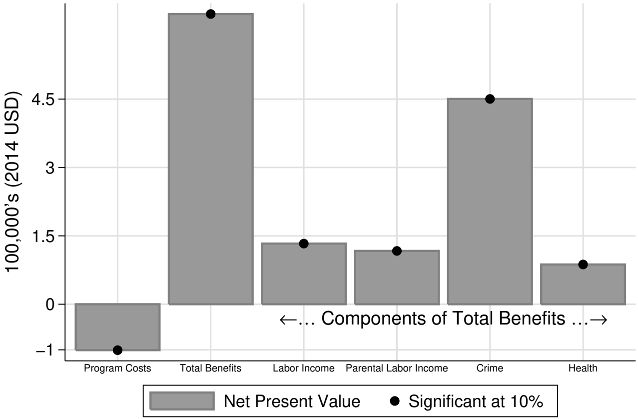
Putting It All Together: Benefit/Cost and Rate of Return Analysis

Program Costs

- The yearly cost of the program: $18,514 per participant (2014 USD).

Link to Detailed Discussion of Program Costs

Figure 13: Net Present Value of Main Components of the Cost/Benefit Analysis Over the Life Cycle per Program Participant, Treatment vs. Next Best



Per−annum Rate of Return: Males and Females 13.7% (s.e. 3%).
Benefit−cost Ratio: Males and Females 7.3 (s.e. 1.8).

Figure 14: Life-cycle Net Present Value of Main Components of the CBA

(a) Males

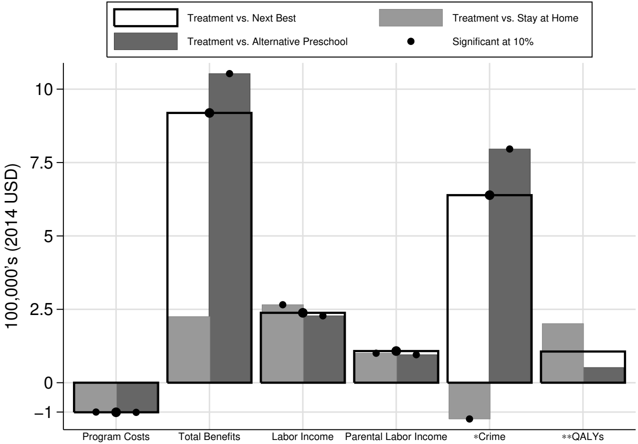


Figure 14: Life-cycle Net Present Value of Main Components of the CBA, Cont'd

(b) Females

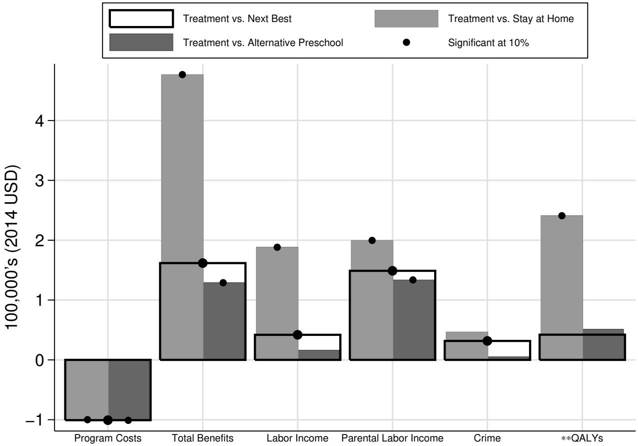
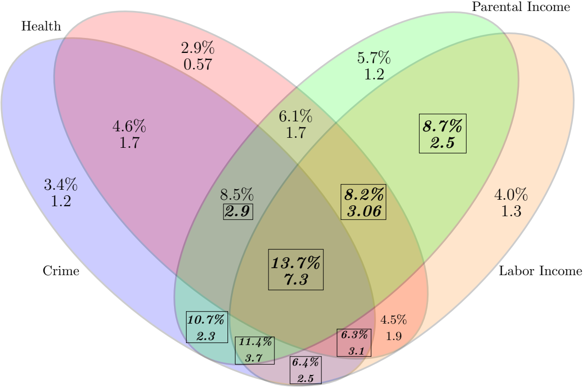


Figure 15: Benefit/Cost Ratio and Internal Rate of Return when Accounting for Different Combinations of the Main Benefits



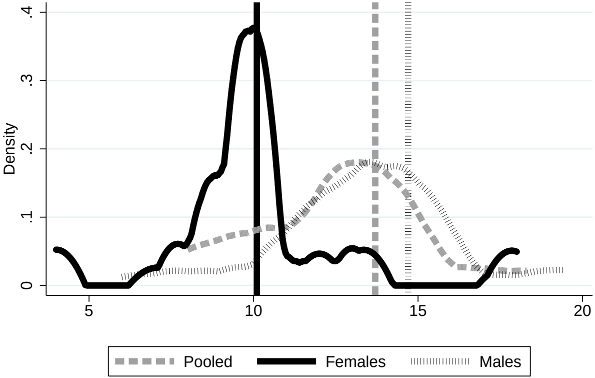
Note: This figure presents all possible combinations of accounting for the benefits from the four major categories in our analysis. The non-overlapping areas present estimates accounting for a single category as the benefit. When two categories overlap, these are the benefits that we account for. And so on for the rest of the plot. The costs remain constant across all calculations and are the same as in Figure 13. Inference is based on non-parametric, one-sided p-values from the empirical bootstrap distribution. We bold point estimates significant at the 10% level.

Table 2: Cost/benefit Analysis of ABC/CARE, Summary

	Females			Males			Pooled		
Removed Component	NPV	IRR	B/C	NPV	IRR	B/C	NPV	IRR	B/C
None	134,240	0.10	**2.45**	935,640	**0.13**	**11.10**	489,010	**0.13**	**6.29**
		(0.08)	(0.79)		(0.06)	(6.35)		(0.05)	(2.11)
Parental Income	101,786	0.04	1.34	68,314	**0.11**	10.36	84,830	**0.09**	**5.36**
		(0.02)	(0.69)		(0.05)	(6.36)		(0.03)	(2.11)
Subject Labor Income	48,007	0.10	**2.16**	258,034	**0.12**	8.22	128,226	**0.11**	**4.86**
		(0.08)	(0.70)		(0.06)	(5.35)		(0.06)	(2.18)
Subject Transfer Income	140	0.10	**2.45**	-7,245	**0.13**	**11.18**	-3,926	**0.13**	**6.33**
		(0.08)	(0.79)		(0.06)	(6.35)		(0.05)	(2.11)
Subject QALY	12,365	0.10	**2.32**	105,122	**0.12**	9.90	81,442	**0.12**	**5.38**
		(0.08)	(0.76)		(0.07)	(6.13)		(0.06)	(2.04)
Medical Expenditures	-11,304	0.09	**2.57**	-44,602	**0.14**	**11.58**	-30,781	**0.14**	**6.62**
		(0.07)	(0.83)		(0.05)	(6.32)		(0.04)	(2.11)
Control Substitution	17,246	0.08	**2.27**	14,283	**0.12**	**10.95**	15,334	**0.11**	**6.12**
		(0.06)	(0.79)		(0.06)	(6.35)		(0.04)	(2.10)
Education Costs	8,989	0.10	**2.36**	3,096	**0.13**	**11.07**	3,084	**0.13**	**6.25**
		(0.07)	(0.79)		(0.06)	(6.34)		(0.04)	(2.11)
Crime Costs	65,209	0.09	1.74	622,376	0.08	4.24	297,855	**0.09**	**3.02**
		(0.08)	(0.72)		(0.04)	(2.72)		(0.05)	(1.14)
Deadweight Loss		0.21	**4.27**		**0.18**	**17.39**		**0.20**	**10.02**
		(0.19)	(1.39)		(0.08)	(9.61)		(0.08)	(3.30)
0% Discount Rate			**5.77**			**29.66**			**16.49**
			(3.55)			(18.12)			(6.97)
7% Discount Rate			**1.66**			4.10			**2.67**
			(0.41)			(2.07)			(0.76)

Objective Sensitivity Analysis (Estimation Error Including Forecast Error)
Subjective Variation (Plausible Ranges) of Non-Monetary Variables

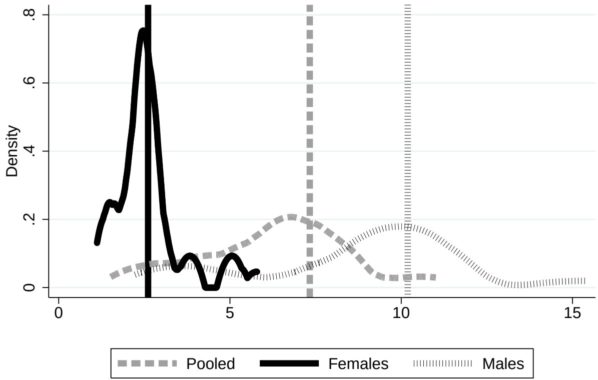
Figure 16: Range of Estimates for Benefit/Cost Ratio



The vertical line represents the baseline estimate.
Average. Pooled: 12.64. Females: 10.07. Males: 13.34.
Median. Pooled: 13.00. Females: 10.00. Males: 13.70.

Note: This figure shows the empirical distribution of all the estimates of the benefit/cost ratio that we provide in the paper, and vary as a function of a wide set of assumptions about the data generating process and the parameters we take from other literature.

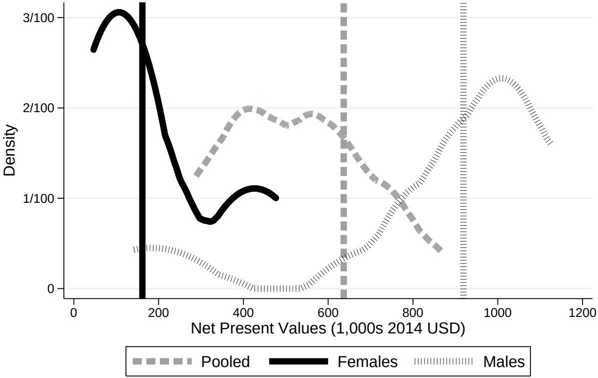
Figure 17: Range of Estimates for Internal Rate of Return



The vertical line represents the baseline estimate.
Average. Pooled: 6.90. Females: 2.63. Males: 9.67.
Median. Pooled: 6.75. Females: 2.49. Males: 9.85.

Note: This figure shows the empirical distribution of all the estimates of the internal rate of return that we provide in the paper, and vary as a function of a wide set of assumptions about the data generating process and the parameters we take from other literature.

Figure 18: Range of Estimates for Net Present Value



The vertical line represents the baseline estimate.
Average. Pooled: 517,608. Females: 190,627. Males: 872,593.
Median. Pooled: 509,466. Females: 129,260. Males: 966,735.

Note: This figure shows the empirical distribution of all the estimates of the net present value that we calculate. They vary as a function of a wide set of assumptions about the data generating process and the parameters we take from other literature.

Link to Sensitivity Tables

Sensitivity to Exogeneity

Link to Sensitivity To Exogeneity Tables

Link to Sensitivity Across Methodologies

Using Our Estimates to Understand Recent Ad-Hoc Benefit/Cost Analyses

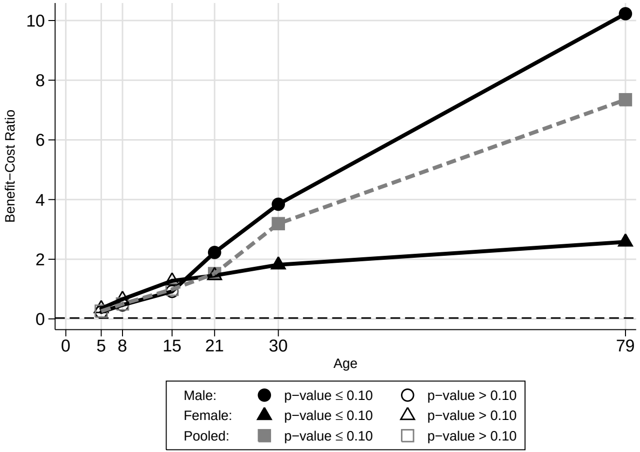
Table 3: Alternative Cost-benefit Analyses Calculations

Age	NPV Source	Component	Kline and Walters (2016) Method	Authors' Method
27	Chetty et al. (2011)	Labor income	0.58 (s.e. 0.28)	
	ABC/CARE-calculated	Labor income	0.09 (s.e. 0.04)	1.09 (s.e. 0.04)
34	ABC/CARE-calculated	Labor income	0.37 (s.e. 0.04)	0.37 (s.e. 0.04)
	ABC/CARE-calculated	All	1.64 (s.e. 0.07)	2.01 (s.e. 0.86)
Life-cycle	ABC/CARE-calculated	Labor income	1.56 (s.e. 0.08)	1.55 (s.e. 0.76)
	ABC/CARE-calculated	All	3.80 (s.e. 0.29)	6.29 (s.e. 2.11)

Note: This table displays benefit/cost ratios based on the methodology in Kline and Walters (2016) and based on our own methodology. Age: age at which we stop calculating the net-present value. NPV Source: source where we obtain the net present value. Component: item used to compute net present value (all refers to the net present value of all the components). Kline and Walters (2016) Method: estimate based on these authors methodology. Author's Method: estimates based on our methodology. Standard errors are based on the empirical bootstrap distribution.

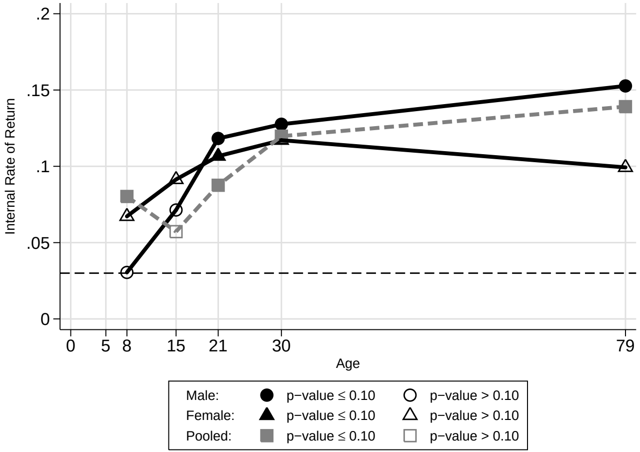
Link to Literature Review

Figure 19: Benefit/Cost Ratio Accounting for Costs and Benefits Up to Certain Ages



Note: This figure displays the estimates of the benefit/cost ratio accounting for the benefits and costs of the program up to the age labeled in the horizontal axis.

Figure 20: Internal Rate of Return Accounting for Costs and Benefits Up to a Certain Age



Note: This figure displays the estimates of the internal rate of return accounting for the benefits and costs of the program up to the age labeled in the horizontal axis.

Summary

- Analyze influential study widely emulated early childhood education intervention with randomized controlled design.
- A template for evaluating social programs confronting a number of methodological and practical issues, going from non-compliance, attrition, multiple hypotheses testing, and the need to predict of long-term outcomes.
- Demonstrate the value of economic and econometric theory in understanding the economic consequences of the program and supplementing the information obtained from social experiments.
- No conflict between structural and causal analysis.
- Economic theory encoded in structural economic models and enabled by econometrics enriches what can be learned from social experiments.
- Develop and apply a procedure for forecasting out-of-sample benefits guided by economic fundamentals

THE UNIVERSITY OF CHICAGO

- Examine the multiple facets of an influential early childhood program.
- Develop and apply principled methods for conducting out of sample cost-benefit analysis and testing assumptions
- Life cycle health costs and benefits calculated using the Future American Model.
- Standard errors on the aggregate estimates to capture the uncertainty of the estimations (sampling error including pretest variation).
- Extensive sensitivity analyses that varies the assumptions of more subjective or controversial decisions about costs and benefits (e.g. discount rate and benefits/costs of crime).
- The results are statistically significant and robust to a variety of perturbations.

- Variety of estimation and testing methods
- Agreement across methods:
  1. Multiple hypothesis testing across blocks
  2. Factor models within blocks
  3. Combining functions by blocks
  4. Benefit/Cost Ratio and Internal Rate of Return within and across blocks

THE UNIVERSITY OF CHICAGO

- High rates of return
- Strong gender differences in benefits and sensitivity to taking male child out of the home
- Controlling for substitution bias matters
- But using economics this "threat to validity" also gives us information about agent choices and their consequences and gives us a richer understanding of the consequences of the program studied.

ABC/CARE is Effective

It supports childcare and educational attainment for mothers and is economically efficient

It works for boys and girls, but in different dimensions and is economically efficient

It reduces crime and is economically efficient

It's expensive, but has a very high rate of return

Program permits long-term follow-up - In addition, program widely replicated and motivates a variety of new initiatives

Control group substitution is informative on differential impact of low quality child-care on boys and benefits of home care

Different methodological approaches agree

The rate of return is 13.7% - it survives extensive sensitivity analysis

Thank You

Institute for
New Economic Thinking

PRITZKER CHILDREN'S INITIATIVE

Buffett Early Childhood Fund
investing in the early years

Robert Wood Johnson Foundation

NIH
National Institutes of Health

THE UNIVERSITY OF CHICAGO

Appendices

Appendices Table of Contents

Appendix: Quantifying the Benefits in Crime Reduction

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.

ABC and CARE Datasets

1. Administrative Youth Arrests Dataset
2. Administrative Adult Arrests Dataset
3. Administrative Sentences Dataset
4. Self-reported Adult Crimes Dataset

Auxiliary Datasets

1. National Crime Victimization Survey (NCVS)
2. **Uniform Crime Reporting Statistics (UCRS)**
3. National Judicial Reporting Program (NJRP)
4. North Carolina Department of Public Safety Dataset (NCDPS)

Crime Categories

Table 4: Crime Categories

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

Methodology for Estimating Crime Costs

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 $\overline{A_j}$ be the national total number of arrests for crime type, j, and let $\overline{S_j}$ be the national total number of sentences. Then, we construct $r_j = \frac{\overline{A_j}}{\overline{S_j}}$, 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: $f_j = \frac{\overline{V_j}}{\overline{A_j}}$. 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.

THE UNIVERSITY OF CHICAGO

Figure 21: Counts of Arrests and Sentences

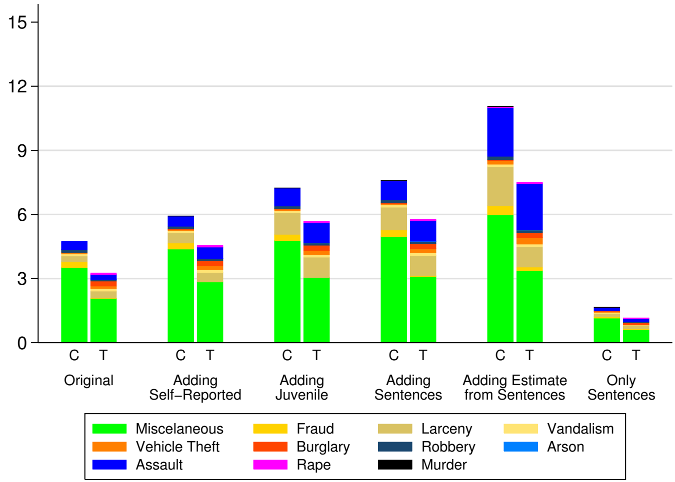
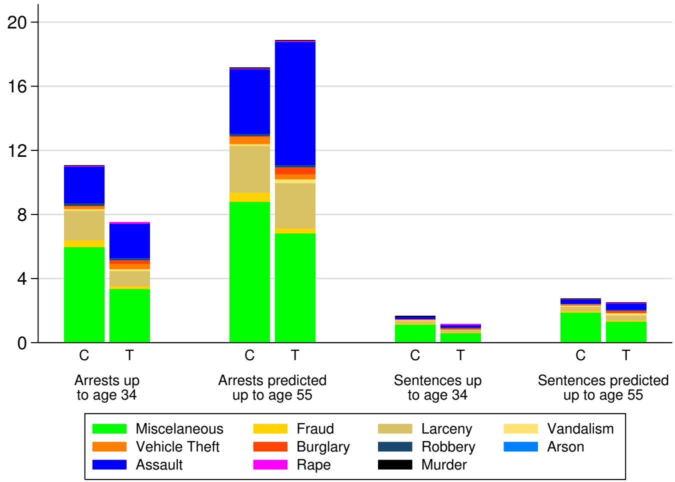


Figure 22: Constructed Predictions



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.
- 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.

Figure 23: Victims-to-Arrests Ratios by Crime

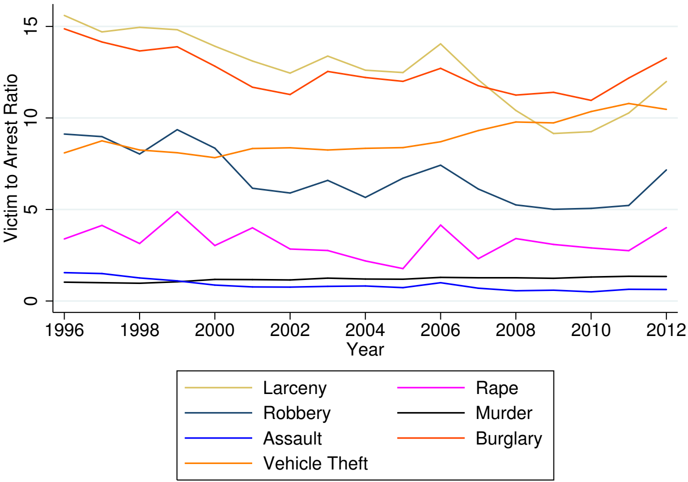
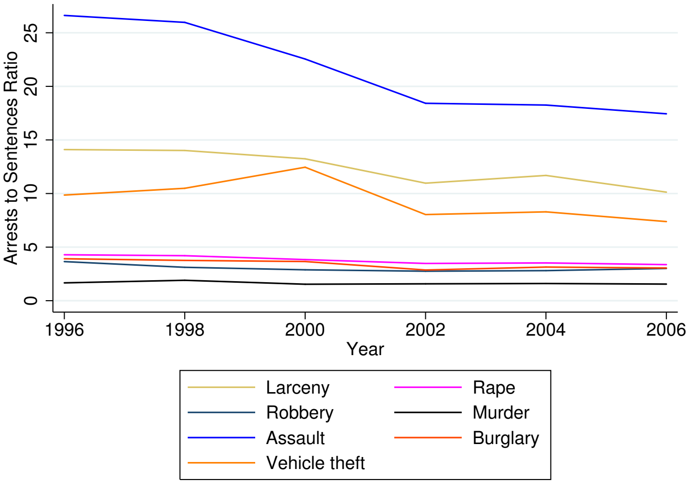


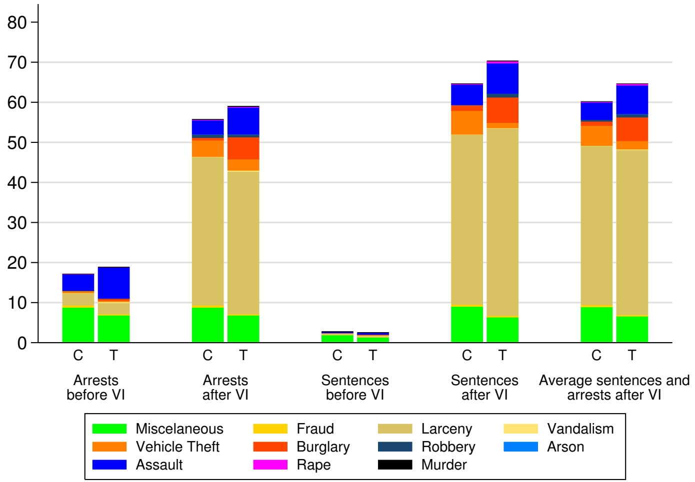
Figure 24: Arrests-to-Sentences Ratio by Crime



Effects on Number of Crimes, After Victimization Inflation

- Figure 25 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 25: Effects on Number of Crimes, After Victimization Inflation



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.

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.
- 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.

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.
   - Being an example of a TD methodology, it is expected that the costs obtained by this method 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.
   - 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.

THE UNIVERSITY OF CHICAGO

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.

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 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.

Costs Used in This Study

- 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.

Costs Used in This Study

- 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 5.
- Notice that there are some strong differences in the cost of crimes, such as assault, burglary, and especially robbery.

Table 5: Monetary 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

Figure 26: Costs of Crime Before Victimization Inflation

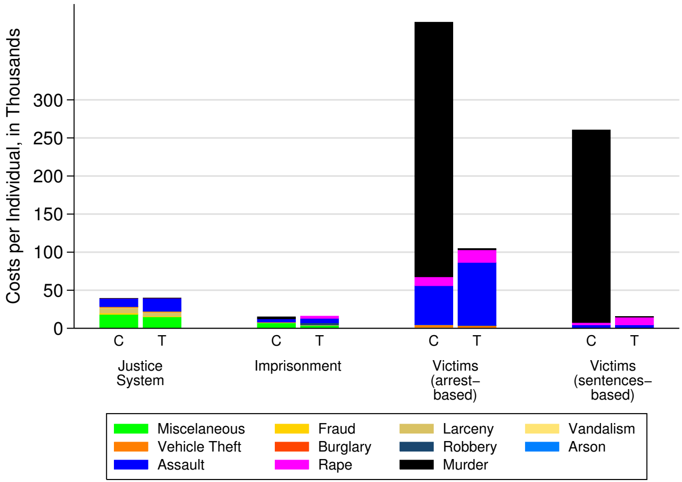
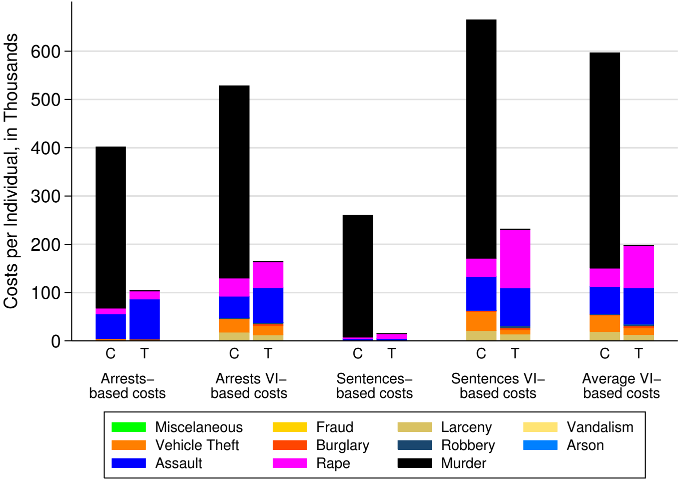


Figure 27: Costs of Crime After Victimization Inflation



Return to main text

Program Details

First-phase Treatment, ABC

- First-phase Treatment, ABC: one control group and one treatment group
  - Control group (56 children):
    1. Iron-fortified formula and monthly supply of diapers, first 15 months of life
  - Treatment group (58 children):
    1. Iron-fortified formula and monthly supply of diapers, first 15 months of life
    2. Breakfast, lunch, and afternoon snack
    3. Medical care from nurses, supervised by a doctor
    4. Center-based childcare

First-phase Treatment, CARE

- First-phase Treatment, CARE: one control group and two treatment groups
  - Control group (23 children):
    1. Iron-fortified formula and monthly supply of diapers, first 15 months of life
  - Family education treatment group (27 children)
    1. Iron-fortified formula and monthly supply of diapers, first 15 months of life
    2. Home visits that aimed to help parents solve common problems related to child rearing.
  - Center-based childcare and family education treatment group (16 children):
    1. Same as the family education treatment group
    2. Center-based childcare

Second-phase Treatment, ABC and CARE

- For both programs: home visits from ages 5 to 8
  $\rightarrow$ similar in objectives to first-phase home visits of CARE
- ABC: re-randomized at age 5 to either receive or not the home visits
  $\rightarrow$ 96 children were re-randomized
- CARE: participants of the two first-phase treatment groups received second-phase treatment

Eligibility

- Both programs targeted disadvantaged children from the semi-rural communities of Chapel Hill close to the Frank Porter Graham Center (FPGC) of the University of North Carolina
  - Mothers in the last trimester of pregnancy were referred by local social service agencies and hospitals
  - Eligibility was determined by a score of 11 or more on the High-risk Index (HRI)
  - Example HRI items:
    - Mother's education level
    - Use of welfare programs
    - Father's presence at home
  - Although race was not a consideration for eligibility, 98% of ABC participants and 90% of CARE participants were African-American

Sample

- ABC
  - Four cohorts of children born between 1972 and 1977
  - 122 individuals recruited
- CARE
  - Two cohorts of children born in 1978 and 1979
  - 67 individuals recruited
- Overall, mothers in CARE were older, more educated, and had higher IQ than the mothers in ABC

ABC and CARE Samples in Context

- We compare the ABC and CARE samples to a comparison group using a cohort of the Panel Study of Income Dynamics born in the same years as the ABC and CARE subjects (1972-1979)
- ABC and CARE subjects were born to younger, less educated mothers many of whom were raising their children without the father present

ABC Randomization

- First Phase: 121 children randomized to one treatment group that received center-based childcare and one control group
    - Effective sample size after randomization compromises: 114 (58 treatment, 56 control)
- Second Phase: 96 of the original children were randomized into one treatment and one control group

CARE Randomization

- First Phase: 67 children randomized into:
  - One treatment group that received center-based childcare and family education (16 children)
  - One treatment group that only received family education (27 children)
  - One control group (23 children)
- Second Phase: Children in the two treatment groups automatically received the second-phase treatment of home visits and the control group remained the same
- No randomization compromises except death and families moving away from the study area

ABC First-phase Randomization Compromises

1. **Left the study before data collection:** We have no data at all for these subjects (4 treatment)
2. **Death before age 5 / Moved out:** We include them in estimations until data are no longer available. Thereafter, they are cases of attrition (2 treatment and 2 control)
3. **Partial treatment:** We assume that they had full treatment (4 children)
4. **Noncompliance to treatment:** We keep the original treatment status for them for ITT estimations (3 treatment)
5. **Crossover from control to treatment:** Three children switched status from control to treatment. We keep the original treatment statuses for ITT estimations (3 children)
6. **Developmental delays:** We drop them because they were not eligible for the program (2 treatment)

ABC Second-phase Randomization Compromises

1. **Not randomized in second phase:**
   - Stopped being followed-up: They are considered cases of attrition
   - Followed-up in later data collection: They are not included when calculating the treatment effects for the second phase, but are included when estimating treatment effects of the first phase on later outcomes
2. **Noncompliance in second phase:** Original treatment statuses are kept for ITT estimations

THE UNIVERSITY OF CHICAGO

Programmatic Elements

- The objectives of both ABC and CARE were to prevent "mental retardation" and develop school readiness
- The different curricula implemented across the programs and cohorts had the following goals:
    - Support language and cognitive development
    - Develop socio-emotional skills considered to enable school-readiness (e.g., task orientation)

THE UNIVERSITY OF CHICAGO

Additional Programmatic Elements

- The ABC treatment group received
  - Daily health screenings and frequent medical check-ups
- The CARE treatment groups received
  - Home visits to help parents form problem-solving skills
- Both ABC and CARE center-based treatment groups received
  - Transportation to and from FPGC
  - Daily nutritious food
- Both ABC and CARE control groups received
  - Iron-fortified formula until the child was 15 months old
  - Unlimited diapers until the child was 3 years old

Programmatic Elements, Second Phase

- Same treatment in ABC and CARE
- State-certified "home-school resource teachers"
- Visited the elementary school and the children's homes twice a month to help
  - Engage the parents with the children's academics
  - Provide one-on-one tutoring to the children
  - Parents with issues related to literacy, housing, and medical care

Baseline Characteristics in ABC and CARE

Table 6: Baseline Characteristics in ABC and CARE

Variable	Age	ABC Obs	CARE Obs	ABC Mean	CARE Mean	*p*-value Single H_0	*p*-value Multiple H_0
Male	0	116	67	0.464	0.596	**(0.060)**	(0.110)
Birth Weight	0	114	64	7.008	7.139	(0.625)	(0.765)
No. Siblings in Household	0	116	67	0.632	0.684	(0.810)	(0.890)
Birth Year	0	116	67	1974	1979	**(0.000)**	**(0.000)**
Mother's Education	0	116	67	10.188	10.868	**(0.010)**	**(0.025)**
Mother's Age	0	116	67	19.828	21.141	**(0.060)**	**(0.100)**
Mother's IQ	0	116	67	84.407	87.164	**(0.070)**	(0.130)
Father at Home	0	116	67	0.283	0.209	(0.270)	(0.380)

Programmatic Elements, First Phase Treatment

Table 7: Elements of First Phase Treatment, ABC and CARE

	ABC	CARE
Treatment	Center-based childcare	Center-based childcare and family education
Center-based Childcare		
Intensity	6.5–9.75 hours a day for 50 weeks per year	6.5–9.75 hours a day for 50 weeks per year
Components	Instruction, medical care, nutrition, social services	Instruction, medical care, nutrition, social services
Staff-to-child Ratio	1:3 during ages 0–1 1:4–5 during age 1–4 1:5–6 during ages 4–5	1:3 during ages 0–1 1:4–5 during age 1–4 1:5–6 during ages 4–5
Staff Qualifications	Mixed diplomas; experienced	Mixed diplomas; experienced
Family Education		
Intensity		One hour-long home visits. 2–3 per month during ages 0–3. 1–2 per month during ages 4–5
Curriculum		Social and mental stimulation; parent-child interaction
Staff-to-child Ratio		1:1
Staff Qualifications		Home visitor training

Programmatic Elements, Second Phase Treatment

Table 8: Elements of Second Phase Treatment, ABC and CARE

	ABC	CARE
Intensity	Every other week	Every other week
Components	Parent-teacher meetings	Parent-teacher meetings
Curriculum	Reading and math	Reading and math
Staff-to-child Ratio	1:1	1:1
Staff Qualifications	Graduate degree and training in special education	Graduate degree and training in special education

Return to main text

Estimated Combining Functions by Outcome Categories

Figure 28: Percentage of Outcomes with Positive Treatment Effects, First Set of Categories

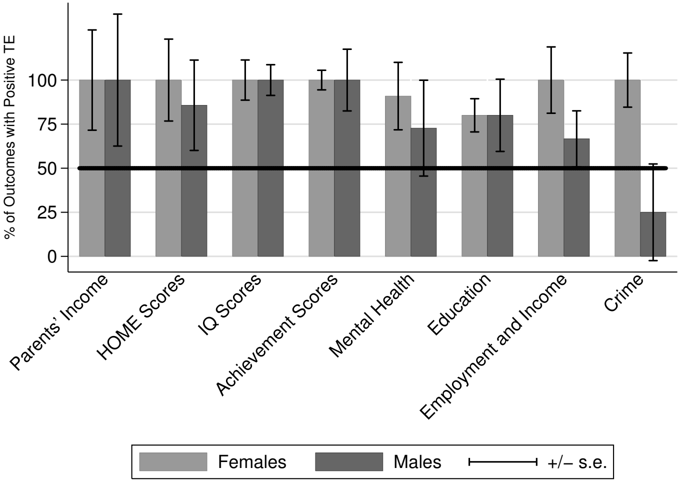


Figure 29: Percentage of Outcomes with Positive Treatment Effects, Second Set of Categories

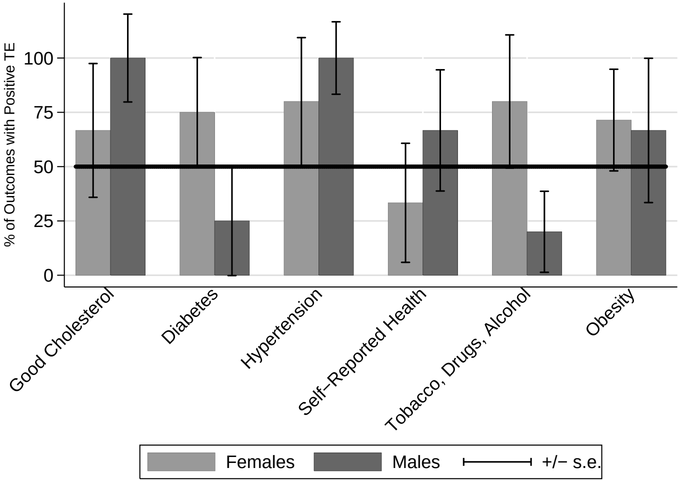


Figure 30: Percentage of Outcomes with Positive Treatment Effects Fixing Control Group to Stay at Home, First Set of Categories

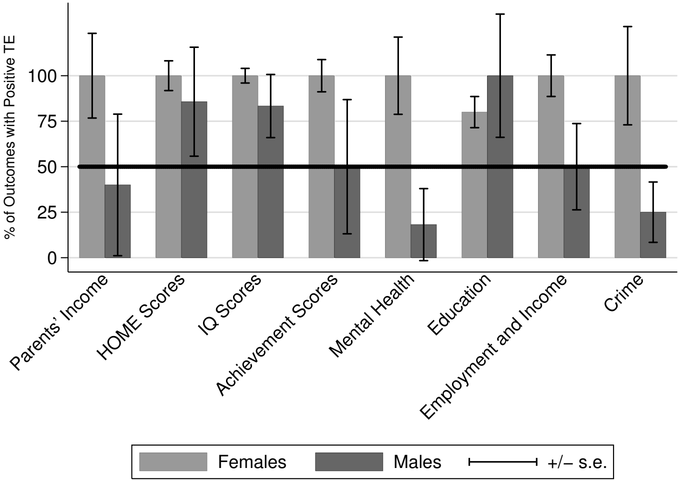


Figure 31: Percentage of Outcomes with Positive Treatment Effects Fixing Control Group to Stay at Home, Second Set of Categories

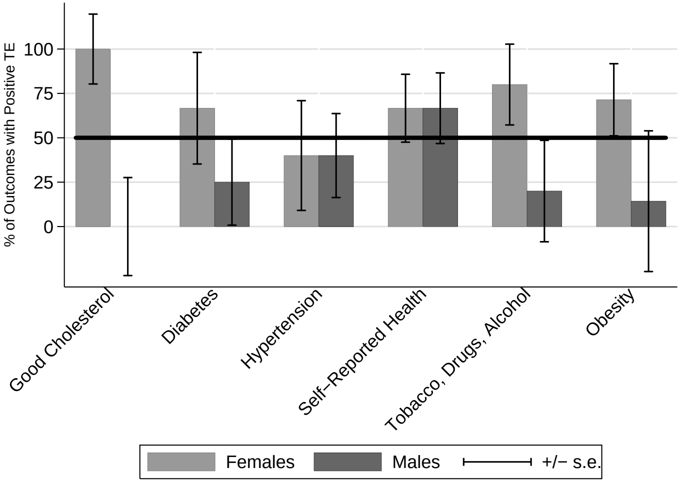


Figure 32: Percentage of Outcomes with Positive Treatment Effects Fixing Control Group to Alternative Preschool, First Set of Categories

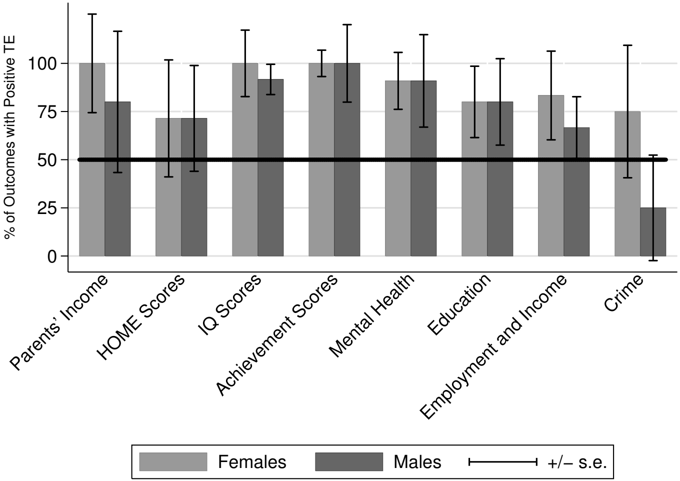
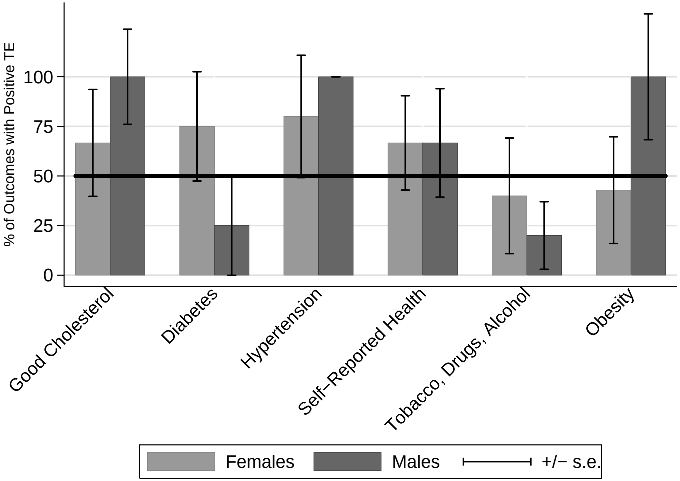


Figure 33: Percentage of Outcomes with Positive Treatment Effects Fixing Control Group to Alternative Preschool, Second Set of Categories



Return to main text

Common Support between Experimental and Non-Experimental Samples

- W (pre-program variables) and X (programs possibly affected by treatment).
- $W =$ male indicator, black indicator, mothers education
- $X =$ PIAT scores at ages 5-7, years of education at age 30 income at ages 21 and 30, BMI at age 34
- How do we choose the variables in W, X?
  - Restriction imposed by the availability (overlap) of the the data in the experimental and non-experimental sources
  - Careful cross-walk to maximize the amount of variables we were able to work with

Figure 34: Support of ABC/CARE and Auxiliary Data: Income at Age 21

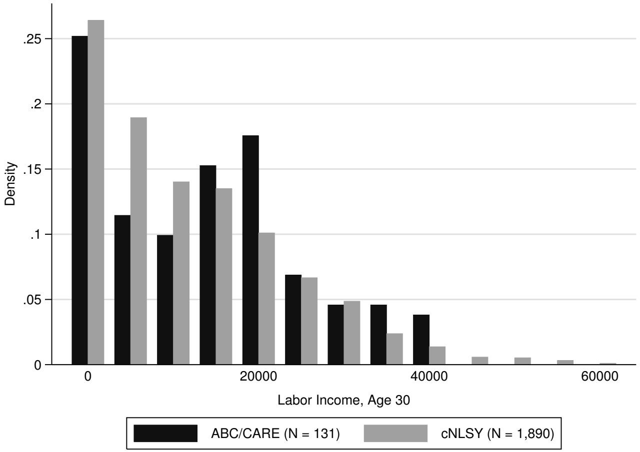


Figure 34: Support of ABC/CARE and Auxiliary Data: Income at Age 30

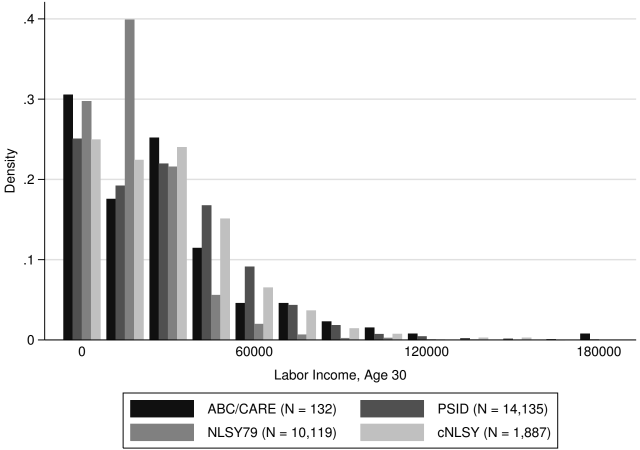


Figure 34: Support of ABC/CARE and Auxiliary Data: Subject's Years of Education

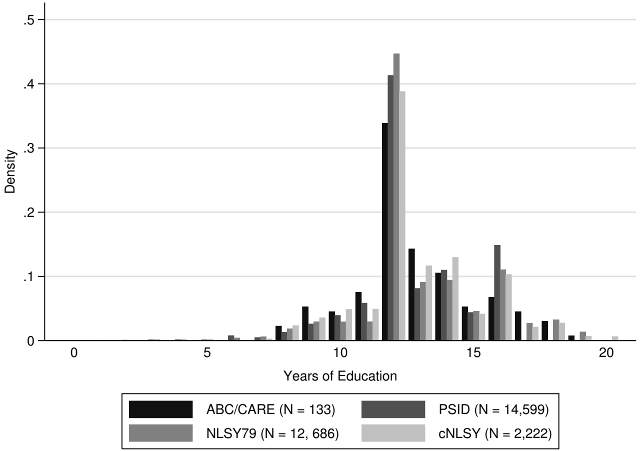


Figure 34: Support of ABC/CARE and Auxiliary Data: Mother's Years of Education

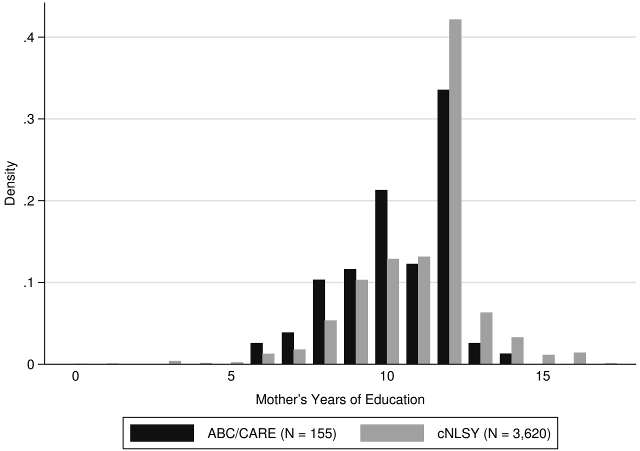


Figure 34: Support of ABC/CARE and Auxiliary Data: Average PIAT Math Scores, Ages 5–7

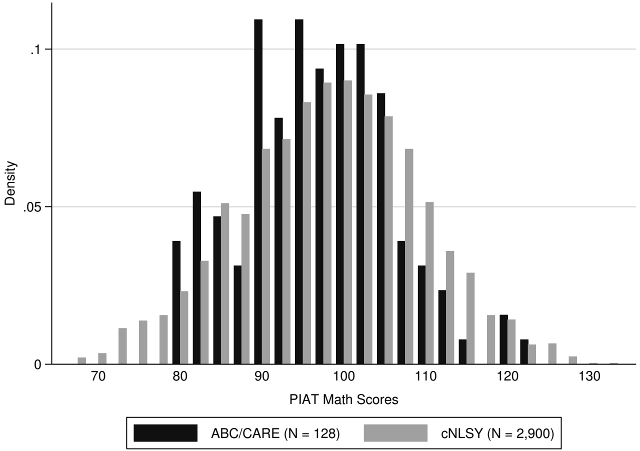
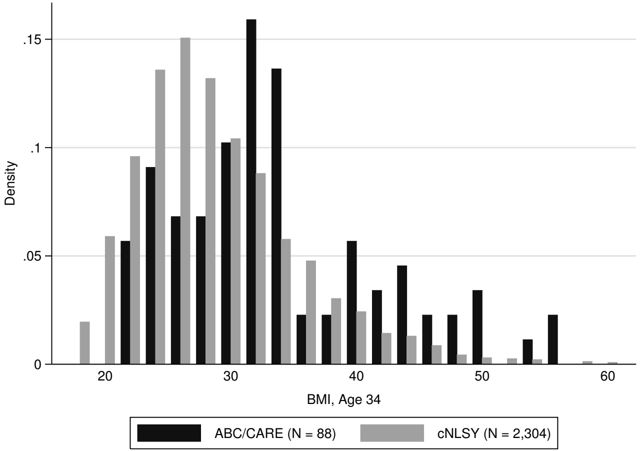


Figure 34: Support of ABC/CARE and Auxiliary Data: Body Mass Index, Age 34



Return to main text

Full Experimental Data Tables

Table 9: 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

Table 9: Early Childhood Data (Part I), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Cognitive Assessments	Language Ability	Auditory association, Verbal expression, etc.	36, 42, 48, 54	30, 42, 54	ITPA^{ABC}, GPB^{ABC}, PLP^{ABC}, 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^{ABC}
		RPM	60	-	RPM^{ABC}
	Quantitative	BSID	3, 6, 9, 12, 18, 24	6, 12, 18, 24	BSID
		MSCD	30, 42, 54	30, 42, 54	MSCD
	Memory	BSID	3, 6, 9, 12, 18, 24	6, 12, 18, 24	BSID
		MSCD	30, 42, 54	30, 42, 54	MSCD
	Motor Development	BSID	3, 6, 9, 12, 18, 24	6, 12, 18, 24	BSID
		MSCD	30, 42, 54	30, 42, 54	MSCD
	Critical Thinking	Curiosity	30, 36, 42, 48, 54, 60, 66, 72	-	Infant Behavior Inventory^{ABC}

Table 9: Early Childhood Data (Part I), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Non-Cognitive Assessments	Social Skills	Positive social response	30, 36, 42, 48, 54, 60, 66, 72	6, 12, 18, 24	Infant Behavior Inventory^{ABC}, Bayley Infant Inventory^{CARE}
		Creativity	30, 36, 42, 48, 54, 60, 66, 72	-	Infant Behavior Inventory^{ABC}
	Self-Control	Locus of control	3, 18	6, 18	RIES
		Distractibility, Attentiveness	30, 36, 42, 48, 54, 60, 66, 72	6, 12, 18, 24	Infant Behavior Inventory^{ABC}, Bayley Infant Inventory^{CARE}
	Emotional Health	KRT	24, 36, 48, 60	24, 30, 36, 42, 48, 60	KRT
	Self-Consciousness	Self-consciousness	30, 36, 42, 48, 54, 60, 66, 72	-	Infant Behavior Inventory^{ABC}

THE UNIVERSITY OF CHICAGO

Table 10: 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	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	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^{CARE}, Parental Attitudes Research Inventory
		Father at home	18, 30, 42, 54	18, 30, 42, 54, 60	Demographic Interview

Table 10: Early Childhood Data (Part II), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Family Environment	Family Economic Environment	Parents' occupation	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54, 60	Demographic Interview
		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
		Family income	Birth, 18, 30, 42, 54	Birth, 18, 30, 42, 54, 60	Demographic Interview
	Parents and Home Environment	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^{ABC}
	Family Members' Physical Health	Health issues of parents	Birth	Birth	Parent Interview
		Pregnancy history	Birth	Birth	Parent Interview

Table 10: Early Childhood Data (Part II), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Childcare	Day-care Experience	Time and location of childcare, Age when begin	Birth, 18, 30, 42, 54	18, 30, 42, 54	Demographic Interview
		Home visits	-	6, 18, 30, 42, 54, 60	Home Visit Data^{CARE}
	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
		Parents' knowledge of childcare	Birth	-	Parent Interview^{ABC}
Physical Health	Growth Data	Height, Weight, Head circumference, etc.	3, 6, 9, 12, 18, 24, 36, 48, 60	Birth, 6, 12, 18, 24, 36, 48, 60	Growth Measures

Table 11: Childhood and Adolescent 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^{ABC}
		MSCD	7	-	MSCD^{ABC}
	Intelligence Tests	SBIS	6	7	SBIS
		WIS	6, 7, 8, 12, 15	6, 8	WIS
		Kaufman^{CARE}	-	6	Kaufman^{CARE}
	Quantitative Skills	MSCD^{ABC}	7	-	MSCD^{ABC}
	Memory	MSCD^{ABC}	7	-	MSCD^{ABC}
	Motor Skills	MSCD^{ABC}	7	-	MSCD^{ABC}

Table 11: Childhood and Adolescent Data (Part I), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Non-Cognitive Assessment	Interpersonal Skills	Gets along with people	6, 8, 12, 15	8, 12	PEI, CAS, PMI^{ABC}, SAI^{ABC}, Subject Interview^{ABC}, Quality Rank^{CARE}
		Relationship with the other sex	15	-	SAI^{ABC}, Subject What I Am Like (Harter)^{ABC}
	Critical Thinking	Thinks for self, questions things	6, 8	8, 12	PEI, Harter Child^{CARE}, CBI
		Concept Attainment Kit	6, 7, 8	-	Concept Attainment Kit^{ABC}
	Self-Control	Distracted in class	6, 7, 8, 12, 15	12	SCAN^{ABC}, CBI, WPB^{ABC}, PMI^{ABC}, SAI^{ABC}, Self-Evaluation Inventory^{ABC}
		Locus of control	15	-	Nowicki-Strickland Data, Pearlin Mastery Scale^{ABC}
	Work Ethic	Task Orientation	6, 7, 8, 12, 15	6, 7, 8, 9, 12	SCAN^{ABC}, CBI, PMI^{ABC}

Table 11: Childhood and Adolescent Data (Part I), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Non-Cognitive Assessment	Emotional Health	Harms self, suicidal thoughts	8, 12, 15	8, 12	Achenbach Parent, Subject Risk Taking Survey^{ABC}
		Depression, anxiety, fear, etc.	6, 7, 8, 12, 15	7, 8, 9, 12	KRT, CAS, ETS, Achenbach Parent
	Social Activities	Athletic activities	8, 12, 15	8, 12	Achenbach Parent, SAI^{ABC}, Subject What I Am Like (Harter)^{ABC}, PEI^{CARE}
		Participant of organizations, e.g. religions	8, 12, 15	8, 12	Achenbach Parent, SAI^{ABC}, Subject Interview^{ABC}
		Reading list	12, 15	12	CAS, SAI^{ABC}
		TV/music	12, 15	12	CAS, SAI^{ABC}, Television Checklist^{ABC}
	Self-Consciousness	Self-conscious emotions	8, 12, 15	8, 12	Achenbach Parent, Subject What I Am Like (Harter)

Table 12: Childhood and Adolescent 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^{ABC}
		Relationship with family members, including father, mother, siblings, etc.	6, 8, 12, 15	8, 12	PEI, FES, SAI, Subject Interview^{ABC}, Adult Self Report^{ABC}, Parent Interview, Achenbach Parent
		Number of siblings	6, 8, 12, 15	7, 8, 12	PEI^{ABC}, Parent Interview
		Marital status of parents	6, 8, 12, 15	7, 8, 12	PEI^{ABC} , Parent Interview
		Father at home	18, 30, 42, 54	18, 30, 42, 54, 60	Demographic Interview
	Parents' Education Style	Role of parents in education	6, 8	8, 12	PEI, Parent Interveiw^{CARE}
		Parents' education beliefs & methods	6, 8	8, 12	PEI, Parent Interview^{CARE}
		Parents' aspiration & attitudes towards child	6, 8, 12, 15	8, 12	PEI, Parent Interview

Table 12: Childhood and Adolescent Data (Part II), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Family Environment	Family Economic Environment	Parents' occupation	6, 8, 12, 15	7, 8, 12	PEI^{ABC}, Parent Interview
		Mother works	9	5, 7, 8	Demographic Interview
		Source of child support	6, 8, 12, 15	7, 8, 12	PEI^{ABC}, Parent Interview
		Family income	6, 8, 12, 15	7, 8, 12	PEI^{ABC}, Parent Interview
	Parents and Home Environment	Parents' authority, warmth, family conflict, etc.	8	8	Parent Interview
	Family Social Status	Parents' education background	6, 8, 12, 15	7, 8, 12	PEI^{ABC}, Parent Interview
		Criminal history and risk taking of family members	8, 12, 15	-	Subject Taylor Life Events^{ABC}, Parent Interview^{ABC}
	Family Members' Physical Health	Health issues of adults in house	8, 12, 15	12	Parent Interview, Subject Taylor Life Events^{ABC}

Table 12: Childhood and Adolescent Data (Part II), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Academic Achievements	Standardized Tests	Reading, mathematics, and language abilities	6, 7, 8, 12	6, 8, 9,12	CAT^{ABC}, PIAT^{ABC}, WJCA
	Performance in Schoolwork	Drop in grades	12, 15	12	CAS
		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
Physical Health	Health Issues	Health issues of subject	8, 12, 15	8, 12	Achenbach Parent, Subject Interview^{ABC}, Adult Self Report^{ABC}, PEI^{CARE}, Parent Interview^{CARE}
	Growth	Vision, weight, height	8	8	Growth Data
	Teenage Pregnancy	Teenage Pregnancy	15	-	Subject Interview^{ABC}
Social Conduct	Law Breaking	Felony, Time spent incarcerated	15	-	MARS^{ABC}, Subject Interview^{ABC}

Table 13: Adult Data (Part I)

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Cognitive Assessments	Intelligence Tests	WIS	21	-	WIS
Non-Cognitive Assessment	Interpersonal Skills	Gets along with people	21, 30	-	Subject Interview
	Self-Control	Locus of control	21, 30	-	Nowicki-Strickland Data^{ABC}, Pearlin Mastery Scale^{ABC}
		Proud of working, interest in working	21, 30	21, 30	Job Satisfaction Survey^{ABC}, Subject Interview
	Emotional Health	Harms self, suicidal thoughts,	21	21	Achenbach, Subject Risk Taking Survey
		depression, anxiety, fear, etc.	21, 30	21, 30	KRT, Achenbach Parent, CAS, Brief Symptom Inventory, ETS
	Social Activities	Athletic activities	21	-	Achenbach,
		Participant of organizations, e.g. religions	21, 30	21, 30	Achenbach, Subject Interview

Table 13: Adult Data (Part I), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Family Environment	Family Members	Number of adults in house	21	-	Parent Interview^{ABC}, Subject Interview
		Relationship with family members, including father, mother, siblings, etc.	21, 30	30	Parent Interview, Achenbach^{ABC}, Subject Interview, Adult Self Report
		Number of siblings	21, 30	30	Parent Interview^{ABC}, Subject Interview
		Marital status of parents	21	-	Parent Interview^{ABC}, Subject Interview
		Number of children, childcare basics	21, 30	30	Subject Interview, Childcare Questionnaire
	Family Economic Environment	Parents' occupation	21	-	Parent Interview^{ABC}, Subject Interview
		Source of child support	21	30	Parent Interview^{ABC}, Subject Interview
		Family income	21	30	Parent Interview^{ABC}, Subject Interview
	Family Members and Children	Relationship quality, family health issues, attitude toward child learning	30	30	Parent Interview, Taylor Life Events^{ABC}, Child Health Questionnaire, PEI

Table 13: Adult Data (Part I), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Family Environment	Marital Status	Marital status, spouse income	21, 30	21, 30	Subject Interview
		Spouse details, marriage history	21, 30	30	Subject Interview
		Relationship with spouse	21, 30	30	Subject Interview, Adult Self Report
Achievement	Education Level	Years in school, plans for future education	21, 30	21, 30	Subject Interview, Adult Self Report
		College type, certificate earned	21, 30	21, 30	Subject Interview, Adult Self Report
	Achievement Test	WJCA	21, 30	-	WJCA

Table 14: Adult 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
	Law Breaking	Felony, Time spent incarcerated	21	21, 30	Subject Interview

Table 14: Adult Data (Part II), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Economic Status	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
		Satisfied with current job	21, 30	21, 30	Subject Interview, Subject What I Am Like (Harter), Adult Self Report

Table 14: Adult Data (Part II), Cont.

Category	Sub-Category	Description	ABC Age	CARE Age	Measure
Economic Status	Transportation	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

Return to main text

ABC/CARE Tables

Table 15: ABC and CARE, Program Comparison

	ABC	CARE	ABC = CARE ?
Program Overview			
Years Implemented	1972–1982	1978–1985	
First-phase Treatment	Birth to 5 years old	Birth to 5 years old	✓
Second-phase Treatment	5 to 8 years old	5 to 8 years old	✓
Initially Recruited Sample	121	67	
# of Cohorts	4	2	
Eligibility	Socio-economic disadvantage according to a multi-factor index	Socio-economic disadvantage according to a multi-factor index	✓
Control			
N	56	23	
Compensation	Diapers from birth to age 3, unlimited formula from birth to 15 months	Diapers from birth to age 3, unlimited formula from birth to 15 months	✓
Control Substitution	75%	74%	

Note: This table compares the main elements of ABC and CARE, summarized in this section. A ✓indicates that ABC and CARE had the same feature. A blank space indicates that the indicated component was not part of the program.

Table 15: ABC and CARE, Program Comparison

	ABC	CARE	ABC = CARE ?
Treatment	Center-based childcare	Center-based childcare and family education	
Center-based Childcare			
N	58	17	
Intensity	6.5–9.75 hours a day for 50 weeks per year	6.5–9.75 hours a day for 50 weeks per year	✓
Components	Stimulation, medical care, nutrition, social services	Stimulation, medical care, nutrition, social services	✓
Staff-to-child Ratio	1:3 during ages 0–1	1:3 during ages 0–1	✓
	1:4–5 during age 1–4	1:4–5 during age 1–4	✓
	1:5–6 during ages 4–5	1:5–6 during ages 4–5	✓
Staff Qualifications	Range of degrees beyond high school; experience in early childcare	Range of degrees beyond high school; experience in early childcare	✓

Note: This table compares the main elements of ABC and CARE, summarized in this section. A ✓indicates that ABC and CARE had the same feature. A blank space indicates that the indicated component was not part of the program.

THE UNIVERSITY OF CHICAGO

Table 15: ABC and CARE, Program Comparison

	ABC	CARE	ABC = CARE ?
Treatment	Center-based childcare	Center-based childcare and family education	
Home Visitation			
N	(not part of the program)	27	
Intensity		Home visits lasting 1 hour. 2–3 times per month during ages 0–3. 1–2 times per month during ages 4–5	
Curriculum		Social and mental stimulation; parent-child interaction	
Staff-to-child Ratio		1:1	
Staff Qualifications		Home visitor training	

Note: This table compares the main elements of ABC and CARE, summarized in this section. A ✓indicates that ABC and CARE had the same feature. A blank space indicates that the indicated component was not part of the program.

Table 15: ABC and CARE, Program Comparison

	ABC	CARE	ABC = CARE ?
School-age Treatment			
Intensity	Every other week	Every other week	✓
Components	Parent-teacher meetings	Parent-teacher meetings	✓
Curriculum	Reading and math	Reading and math	✓
Staff Qualifications	Range of degrees beyond high school; experience in early childcare	Range of degrees beyond high school; experience in early childcare	✓

Note: This table compares the main elements of ABC and CARE, summarized in this section. A ✓indicates that ABC and CARE had the same feature. A blank space indicates that the indicated component was not part of the program.

Return to main text

Table 16: Data Availability for ABC and CARE (Part I)

Category	Age 0-5	Age 5-15	Adult
Physical Health			
Growth data	✔	-	-
Health issues	-	✔	✔
Full medical sweep	-	-	✔
Family Environment			
Family Members	✔	✔	✔
Economic Environment	✔	✔	✔
Family Social Status	✔	✔	-
Family Physical Health	✔	✔	-
Marital Status/Number of Children	-	-	✔
Childcare			
Daycare/Parental Care Info	✔	-	-
Cognitive Assessments			
Intelligence Levels	✔	✔	Only ABC
Language Ability	✔	✔	-
Motor Development	✔	✔	-
Critical Thinking	Only ABC	✔	-

Table 17: Data Availability for ABC and CARE (Part II)

Category	Age 0-5	Age 5-15	Adult
Non-Cognitive Assessments			
Social Skills	✔	✔	✔
Self Control	✔	✔	Only ABC
Self-Consciousness	Only ABC	✔	-
Work Ethic	-	✔	-
Social Activities	-	✔	✔
Academic Achievements			
Standardized Tests	-	✔	-
Performance in School	-	✔	-
Education Level	-	-	✔
Economic Status			
Living Circumstances	-	-	✔
Income/Working Condition	-	-	✔
Social Conduct			
Administrative Criminal Records	-	-	✔
Law Breaking	-	✔	✔
Smoking, Drinking, and Drugs	-	-	✔

Return to main text

Tests of Exogeneity

Testing Assumption A–3: Exogeneity

- The following framework help us to test both Assumptions A–3 and A–4.
- Define an outcome vector as

$$Y_{k,a} = X_{k,a}^{d}\gamma + \varepsilon_{a}^{d} \qquad (a)$$

- With an associated measurement system

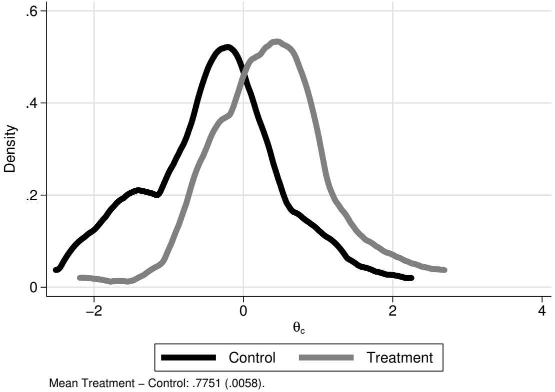
$$\begin{aligned} \boldsymbol{\varepsilon}_a^d &= \boldsymbol{\beta}^d \boldsymbol{\theta}_a^d + \boldsymbol{\omega}_a^d && (b) \\ M_a^d &= \boldsymbol{\lambda}^d \boldsymbol{\theta}_a^d + \boldsymbol{\eta}_{a^d}, && (c) \end{aligned} \tag{19}$$

- where $\boldsymbol{\theta}^d \perp\!\!\!\perp \boldsymbol{\eta}_a^d, \boldsymbol{\omega}_a^d$ and $\boldsymbol{\eta}_a^d \perp\!\!\!\perp \boldsymbol{\omega}_a^d$ for all $a \in \{0, \ldots, A\},\ d \in \{0, 1\}$.
- We use predictors in these equations.
- For sake of simplicity, we omit an explicit representation of them here.

- When the auxiliary measurement system M_a^d consists of at least three measures, we are able to identify the vectors of coefficients characterizing this system, $\boldsymbol{\lambda}^d, \boldsymbol{\beta}^d$, as well as the respective covariance matrices, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}_a^d}, \boldsymbol{\Sigma}_{\boldsymbol{\eta}_a^d}, \boldsymbol{\Sigma}_{\boldsymbol{\omega}_a^d}$, and use the method of Bartlett (1938) to obtain an estimate of $\boldsymbol{\theta}_a^d$ (Heckman et al., 2013).
- Identifying and estimating the elements in System (19) helps two purposes: (i) propose a test of Assumption A–3; and (ii) use estimates of $\boldsymbol{\theta}_a^d$ as control functions when testing Assumption A–4 in the next appendix, i.e. use these estimates to "control" for endogeneity.

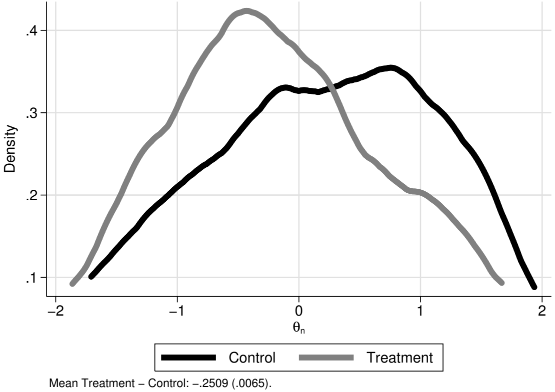
- We start by providing estimates for the elements in System (19) in the experimental sample.
- We assume dedicated measures for these skills at one time period. Put simply, we have two independent systems, one to measure $\boldsymbol{\theta}_c^d$ and one to measure $\boldsymbol{\theta}_n^d$, where $\boldsymbol{\theta}_a^d := \left[\boldsymbol{\theta}_c^d, \boldsymbol{\theta}_n^d\right]$.
- We use a set of IQ measures from ages 2 to 8 to obtain an estimate of $\boldsymbol{\theta}_c^d$ and a set of measures of somatization, hostility, depression, and mental health all at age 21 to measure to estimate $\boldsymbol{\theta}_n^d$.
- Figure 35 shows our estimates by treatment status.

Figure 35: Estimates of Cognitive (θ_c^d) Skills



Note: This figure displays a factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. "Less" in the factor measuring non-cognitive skills is "positive" given the measures we rely on to construct it. The mean difference between treatment and control is displayed below each panel, with standard error in parentheses. Standard errors are based on the empirical bootstrap distribution.

Figure 36: Estimates of Non-cognitive Skills (θ_n^d)



Note: This figure displays an analogous set of graphs for measures of somatization, hostility, depression, and mental health at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. "Less" in the factor measuring non-cognitive skills is "positive" given the measures we rely on to construct it. The mean difference between treatment and control is displayed below each panel, with standard error in parentheses. Standard errors are based on the empirical bootstrap distribution.

- We can also estimate $\boldsymbol{\theta}_a^d$ in the auxiliary sample.
- Once these estimates are available, we can test Assumption A–3 in the experimental and auxiliary samples.
- The test consists of the following.
- Let γ^E be the parameter associated to $X_{k,a}^d$ in Equation (a) in System (19) when not accounting for $\boldsymbol{\theta}_a^d$.
- Similarly, let γ^I be the parameter associated with $X_{k,a}^d$ in Equation (a) in System (19) when accounting for $\boldsymbol{\theta}_a^d$.
- Under the null hypotheses, Assumption A–3 holds and $\boldsymbol{\theta}_a^d$ is an irrelevant predictor in Equation (a) in System (19).
- This makes the OLS estimate of γ^E inconsistent.

- If the null hypotheses are false, $X^d_{k,a}$ and ε^d_a are not independent, γ^I is consistent and γ^E is not.
- We test the null hypothesis by asking if the elements in $\boldsymbol{\theta}^d_a$ are relevant predictors of a set of outcomes at age 30, so that we can perform the tests both the experimental and the auxiliary samples.
- We contrast specifications with and without including estimates of $\boldsymbol{\theta}^d_a$, and report the F-statistic corresponding to this comparison.
- This is a version of a Durbin-Wu-Hausman test (see Durbin, 1954; Hausman, 1978; Wu, 1973).
- Tables 22 to 25 present the results.
- In most cases, we are not able to reject the null hypothesis that Assumption A–3 holds.

Table 18: Prediction of Labor Income at Age 30 Accounting for B_k and $\boldsymbol{\theta}, X_{k,a}$, ABC/CARE Control Group

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
Mother's Education	1,599.57	0.17	867.41	0.34	-769.20	0.68	-580.88	0.62
PIAT (5-7)	.	.	.	.	45.98	0.41	423.44	0.20
Education (30)	.	.	.	.	3,415.53	0.03	4,505.94	0.04
Labor Income (21)	.	.	.	.	0.69	0.02	0.97	0.03
Cognitive	.	.	758.28	0.43	.	.	-8,009.28	0.93
Non Cognitive	.	.	-342.62	0.52	.	.	7,275.49	0.09
Constant	10,239.82	0.28	16,530.50	0.22	-23,140.28	0.80	-80,679.09	0.96
F-stat	2.27		1.80		11.89		7.91	
p-value	0.42		0.41		0.42		0.01	
R^2	0.03		0.07		0.30		0.40	
Observations	66		51		65		63	
F-stat: exclude Cognitive, Non-Cognitive		1.70				4.14		
p-value		0.45				0.09		

F-stat: exclude Cognitive and Non-Cognitive: *F*-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.
Note: Prediction of labor income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and *p*-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 19: Prediction of Labor Income at Age 30 Accounting for B_k and $\boldsymbol{\theta}, X_{k,a}$, ABC/CARE Treatment Group

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
Mother's Education	3,134.16	0.23	2,600.34	0.35	2,913.44	0.28	5,835.67	0.22
PIAT (5-7)	.	.	.	.	-263.29	0.66	-871.06	0.76
Education (30)	.	.	.	.	11,600.24	0.00	13,069.48	0.00
Labor Income (21)	.	.	.	.	-0.18	0.64	-0.62	0.75
Cognitive	.	.	2,766.35	0.40	.	.	4,828.93	0.34
Non Cognitive	.	.	7,600.33	0.18	.	.	6,223.32	0.19
Constant	3,900.73	0.47	10,553.93	0.42	-122,709.85	0.91	-109,410.81	0.76
F-stat	1.72		2.45		4.59		4.95	
p-value	0.38		0.21		0.38		0.06	
R^2	0.02		0.10		0.26		0.33	
Observations	64		49		65		63	
F-stat: exclude Cognitive, Non-Cognitive	2.49				2.03			
p-value	0.21				0.31			

F-stat: exclude Cognitive and Non-Cognitive: F-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.
Note: Prediction of labor income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 20: Prediction of Labor Income at Age 30 Accounting for B_k and $\boldsymbol{\theta}, X_{k,a}$, ABC/CARE Control and Treatment Groups

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
Mother's Education	2,668.48	0.12	2,200.35	0.25	794.11	0.36	1,724.88	0.31
PIAT (5-7)	.	.	.	.	-126.19	0.67	-400.57	0.72
Education (30)	.	.	.	.	8,601.33	0.00	9,706.02	0.00
Labor Income (21)	.	.	.	.	0.14	0.37	0.21	0.37
Cognitive	.	.	4,260.39	0.16	.	.	1,427.18	0.44
Non Cognitive	.	.	2,899.66	0.25	.	.	7,557.01	0.05
Constant	4,443.37	0.41	9,166.30	0.38	-78,053.28	0.95	-75,621.84	0.87
F-stat	2.50		1.90		5.87		5.37	
p-value	0.29		0.31		0.29		0.01	
R^2	0.02		0.04		0.20		0.25	
Observations	132		100		130		133	
F-stat: exclude Cognitive, Non-Cognitive	2.07				2.92			
p-value	0.31				0.19			

F-stat: exclude Cognitive and Non-Cognitive: F-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.
Note: Prediction of labor income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 21: Prediction of Labor Income at Age 30 Accounting for B_k and $\boldsymbol{\theta}, X_{k,a}$, CNLSY

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
Mother's Education	2,292.54	0.00	1,528.20	0.00	117.79	0.25	-47.06	0.50
PIAT (5-7)	.	.	.	.	262.38	0.00	447.61	0.00
Education (30)	.	.	.	.	3,722.75	0.00	4,202.69	0.00
Labor Income (21)	.	.	.	.	0.62	0.00	0.82	0.00
Cognitive	.	.	2,859.63	0.00	.	.	-4,149.19	0.88
Non Cognitive	.	.	-2,921.97	1.00	.	.	-590.26	0.75
Constant	2,840.27	0.00	11,377.06	0.00	-53,962.05	1.00	-78,072.63	1.00
F-stat	46.92		4.89		83.55		18.31	
p-value	0.00		0.04		0.00		0.00	
R^2	0.03		0.04		0.19		0.33	
Observations	1,862		350		1,860		1,862	
F-stat: exclude Cognitive, Non-Cognitive	4.18				1.77			
p-value	0.04				0.34			

F-stat: exclude Cognitive and Non-Cognitive: F-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.
Note: Prediction of labor income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and measures of reading and comprehension of the PIAT, as weel as the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior) (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 22: Prediction of Transfer Income at Age 30 Accounting for B_k and $\theta, X_{k,a}$, ABC/CARE Control Group

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
Mother's Education	-413.76	0.78	-406.14	0.69	48.75	0.49	51.97	0.47
PIAT (5-7)	.	.	.	.	27.10	0.29	-101.93	0.77
Education (30)	.	.	.	.	-684.75	0.99	-693.44	0.91
Labor Income (21)	.	.	.	.	-0.14	0.99	-0.15	0.93
Cognitive	.	.	-348.53	0.69	.	.	1,696.96	0.13
Non Cognitive	.	.	1,622.92	0.05	.	.	887.17	0.19
Constant	6,664.39	0.11	6,614.39	0.18	9,942.56	0.18	22,736.59	0.10
F-stat	1.93		2.96		3.53		2.68	
p-value	0.34		0.15		0.21		0.27	
R^2	0.04		0.15		0.21		0.27	
Observations	68		52		70		70	
F-stat: exclude Cognitive, Non-Cognitive			3.38				2.42	
p-value			0.19				0.27	

F-stat: exclude Cognitive and Non-Cognitive: F-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.
Note: Prediction of transfer income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\theta}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\theta}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 23: Prediction of Transfer Income at Age 30 Accounting for B_k and $\boldsymbol{\theta}, X_{k,a}$, ABC/CARE Treatment Group

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
Mother'sEducation	-212.39	0.75	-336.44	0.81	-199.39	0.74	-302.60	0.79
PIAT(5-7)	.	.	.	.	-46.36	0.86	-22.41	0.65
Education(30)	.	.	.	.	-35.62	0.56	-72.66	0.59
LaborIncome(21)	.	.	.	.	-0.05	0.94	-0.05	0.90
Cognitive	.	.	-421.59	0.75	.	.	-273.48	0.62
Non Cognitive	.	.	-825.26	0.95	.	.	-987.11	0.98
Constant	3,348.22	0.16	4,937.75	0.14	9,041.98	0.09	8,432.47	0.18
F-stat	1.23		2.59		1.81		2.27	
p-value	0.45		0.18		0.45		0.25	
R^2	0.03		0.15		0.13		0.25	
Observations	63		49		65		63	
F-stat: exclude Cognitive, Non-Cognitive	3.08				2.79			
p-value	0.20				0.18			

F-stat: exclude Cognitive and Non-Cognitive: *F*-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.

Note: Prediction of transfer income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and *p*-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 24: Prediction of Transfer Income at Age 30 Accounting for B_k and $\theta, X_{k,a}$, ABC/CARE Control and Treatment Groups

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
Mother's Education	-299.72	0.81	-411.12	0.85	-135.23	0.62	-211.76	0.68
PIAT (5-7)	.	.	.	.	-34.90	0.81	-66.99	0.80
Education (30)	.	.	.	.	-430.88	0.96	-453.82	0.96
Labor Income (21)	.	.	.	.	-0.09	1.00	-0.08	0.96
Cognitive	.	.	-753.98	0.93	.	.	153.54	0.42
Non Cognitive	.	.	631.74	0.17	.	.	264.49	0.34
Constant	5,135.83	0.06	6,460.15	0.07	13,548.68	0.03	17,791.02	0.05
F-stat	1.78		3.04		3.86		2.75	
p-value	0.35		0.12		0.35		0.09	
R^2	0.02		0.10		0.15		0.18	
Observations	133		101		135		133	
F-stat: exclude Cognitive, Non-Cognitive	3.38				1.23			
p-value	0.17				0.44			

F-stat: exclude Cognitive and Non-Cognitive: F-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.

Note: Prediction of transfer income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\theta}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\theta}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 25: Prediction of Transfer Income at Age 30 Accounting for B_k and $\theta, X_{k,a}$, CNLSY

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
Mother'sEducation	366.18	0.38	10,450.96	0.00	2,337.63	0.12	10,634.87	0.00
PIAT(5-7)	.	.	.	.	-872.86	0.88	-364.63	0.50
Education(30)	.	.	.	.	-8,126.93	1.00	-6,206.13	0.88
LaborIncome(21)	.	.	.	.	0.79	0.25	-0.99	1.00
Cognitive	.	.	-9,680.93	0.88	.	.	-5,092.70	0.50
NonCognitive	.	.	18,373.65	0.03	.	.	6,585.57	0.22
Constant	.	.	20,921.34	0.00	.	.	9,015.29	0.12
F-stat	0.14		1.80		1.06		1.17	
p-value	.75		0.18		0.49		0.37	
R^2	0.00		0.02		0.01		0.06	
Observations	1,101		239		1,100		1,099	
F-stat: exclude Cognitive, Non-Cognitive	1.70				0.71			
p-value	0.26				0.52			

F-stat: exclude Cognitive and Non-Cognitive: F-statistic contrasting the specifications in columns (1) and (3) and (5) and (7), respectively.
Note: Prediction of transfer income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value for the treatment and control groups and a test for the treatment-control difference. $\hat{\theta}_c$: factor score estimated based on the measurement system in (19) and measures of IQ at ages 2, 3, 4, 5, 7, and 8 (cognitive skill). $\hat{\theta}_n$: factor score estimated based on the measurement system in (19) and measures of somatization, hostility, depression, and a global mental health index at age 21 (non-cognitive skill). Both measures of skills are standardized to a mean of 0 and a standard deviation of 1. Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Return to main text

Testable Implications

Testing Assumption A–4: Structural Invariance

- We show that Assumption A–4, together with Assumption A–3, implies:

$$\mathbb{E}\left[Y^d_{k,j,a}|X^d_{k,a}=x, B_k=b, D=d\right]=\mathbb{E}\left[Y_{k,j,a}|X^d_{k,a}=x, B_k=b\right], \tag{20}$$

for $a \in \{1,\ldots,A\}$, $k \in \{e,n\}$, and $d \in \{0,1\}$.

- A direct test of this hypothesis is to use the experimental sample and ask if once we account for a set of the variables in $X_{k,a}$, R (randomization to treatment assignment in ABC/CARE, which, as discussed in text, is equivalent to D) predicts the outcome of interest, conditional on B_k.

Table 26: Prediction of High School Graduation at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Pooled Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	0.130	0.040	0.106	0.170	0.019	0.415	0.019	0.425
Mother's Education	0.093	0.000	0.081	0.000	0.051	0.000	0.043	0.050
PIAT (5-7)	.	.	.	.	-0.005	0.910	-0.004	0.765
Education (30)	.	.	.	.	0.119	0.000	0.124	0.000
Labor Income (21)	.	.	.	.	0.000	0.005	0.000	0.010
Cognitive	.	.	0.020	0.360	.	.	-0.038	0.740
Non Cognitive	.	.	-0.027	0.720	.	.	0.011	0.395
Constant	-0.410	0.985	-0.283	0.865	-1.082	1.000	-1.148	0.985
F-stat	14.497		5.819		40.509		25.147	
R^2	0.151		0.143		0.440		0.434	
Observations	134		102		135		133	

Note: Prediction of high school graduation at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 27: Prediction of High School Graduation at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Female Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	0.196	0.010	0.091	0.250	0.031	0.385	0.037	0.425
Mother's Education	0.077	0.000	0.062	0.050	0.036	0.055	0.023	0.170
PIAT (5-7)	.	.	.	.	-0.008	0.860	-0.004	0.655
Education (30)	.	.	.	.	0.093	0.000	0.102	0.005
Labor Income (21)	.	.	.	.	0.000	0.000	0.000	0.000
Cognitive	.	.	0.051	0.285	.	.	-0.073	0.775
Non Cognitive	.	.	-0.076	0.895	.	.	0.051	0.200
Constant	-0.266	0.800	-0.065	0.540	-0.444	0.790	-0.872	0.870
F-stat	10.180		5.545		35.887		31.753	
R^2	0.172		0.197		0.556		0.612	
Observations	68		53		70		70	

Note: Prediction of high school graduation at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 28: Prediction of High School Graduation at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Male Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	0.084	0.220	0.146	0.140	0.088	0.190	0.067	0.305
Mother's Education	0.116	0.000	0.131	0.000	0.057	0.045	0.072	0.095
PIAT (5-7)	.	.	.	.	0.000	0.470	-0.000	0.500
Education (30)	.	.	.	.	0.152	0.000	0.156	0.000
Labor Income (21)	.	.	.	.	0.000	0.170	0.000	0.440
Cognitive	.	.	-0.023	0.645	.	.	-0.009	0.525
Non Cognitive	.	.	0.051	0.195	.	.	0.004	0.475
Constant	-0.636	0.970	-0.824	0.950	-2.092	1.000	-2.227	0.955
F-stat	11.144		6.555		17.009		10.294	
R^2	0.190		0.215		0.467		0.460	
Observations	67		49		65		70	

Note: Prediction of high school graduation at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 29: Prediction of Employment at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Pooled Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	0.123	0.040	0.121	0.115	-0.008	0.550	0.057	0.265
Mother's Education	0.033	0.060	0.017	0.230	0.031	0.150	0.029	0.155
PIAT (5-7)	.	.	.	.	0.008	0.020	0.012	0.060
Education (30)	.	.	.	.	0.046	0.005	0.026	0.080
Labor Income (21)	.	.	.	.	-0.000	0.850	-0.000	0.875
Cognitive	.	.	0.077	0.060	.	.	-0.016	0.595
Non Cognitive	.	.	0.034	0.285	.	.	0.060	0.170
Constant	0.361	0.075	0.530	0.020	-0.877	0.975	-0.966	0.895
F-stat	3.903		4.073		5.239		3.979	
R^2	0.057		0.124		0.177		0.229	
Observations	133		101		135		133	

Note: Prediction of employment at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 30: Prediction of Employment at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Female Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	0.124	0.135	0.003	0.495	-0.078	0.705	-0.092	0.745
Mother's Education	-0.000	0.520	-0.000	0.500	-0.012	0.670	-0.000	0.500
PIAT (5-7)	.	.	.	.	0.010	0.030	0.008	0.145
Education (30)	.	.	.	.	0.040	0.035	0.030	0.085
Labor Income (21)	.	.	.	.	0.000	0.260	-0.000	0.520
Cognitive	.	.	0.151	0.005	.	.	0.065	0.240
Non Cognitive	.	.	-0.027	0.655	.	.	0.019	0.425
Constant	0.702	0.000	0.754	0.000	-0.624	0.865	-0.359	0.655
F-stat	1.873		5.089		3.432		3.918	
R^2	0.048		0.207		0.229		0.289	
Observations	67		52		65		70	

Note: Prediction of employment at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 31: Prediction of Employment at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Male Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	0.132	0.065	0.228	0.035	0.083	0.220	0.271	0.005
Mother's Education	0.066	0.020	0.045	0.150	0.094	0.020	0.116	0.000
PIAT (5-7)	.	.	.	.	0.008	0.090	0.022	0.015
Education (30)	.	.	.	.	0.023	0.235	-0.009	0.635
LaborIncome (21)	.	.	.	.	-0.000	0.990	-0.000	0.940
Cognitive	.	.	-0.030	0.665	.	.	-0.180	0.970
Non Cognitive	.	.	0.110	0.020	.	.	0.138	0.030
Constant	-0.002	0.500	0.203	0.350	-1.202	0.940	-2.416	0.975
F-stat	4.050		3.140		3.899		5.322	
R^2	0.114		0.192		0.240		0.443	
Observations	66		49		65		63	

Note: Prediction of employment at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 32: Prediction of Labor Income at Age 30 Accounting for R, B_k, $\boldsymbol{\theta}$, and $X_{k,a}$ Pooled Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	10576.303	0.065	11,165.829	0.125	283.356	0.490	1,836.270	0.410
Mother's Education	1,851.130	0.205	1,131.843	0.375	496.581	0.430	1,052.668	0.365
PIAT (5-7)					-81.009	0.595	-320.784	0.705
Education (30)					8,097.138	0.000	9,141.309	0.000
Labor Income (21)					0.130	0.330	0.192	0.325
Cognitive			2,308.860	0.305			785.891	0.465
Non Cognitive			2,665.092	0.190			6,876.181	0.065
Constant	7,067.552	0.405	14,188.359	0.340	-73,300.00	0.965	-70,500.00	0.920
F-stat	1.965		1.522		5.746		4.742	
R^2	0.031		0.056		0.210		0.251	
Observations	132.000		101.000		130.000		133.000	

Note: Prediction of employment at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 33: Prediction of Labor Income at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Female Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	3,401.892	0.305	1,194.706	0.410	-6,899.006	0.915	-5,862.320	0.840
Mother's Education	-852.061	0.675	-1,688.467	0.835	-2,581.049	0.975	-2,473.902	0.965
PIAT (5-7)					260.764	0.165	347.907	0.170
Education (30)					3,580.642	0.000	3,916.084	0.005
LaborIncome (21)					0.329	0.175	0.392	0.160
Cognitive			3,828.286	0.130			-2,905.637	0.785
Non Cognitive			-1,663.392	0.655			2,051.882	0.300
Constant	32,117.510	0.055	39,943.031	0.025	-21,500.00	0.800	-36,600.00	0.840
F-stat	1.234		2.812		9.052		8.916	
R^2	0.039		0.143		0.354		0.393	
Observations	67		52		65		70	

Note: Prediction of labor income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 34: Prediction of Labor Income at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Male Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	18169.158	0.075	21,891.223	0.150	15,649.704	0.115	18,835.850	0.185
Mother's Education	5,722.000	0.090	6,064.495	0.260	4,618.608	0.155	8,200.867	0.160
PIAT (5-7)					459.787	0.180	1,828.085	0.110
Education (30)					15,803.528	0.000	22,139.904	0.015
Labor Income (21)					0.107	0.410	0.193	0.365
Cognitive			-896.956	0.525			-13,700	0.815
Non Cognitive			10,273.761	0.105			7,533.493	0.175
Constant	-31,600.00	0.780	-34,800.00	0.630	-272,000.00	0.985	-526,000.00	0.965
F-stat	2.327		1.963		4.833		7.182	
R^2	0.068		0.128		0.343		0.465	
Observations	66		48		65		63	

Note: Prediction of labor income at age 30 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 35: Prediction of Body-Mass Index at Age 34 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Pooled Sample, ABC/CARE

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
R	1.027	0.270	2.864	0.150	1.213	0.250	3.367	0.090
Mother's Education	-0.130	0.615	-0.116	0.560	0.003	0.500	-0.273	0.665
PIAT (5-7)					0.076	0.260	0.277	0.060
Education (30)					-0.116	0.575	-0.295	0.610
Labor Income (21)					0.000	0.290	0.000	0.095
Cognitive			-1.675	0.935			-3.431	0.960
Non Cognitive			1.615	0.195			2.392	0.100
Constant	34.913	0.000	33.909	0.000	26.682	0.070	9.604	0.330
F-stat	1.366		2.612		1.663		2.830	
R^2	0.027		0.110		0.090		0.209	
Observations	87		66		85		84	

Note: Prediction of body-mass index at age 34 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and *p*-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 36: Prediction of Body-Mass Index at Age 34 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Female Sample, ABC/CARE

	(1) Estimate	(2) *p*-value	(3) Estimate	(4) *p*-value	(5) Estimate	(6) *p*-value	(7) Estimate	(8) *p*-value
R	3.675	0.110	7.167	0.035	4.623	0.090	6.526	0.020
Mother's Education	-0.148	0.580	-0.654	0.820	-0.492	0.715	-0.909	0.835
PIAT (5-7)					-0.119	0.775	0.040	0.440
Education (30)					0.238	0.445	0.269	0.435
Labor Income (21)					0.000	0.340	0.000	0.385
Cognitive			-2.171	0.925			-2.366	0.815
Non Cognitive			2.285	0.155			2.536	0.145
Constant	36.244	0.000	39.310	0.000	46.750	0.020	33.957	0.075
F-stat	1.837		3.151		2.442		3.206	
R^2	0.065		0.206		0.191		0.285	
Observations	51		41		50		49	

Note: Prediction of body-mass index at age 34 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and *p*-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 37: Prediction of Body-Mass Index at Age 34 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Male Sample, ABC/CARE

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
R	-0.189	0.510	1.262	0.370	-0.397	0.545	1.150	0.400
Mother's Education	-0.513	0.805	0.448	0.380	-0.091	0.510	1.074	0.215
PIAT (5-7)					0.224	0.050	0.651	0.075
Education (30)					0.445	0.250	1.482	0.220
Labor Income (21)					0.000	0.165	0.000	0.100
Cognitive			-1.677	0.800			-4.854	0.920
Non Cognitive			0.119	0.475			0.563	0.380
Constant	36.443	0.000	26.285	0.030	3.330	0.455	-64.561	0.885
F-stat	1.835		2.330		5.387		31.866	
R^2	0.076		0.180		0.230		0.504	
Observations	37		25		35		35	

Note: Prediction of body-mass index at age 34 based on the variables listed in the row. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

- We show that Assumption A–4, together with Assumption A–3, implies equality of conditional expectations in the experimental and auxiliary samples.

$$\mathbb{E}\left[Y_{e,j,a}|X^d_{e,a}=x, B_e=b\right] = \mathbb{E}\left[Y_{n,j,a}|X^d_{n,a}=x, B_e=b\right], \quad d \in \{0,1\}, \quad j \in \mathcal{J}_a. \tag{21}$$

- We test this hypothesis at $a=a^*$, where we observe the predicted outcomes at ages 30.
- Our non-experimental source at $a=a^*$ is the CNLSY.

Table 38: Prediction of Labor Income at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Female Sample, ABC/CARE and CNLSY

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
K^*	4,396.848	0.060	2,292.980	0.250	456.565	0.410	539.899	0.445
Mother's Education	289.800	0.400	-1253.548	0.800	-1878.064	0.985	-2,126.096	0.960
PIAT (5-7)					207.361	0.090	221.599	0.215
Education (30)					3381.137	0.000	3652.225	0.000
Labor Income (21)					0.345	0.020	0.366	0.050
Cognitive			4,078.844	0.055			-,1479.220	0.670
Non Cognitive			-1370.089	0.640			2229.399	0.195
Constant	1,7358.422	0.100	33,633.047	0.030	-25,100.00	0.960	-27,400.00	0.840
F-stat	1.924		2.882		13.153		9.163	
R^2	0.022		0.106		0.279		0.312	
Observations	382		128		380		385	

* $K = 1$ if $k = e$; $K = 0$ if $k = n$.

Note: Prediction of labor income at age 30 based on the variables listed in the row using the ABC/CARE and the CNLSY sample constructed according to our procedure. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) in CNLSY and IQ at ages 2, 3, 4, 5, 7, and 8 in ABC/CARE (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior) in CNLSY and measures of somatization, hostility, depression, and a global mental health index at age 21 in ABC/CARE (non-cognitive skill). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 39: Prediction of Labor Income at Age 30 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Male Sample, ABC/CARE and CNLSY

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
K^*	17984.393	0.010	21406.389	0.015	4864.750	0.260	3301.140	0.305
Mother's Education	4,182.211	0.035	2,885.837	0.295	1,991.183	0.150	3,960.881	0.210
PIAT (5-7)					13.463	0.480	608.659	0.210
Education (30)					11,855.479	0.000	18,995.199	0.010
Labor Income (21)					0.289	0.165	0.243	0.260
Cognitive			5,012.976	0.205			-1498.498	0.560
Non Cognitive			6,902.538	0.115			6,335.481	0.070
Constant	-23,300.00	0.805	-1.13e+04	0.575	-1.50e+05	0.985	-318,000.00	0.965
F-stat	4.333		2.187		9.588		8.790	
R^2	0.059		0.087		0.283		0.403	
Observations	312		102		310		315	

* $K = 1$ if $k = e$; $K = 0$ if $k = n$.

Note: Prediction of labor income at age 30 based on the variables listed in the row using the ABC/CARE and the CNLSY sample constructed according to our procedure. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) in CNLSY and IQ at ages 2, 3, 4, 5, 7, and 8 in ABC/CARE (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior) in CNLSY and measures of somatization, hostility, depression, and a global mental health index at age 21 in ABC/CARE (non-cognitive skill). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 40: Prediction of Body-Mass Index at Age 34 Accounting for $R, B_k, \boldsymbol{\theta}$, and $X_{k,a}$ Female Sample, ABC/CARE and CNLSY

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
K^*	4.380	0.000	3.620	0.015	4.538	0.020	3.731	0.115
Mother's Education	-0.110	0.585	-0.273	0.675	-0.225	0.705	-0.433	0.735
PIAT (5-7)					-0.006	0.530	0.076	0.285
Education (30)					0.001	0.500	0.337	0.420
Labor Income (21)					0.000	0.315	-0.000	0.525
Cognitive			-0.480	0.680			-0.773	0.705
Non Cognitive			0.858	0.255			0.805	0.275
Constant	32.921	0.000	34.948	0.000	34.288	0.000	25.174	0.085
F-stat	6.255		3.312		3.929		2.370	
R^2	0.075		0.110		0.122		0.167	
Observations	366		117		365		364	

* $K = 1$ if $k = e$; $K = 0$ if $k = n$.

Note: Prediction of body-mass index at age 34 based on the variables listed in the row using the ABC/CARE and the CNLSY sample constructed according to our procedure. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) in CNLSY and IQ at ages 2, 3, 4, 5, 7, and 8 in ABC/CARE (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior) in CNLSY and measures of somatization, hostility, depression, and a global mental health index at age 21 in ABC/CARE (non-cognitive skill). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Table 41: Prediction of Body-Mass Index at Age 34 Accounting for R, B_k, θ, and $X_{k,a}$ Male Sample, ABC/CARE and CNLSY

	(1) Estimate	(2) p-value	(3) Estimate	(4) p-value	(5) Estimate	(6) p-value	(7) Estimate	(8) p-value
K^*	2.327	0.025	2.812	0.065	1.822	0.100	3.135	0.060
Mother's Education	-0.180	0.700	0.347	0.320	-0.029	0.515	0.518	0.180
PIAT (5-7)					0.080	0.085	0.236	0.050
Education (30)					0.161	0.300	0.399	0.270
Labor Income (21)					0.000	0.080	0.000	0.020
Cognitive			-1.270	0.835			-2.362	0.970
Non Cognitive			0.188	0.385			0.482	0.265
Constant	30.350	0.000	25.170	0.000	18.323	0.010	-6.799	0.600
F-stat	2.828		3.562		3.171		3.867	
R^2	0.050		0.142		0.096		0.260	
Observations	283		79		285		280	

* $K = 1$ if $k = e$; $K = 0$ if $k = n$.

Note: Prediction of body-mass index at age 34 based on the variables listed in the row using the ABC/CARE and the CNLSY sample constructed according to our procedure. Empty cells indicate that the variable was not used in the prediction. For each coefficient we provide point estimate and p-value. $\hat{\boldsymbol{\theta}}_c$: factor score estimated based on the measurement system in (19) and PIAT sections that we do not use to predict (reading and comprehension) as well as by the Peabody Picture Vocabulary Test (PPVT) in CNLSY and IQ at ages 2, 3, 4, 5, 7, and 8 in ABC/CARE (cognitive skill). $\hat{\boldsymbol{\theta}}_n$: factor score estimated based on the measurement system in (19) and on six scales of the Behavior Problems Index (e.g., anxiety, dependency, social behavior) in CNLSY and measures of somatization, hostility, depression, and a global mental health index at age 21 in ABC/CARE (non-cognitive skill). Inference is based on the empirical bootstrap distribution. If the estimates for the constant terms are in the ten or hundred thousands, we report a figure that has been rounded to the thousands.

Return to main text

Treatment Effects Accounting Correcting the p-values Using Step-down

Table 42: Treatment Effects on Parental Labor Income, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Parental Labor Income							
	3.5	1,036	-1,185	3,085	1,452	112	703
		(0.604)	(0.534)	(1.000)	(0.759)	(0.672)	(0.657)
	12	7,085	10,384	15,563	12,682	4,773	5,411
		(0.283)	(0.210)	(1.000)	(0.429)	(0.465)	(0.476)
	15	8,488	7,185	6,697	4,915	7,603	4,379
		(0.283)	(0.412)	(1.000)	(0.759)	(0.407)	(0.657)
	21	12,732	12,650	1,568	-1,000	15,124	10,323
		(0.091)	(0.299)	(1.000)	(0.759)	**(0.003)**	(0.204)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 43: Treatment Effects on Education, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Graduated High School	30	0.073	0.130	0.114	0.084	0.077	0.063
		(0.629)	(0.377)	(1.000)	(0.763)	(0.568)	(0.539)
Graduated 4-year College	30	0.170	0.178	0.124	0.100	0.179	0.142
		(0.272)	(0.359)	(1.000)	(0.763)	(0.300)	(0.408)
Years of Edu.	30	0.525	0.785	0.857	0.782	0.385	0.343
		(0.540)	(0.351)	(1.000)	(0.514)	(0.568)	(0.539)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped p-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 44: Treatment Effects on Subject Employment and Income, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Labor Income	30	19,810	27,373	17,909	23,950	20,065	21,068
		(0.346)	(0.441)	(1.000)	(0.316)	(0.308)	(0.451)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 45: Treatment Effects on Crime, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Total Felony Arrests	Mid-30s	0.196	0.392	0.946	1.338	0.096	0.184
		(0.439)	(0.508)	(1.000)	(0.135)	(0.522)	(0.446)
Total Misdemeanor Arrests	Mid-30s	-0.501	-0.243	-0.251	-0.033	-0.666	-0.508
		(0.381)	(0.508)	(1.000)	(0.439)	(0.301)	(0.382)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped p-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 46: Treatment Effects on Tobacco, Drugs, Alcohol, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Self-reported drug user	Mid-30s	-0.333	-0.398	-0.500	-0.557	-0.309	-0.330
		(0.114)	**(0.054)**	(1.000)	(0.143)	(0.397)	(0.183)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 47: Treatment Effects on Hypertension, Male Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Systolic Blood Pressure (mm Hg)	Mid-30s	-9.791	-13.511	15.280	14.979	-19.920	-18.537
		(0.193)	(0.131)	(0.979)	**(0.001)**	**(0.040)**	**(0.029)**
Diastolic Blood Pressure (mm Hg)	Mid-30s	-10.854	-16.689	-8.640	-8.741	-14.240	-13.988
		(0.090)	**(0.002)**	(0.979)	**(0.079)**	**(0.004)**	**(0.029)**
Hypertension	Mid-30s	-0.291	-0.352	-0.053	-0.075	-0.420	-0.435
		(0.102)	(0.117)	(0.979)	(0.573)	**(0.031)**	**(0.029)**

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 48: Treatment Effects on Parental Labor Income, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Parental Labor Income							
	3.5	2,756	3,277	5,146	8,601	2,802	3,762
		(0.456)	(0.461)	(1.000)	(0.141)	(0.551)	(0.487)
	12	13,633	19,386	22,294	26,474	11,570	18,629
		(0.313)	(0.224)	(1.000)	**(0.034)**	(0.399)	(0.115)
	15	8,565	9,322	2,829	8,435	9,819	10,480
		(0.313)	(0.392)	(1.000)	(0.579)	(0.166)	(0.232)
	21	5,708	6,944	25,270	25,135	4,446	3,926
		(0.438)	(0.507)	(1.000)	**(0.002)**	(0.551)	(0.507)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 49: Treatment Effects on Education, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Graduated High School	30	0.253	0.110	0.642	0.596	0.137	0.066
		(0.077)	(0.580)	(1.000)	**(0.003)**	(0.479)	(0.661)
Graduated 4-year College	30	0.134		0.217	0.219	0.106	0.094
		(0.292)		(1.000)	**(0.040)**	(0.479)	(0.630)
Years of Edu.	30	2.143	1.715	4.025	3.925	1.567	1.412
		(0.004)	**(0.007)**	(1.000)	**(0.001)**	**(0.064)**	(0.120)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

THE UNIVERSITY OF CHICAGO

Table 50: Treatment Effects on Subject Employment and Income, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Labor Income							
	30	2,548	2,412	14,356	14,862	-425	-822
		(0.393)	(0.628)	(1.000)	**(0.059)**	(0.637)	(0.630)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 51: Treatment Effects on Crime, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Total Felony Arrests	Mid-30s	-0.328	-0.394	-1.345	-0.965	-0.077	0.005
		(0.153)	(0.179)	(1.000)	(0.178)	(0.245)	(0.488)
Total Misdemeanor Arrests	Mid-30s	-0.973	-1.212	-2.708	-2.448	-0.588	-0.201
		(0.131)	(0.179)	(1.000)	(0.192)	(0.232)	(0.389)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped p-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 52: Treatment Effects on Tobacco, Drugs, Alcohol, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Self-reported drug user	Mid-30s	-0.033	-0.039	-0.116	-0.101	-0.010	0.033
		(0.683)	(0.622)	(0.999)	(0.557)	(0.707)	(0.770)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped p-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Table 53: Treatment Effects on Hypertension, Female Sample

Variable	Age	(1)	(2)	(3)	(4)	(5)	(6)
Systolic Blood Pressure (mm Hg)	Mid-30s	-2.899	-4.316	1.065	-0.827	-3.980	-6.805
		(0.419)	(0.410)	(0.998)	(0.481)	(0.356)	(0.334)
Diastolic Blood Pressure (mm Hg)	Mid-30s	-0.002	1.323	4.725	4.120	-1.291	-2.186
		(0.485)	(0.496)	(0.998)	(0.481)	(0.401)	(0.406)
Hypertension	Mid-30s	0.172	0.151	0.232	0.162	0.156	0.107
		(0.207)	(0.340)	(0.998)	(0.481)	(0.276)	(0.406)

Note: This table presents estimates for the treatment effects for each of the variables listed in the rows. At the bottom of each table we also present treatment effects on a factor estimated using the method in our Appendix and the measures indicated in the table. One-tailed, bootstrapped *p*-values are in parentheses. We adjust them for multiple hypotheses testing using the step-down algorithm in Romano and Wolf (2016).

Return to main text

Exploring the Impact of Using Different Prediction Models

- Our analysis is based on a causal model for treatment ($d = 1$) and control ($d = 0$) outcomes for measure j at age a in sample $k \in \{e, n\}$ where e denotes membership in the experimental sample and n denotes membership in the auxiliary sample:

$$Y^d_{k,j,a} = \phi^d_{k,j,a}(X^d_{k,a}, B_k) + \varepsilon^d_{k,j,a}, \quad k \in \{n, e\}, \quad j \in \mathcal{J}_a, \quad d \in \{0, 1\}. \tag{22}$$

- $\phi^d_{k,j,a}(\cdot, \cdot)$ is an invariant structural production relationship mapping inputs $X^d_{k,a}, B_k$ into output $Y^d_{k,j,a}$ holding error term $\varepsilon^d_{k,j,a}$ fixed.

- In general, we let

$$\begin{aligned} \varepsilon^d_{k,j,a} &= f^d + \omega^d_{k,j,a} \\ \omega^d_{k,j,a} &= \rho^d \omega^d_{k,j,a-1} + U^d_{k,j,a}, \end{aligned} \tag{23}$$

$U^d_{k,j,a} \perp\!\!\!\perp X^d_{k,a}$.

- In this appendix, we present different structures for $\phi^{d}_{k,j,a}(\cdot,\cdot)$ and $\varepsilon^{d}_{k,j,a}$ and investigate the robustness of our estimates to different assumptions about the structure of both these elements.
- We do this exercise for labor income.
- We describe the precise steps that we follow to construct out-of-sample predictions based on these different structures and frame our estimations in a general method of moments framework.

- Note that Assumption A–4 (Invariance) implies that $\phi^{d}_{k,j,a}(\cdot,\cdot) = \phi_{k,j,a}(\cdot,\cdot) = \phi_{j,a}(\cdot,\cdot)$.
- That is, invariance holds across the treatment and the control groups and invariance holds across the experimental and the auxiliary samples.
- It is important to note that invariance across the treatment and the control groups implies that the variables $X^{d}_{k,a}$ summarize the effect of the treatment on the outcome.
- Given this and Assumption A–3 (Exogeneity), the distribution of $\varepsilon^{d}_{k,j,a}$ is the same across the treatment and the control groups.
- We then drop the superscript in $\varepsilon^{d}_{k,j,a}$.

- We test invariance across the treatment and the control groups and invariance across the experimental and the auxiliary samples.
- We also document that the support of $Y^d_{n,j,a}, X^d_{n,a}, B_n$ covers the support of $Y^d_{e,j,a}, X_{e,a}, B_e$ for $d \in \{0, 1\}$.
- We drop the d superscript in $Y^d_{n,j,a}, X^d_{n,a}$ given that we estimate an invariant model.

- In our empirical analysis, we work with a linear specification of $\phi_{j,a}$.
- We explore different alternatives for prediction under this specification as well.
- The system of interest is:

$$
\begin{aligned}
Y_{k,j,a} &= \lambda_0 + \lambda_1 Y_{k,j,a-1} + \lambda_2 X_{n,a} + \varepsilon_{k,j,a} \\
\varepsilon_{k,j,a} &= \underbrace{f}_{\text{Fixed Effect}} + \underbrace{\omega_{k,j,a}}_{\substack{\text{Possibly Serially} \\ \text{Correlated Component}}} \\
\omega_{k,j,a} &= \rho\omega_{k,j,a-1} + \underbrace{U_{k,j,a}}_{\substack{\text{Independent} \\ \text{Innovation}}} ,
\end{aligned}
\tag{24}
$$

- $U_{k,j,a} \perp\!\!\!\perp X_{k,a}$.

- Table 70 summarizes the results from our exploration through two statistics:
  - (i) the net present value (discounted to birth treatment - control) of predicted labor income under different assumptions; and
  - (ii) the overall cost-benefit ratio when the predictions are done based on the different proposed alternatives.
- The results indicate that the model that we base our predictions on in the main text has little sensitivity to the deviations that we propose.

Table 54: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Predictions

	Specification 1: $\lambda_1 \neq 0$, $\rho = 0$, $f = 0$			Specification 2: $\lambda_1 = 0$, $\rho \neq 0$, $f = 0$		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	71,345	0.13	6.29	154,547	0.26	12.39
	(86,343)	(.05)	(2.11)	(187,036)	(0.11)	(5.16)
Males	300,896	0.13	11.1	200,509	0.09	7.62
	(241,588)	(0.06)	(6.35)	(160,988)	(0.04)	(3.73)
Females	59,390	0.10	2.45	79,441	0.15	3.61
	(63,060)	(0.07)	(0.79)	(99,416)	(0.11)	(1.56)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the four different specifications for prediction that are explained below. Specification 1: prediction based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: prediction based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: prediction based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: prediction based on lagged outcome; no serial autocorrelation; and fixed effect.

Table 54: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Predictions

	Specification 3: $\boldsymbol{\lambda_1} \neq 0$, $\boldsymbol{\rho} \neq 0$, $f = 0$			Specification 4: $\boldsymbol{\lambda_1} = 0$, $\boldsymbol{\rho} = 0$, $f \neq 0$		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	268,179	0.49	23.64	46,953	0.09	4.14
	(211,089)	(0.12)	(5.16)	(25,323)	(0.01)	(0.62)
Males	456,078	0.2	16.82	74,775	0.03	2.76
	(358,534)	(0.09)	(9.42)	(54,752)	(0.01)	(1.44)
Females	31,303	0.05	1.29	19,959	0.03	0.82
	(168,160)	(0.19)	(2.11)	(34,142)	(0.04)	(0.43)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the four different specifications for prediction that are explained below. Specification 1: prediction based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: prediction based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: prediction based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: prediction based on lagged outcome; no serial autocorrelation; and fixed effect.

Table 54: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Predictions

	Specification 5: Non-parametric		
	NPV	IRR	B/C
Pooled	62,080	0.10	4.98
	(75,030)	(0.03)	(2.07)
Males	289,471	0.13	11.01
	(232,471)	(0.06)	(5.39)
Females	59,163	0.11	2.69
	(74,039)	(0.08)	(1.16)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the four different specifications for prediction that are explained below. Specification 1: prediction based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: prediction based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: prediction based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: prediction based on lagged outcome; no serial autocorrelation; and fixed effect.

- In the auxiliary sample, we observe the outcome $Y_{n,j,a}$ for $a \in [a^*, \ldots, A]$.
- In the experimental sample, we observe the outcome $Y_{e,j,a}$ for at most two ages, depending on the outcome.
- For the time being, suppose that we observe the outcome at one age ($a = a^*$).
- We return to this point below.
- By out-of-sample predictions we mean using the information in the auxiliary sample at $a \in [a^*, \ldots, A]$ to form extrapolations in the experimental sample, where we do not observe the outcome of interest during this age periods.
- We produce out-of-sample predictions and calculate the net present value of labor income (treatment - control) under different assumptions.

Specification 1: Lagged Component ($\lambda_1 \neq 0$); No Serial Correlation ($\rho = 0$); and No Fixed Effect ($f = 0$)

1. The predictions across the paper are constructed under this framework: labor and transfer income, crime, and health.
2. The comparisons between realizations and predictions are close.
3. Additional tests show the following.
4. We fail to reject: invariance across the treatment and the control groups, invariance across the experimental and auxiliary samples, and we fail to reject exogeneity both in the experimental and the auxiliary samples.
5. The tests are at $a = a^*$.

Specification 2: No Lagged Component ($\lambda_1 = 0$); Serial Correlation ($\rho \neq 0$); and No Fixed Effect ($f = 0$)

1. Given that $Y_{k,j,a-1}$ is not one of the elements in $X_{k,a}$, Assumption A–3 (Exogeneity) holds even when we do not restrict ρ.
2. It is straightforward to account for serial correlation in this case: serial correlation is a particular case of arbitrary heteroskedasticity.
3. We do not even need to take a stand on the serial autocorrelation structure.
4. We can simply invoke the assumption of $Y_{k,j,a}$ not being one of the elements in $X_{k,a}$ and proceed to account for arbitrary forms of heteroskedasticity.

Specification 2: No Lagged Component ($\lambda_1 = 0$); Serial Correlation ($\rho \neq 0$); and No Fixed Effect ($f = 0$)

5. The predictions in the paper are extremely similar in this case.
6. That is, the lag does not help the predictions as much as we would initially think.
7. This is more evidence in favor of $X_{k,a}$ summarizing the treatment effects.

Specification 3: Lagged Component ($\lambda_1 \neq 0$); Serial Correlation ($\rho \neq 0$); and No Fixed Effect ($f = 0$)

1. Is serial correlation present in the data?
2. The estimates indicate that it is.
3. From ages 21 to 30 we estimate the model in the CNLSY and the estimate for ρ is .7465.
4. From ages 30 to 67 (assumed retirement) we estimate the model in the NLSY79/PSID and the estimate for ρ is .5426.
5. When we restrict the sample to people who earn 30,000 at each of these ages, the analogous estimates of ρ are .7370 and .5316.
6. These estimates are statistically significant at the 1% level.
7. We could invoke more formal tests, but with the size of the point estimates and their precision, we will never fail to reject the null of no autocorrelation.
8. We present details on this estimation below.

- We can ρ-transform the system of interest to obtain consistent estimates as follows.
- We drop the j index for simplicity and write:

$$Y_{k,a} = \lambda_0 (1-\rho) + (\lambda_1 + \rho) Y_{k,a-1} - \lambda_1 \rho Y_{k,a-2} + \lambda_2 (X_{k,a} - \rho X_{k,a-1}) + U_{k,a} \tag{25}$$

- OLS produces consistent estimates of the coefficients.
- This enables us to construct predictions, as we explain in Appendix 116, as the transformed model has very similar features to Specification 1 (lagged component—in this case two lagged components— and no serial correlation—by construction).

Specification 4:

- Permanent-Transitory Decomposition of Unobserved Components ($\lambda_1 \neq 0$; $\rho = 0$; $f \neq 0$)

Specification 5: Non-Parametric Predictions

- An alternative to any of these scenarios is to form non-parametric predictions.
- That is:
  (i) for each individual i in the experimental sample, e, find an individual(s) $I(i)$ in the non-experimental sample, n, using Algorithm 1;
  (ii) impute the post-a^* trajectory of $Y_{k,j,a}$ of individual(s) $I(i)$ in the non-experimental sample, n, to individual i in the experimental sample, e.

THE UNIVERSITY OF CHICAGO

Return to main text

Estimation Procedure and Data Combination Estimator in the GMM Framework

Use Standard Panel Data Methods: Example

- $Y_a = \lambda_0 + \lambda_1 \; Y_{a-1} + \lambda_2 \; X_a + \varepsilon_a$
- $\varepsilon_a = \underbrace{f}_{\text{fixed effect}} + \underbrace{\omega_a}_{\text{serially correlated process}}$
- $\omega_a = \rho \; \omega_{a-1} + U_a$
- $U_a \perp\!\!\!\perp X_a, Y_{a-1}$

Standard Methods: Example

$$
\begin{aligned}
Y_a &= \lambda_0 + \lambda_1 \; Y_{a-1} + \lambda_2 \; X_a + f + \rho \; \omega_{a-1} + U_a \\
\Delta Y_a &= (\lambda_1 + \rho) \; \Delta Y_{a-1} - \rho \lambda_1 \; \Delta Y_{a-2} \\
&\quad + \lambda_2 \Delta (X_a - \rho X_{a-1}) + \Delta U_a
\end{aligned}
$$

- Form moment conditions in usual way
  (e.g., instrument using lagged values of Y_a)

Pool Across Samples

- Use initial conditions from experimental sample

- Let W be a positive definite matrix. We estimate $\boldsymbol{\theta}$ by minimizing

$$Q := \begin{bmatrix} \bar{m}\left(\cdot;\boldsymbol{\theta}\right) \\ \bar{h}\left(\cdot;\boldsymbol{\theta}\right) \end{bmatrix}' W^{-1} \begin{bmatrix} \bar{m}\left(\cdot;\boldsymbol{\theta}\right) \\ \bar{h}\left(\cdot;\boldsymbol{\theta}\right) \end{bmatrix}, \tag{26}$$

- $\bar{m}$ denotes suitably transformed version of outcome equations.
- $\bar{u}$ denotes the empirical counterpart of ε vector is determined from data.
- W is not restricted to be diagonal so that these moments are allowed to be correlated (but see Altonji and Segal, 1996).
- Iterated, feasible procedures to obtain an estimate of W jointly with the parameters of interest guarantee efficiency and are straightforward to implement.

1. Use the auxiliary sample (n) to estimate the the coefficients characterizing $\phi_{j,a}(\cdot, \cdot)$.
   - In practice, we use a weighted version of the auxiliary samples.
   - The weights give relatively high importance to the individuals in the auxiliary sample whose characteristics B_k are close to the those of the individuals in the experimental sample.

- We denote these coefficients by $\boldsymbol{\theta}_{j,a}$ and the estimate of this function as $\hat{\phi}_{j,a}(\cdot,\cdot)$.
- At each age, we are able to compute the residuals from this estimation procedure as follows:

$$Y_{n,j,a} - \hat{\phi}_{j,a}(X_{k,a}, B_k) := \hat{\varepsilon}_{n,j,a}. \quad (27)$$

- For outcome j, we form the vector of residuals $\hat{\boldsymbol{\varepsilon}}_{n,j} := [\hat{\varepsilon_{n,j,a^*+1}}, \ldots, \hat{\varepsilon_{n,j,A}}]$.

Initially Assume Independence

2. At age $a^* + 1$, we construct the predicted outcome for the experimental sample (e) for each individual as follows:

$$\hat{Y}_{e,j,a^*+1} = \hat{\phi}_{j,a^*+1}\left(X_{e,a^*+1}, B_e\right). \tag{28}$$

- We are able to evaluate $\hat{\phi}_{j,a^*+1}$ at X_{e,a^*+1}, B_e even when X_{e,a^*+1} contains a one-period lag of Y_{e,j,a^*+1} because we observe Y_{e,j,a^*}.
- This prediction does not account for estimation error.
- We discuss estimation error below.

3. At age $a^* + 2$, we construct the predicted outcome in the experimental sample (e) as follows:

$$\hat{Y}_{e,j,a^*+2} = \hat{\phi}_{j,a^*+1}\left(X_{e,a^*+1}, B_e\right). \tag{29}$$

- We are able to evaluate $\hat{\phi}_{j,a^*+2}$ at X_{e,a^*+2}, B_e even when X_{e,a^*+2} contains a one-period lag of Y_{e,j,a^*+2} because we can predict Y_{e,j,a^*+1} from the previous step.

4. We iterate this procedure up to age A. For outcome j, we form the vector of predictions $\hat{Y}_{e,j} := \left[\hat{Y}_{e,j,a^*+1}, \ldots, \hat{Y}_{e,j,A}\right]$.

5. Under Assumption A–4 (Invariance), the distribution of $\hat{\varepsilon}_{n,j}$ is a consistent estimator of the distribution of $\hat{\varepsilon}_{e,j}$. We form a prediction that accounts for prediction error as follows:

$$\tilde{Y}_{e,j} = \hat{Y}_{e,j} + \hat{\varepsilon}_{n,j}. \tag{30}$$

- In practice, we randomly sample a vector of residuals from an individual j in the auxiliary sample (n) and pair it with the vector $\hat{Y}_{e,j}$ of individual i in the experimental sample (e) to form the prediction $\tilde{Y}_{e,j}$ for individual i in the experimental sample.
- That is, the pairing of individual j in the auxiliary sample (n) with individual i in the experimental sample (e) is random.
- Random pairing is valid under invariance and exogeneity, i.e. under this assumption the vector of residuals from any individual j in the auxiliary sample is a valid estimate for the vector of residuals of any individual i in the experimental sample.
- We form the pairing one time for the main point estimates, and then bootstrap this pairing when producing inference.

GMM Formulation

- Note that Assumption A–3 (Exogeneity) and Assumption A–4 (Invariance) imply the following moment condition:

$$\mathbb{E}\left[m_{j,a}\left(X_{n,a}^{d}, B_{n}; \boldsymbol{\theta}_{j,a}\right)\right] = 0, \quad k \in \{n, e\}, \quad j \in \mathcal{J}_a \tag{31}$$

where $m_{j,a}\left(X_{n,a}, B_n; \boldsymbol{\theta}_{j,a}\right) := X_{\mathbf{n},a}{}' \left(Y^d_{n,j,a} - \phi_{j,a}\left(X^d_{n,a}, B_n\right)\right)$ for $a \in [0, \ldots A]$.

- We use the auxiliary sample (n) to estimate the vector of coefficients.
- Let $m\left(\cdot, \boldsymbol{\theta}\right)$, stack the function $m_{j,a}\left(X_{n,a}, B_n; \boldsymbol{\theta}_{j,a}\right)$ for all $j \in \mathcal{J}_a$, all $a \in [0, \ldots A]$, and $k = n$.

- Observing the outcomes at age a^* provides us with additional moment conditions.
- To see this, note that, in our analysis, X_{k,a^*+1} contains a lagged variable of the outcome to predict and define the moment: $h_{j,a^*+1}\left(X_{e,a^*+1}, B_n; \boldsymbol{\theta}_{j,a^*+1}\right) =:$ $X_{e,a^*+1}{}' \left(\hat{Y}_{e,j,a^*+1} - \phi_{j,a^*+1}\left(X_{e,a^*+1}, B_e\right)\right)$, where $\hat{Y}_{e,j,a^*+1}$ is defined as before.
- Although this moment uses information in the auxiliary sample (through the construction of $\hat{Y}_{e,j,a^*+1}$), it provides additional information (not in (31)) through X_{e,a^*+1}.
- It is key moment: it initializes the out-of-sample predictions.

- For some outcomes, there are gaps in the experimental sample.
- For example, we observe labor and transfer income at ages 21 and 30.
- In this case, we have two additional moments, not only one.
- Stack these set of additional moments and denote them by $h(\cdot, \boldsymbol{\theta})$ (and helps us initialize the out-of-sample predictions).
- These additional set of moments overidentify the parameter vector of interest, $\boldsymbol{\theta}$.
- Standard procedures allow us to use these set of additional moments to improve efficiency.

- Let W be a positive definite matrix. We estimate $\boldsymbol{\theta}$ by minimizing

$$Q := \begin{bmatrix} \bar{m}\left(\cdot;\boldsymbol{\theta}\right) \\ \bar{h}\left(\cdot;\boldsymbol{\theta}\right) \end{bmatrix}' W^{-1} \begin{bmatrix} \bar{m}\left(\cdot;\boldsymbol{\theta}\right) \\ \bar{h}\left(\cdot;\boldsymbol{\theta}\right) \end{bmatrix}, \tag{32}$$

  where $\bar{u}$ denotes the empirical counterpart of u.
- W is not restricted to be diagonal so that these moments are allowed to correlate.
- Iterated, feasible procedures to obtain an estimate of W jointly with the parameters of interest guarantee efficiency and are straightforward to implement (Amemiya, 1985; Hansen, 1982).
- Altonji and Segal (1996) show that GMM presents downwards bias in absolute value in small-sample size setting, which could be a concern in our setting.

Specification 2

1. Use the auxiliary sample (n) to estimate the the coefficients characterizing $\phi_{j,a}(\cdot,\cdot)$.

- In practice, we use a weighted version of the auxiliary samples.
- The weights give relatively high importance to the individuals in the auxiliary sample whose characteristics B_k are close to the those of the individuals in the experimental sample.
- We denote these coefficients by $\boldsymbol{\theta}_{j,a}$ and the estimate of this function as $\hat{\phi}_{j,a}(\cdot,\cdot)$.
- At each age, we are able to compute the residuals from this estimation procedure as follows:

$$Y_{n,j,a} - \hat{\phi}_{j,a}(X_{k,a}, B_k) := \hat{\varepsilon}_{n,j,a}. \tag{33}$$

For outcome j, we form the vector of residuals $\hat{\boldsymbol{\varepsilon}}_{n,j} := [\hat{\varepsilon}_{n,j,a^*+1}, \ldots, \hat{\varepsilon}_{n,j,A}]$.

2. At age $a \geq a^* + 1$, we construct the predicted outcome for the experimental sample (e) for each individual as follows:

$$\hat{Y}_{e,j,a} = \hat{\phi}_{j,a}\left(X_{e,a}, B_e\right). \tag{34}$$

- We are able to evaluate $\hat{\phi}_{j,a^*+1}$ at X_{e,a^*+1}, B_e because X_{e,a^*+1} is fully observed in the experimental data.
- We stack the predictions across ages in the following vector $\hat{Y}_{e,j} := \left[\hat{Y}_{e,j,a^*+1}, \ldots, \hat{Y}_{e,j,A}\right]$.
- These predictions do not account for estimation error. We discuss estimation error below.

3. Under Assumption A–4 (Invariance), the distribution of $\hat{\varepsilon}_{n,j}$ is a consistent estimator of the distribution of $\hat{\varepsilon}_{e,j}$. We form a prediction that accounts for prediction error as follows:

$$\tilde{Y}_{e,j} = \hat{Y}_{e,j} + \hat{\varepsilon}_{n,j}. \tag{35}$$

- In practice, we randomly sample a vector of residuals from an individual j in the auxiliary sample (n) and pair it with the vector $\hat{Y}_{e,j}$ of individual i in the experimental sample (e) to form the prediction $\tilde{Y}_{e,j}$ for individual i in the experimental sample.
- That is, the pairing of individual j in the auxiliary sample (n) with individual i in the experimental sample (e) is random.
- Random pairing is valid under invariance and exogeneity, i.e. under this assumption the vector of residuals from any individual j in the auxiliary sample is a valid estimate for the vector of residuals of any individual i in the experimental sample.
- We form the pairing one time for the main point estimates, and then bootstrap this pairing when producing inference.

- In this specification, there is no "initialization" of the prediction out of sample. Thus, the GMM estimate consists of minimizing

$$Q := \left[\bar{m}\left(\cdot;\boldsymbol{\theta}\right)\right]' W^{-1}\left[\bar{m}\left(\cdot;\boldsymbol{\theta}\right)\right], \tag{36}$$

where $m_{j,a}\left(X_{n,a}, B_n; \boldsymbol{\theta}_{j,a}\right) := X_{n,a}{}' \left(Y^d_{n,j,a} - \phi_{j,a}\left(X^d_{n,a}, B_n\right)\right)$ for $a \in [0, \ldots A]$ and $X_{n,a}$ contains no lags of $Y^d_{n,j,a}$.

Specification 3

$$Y_{k,a} = \lambda_0 (1 - \rho) + (\lambda_1 + \rho) Y_{k,a-1} - \lambda_1 \rho Y_{k,a-2} + \lambda_2 (X_{k,a} - \rho X_{k,a-1}) + U_{k,a} \tag{37}$$

- This is a model with two lags and no serial correlation.
- The estimation procedure and the GMM framework are analogous to those of **Specification 1**.
- The two lags are not an issue for estimation in the auxiliary sample because we observe labor income for the full range of relevant ages, thus we estimate the prediction function.
- To initialize the procedure in the experimental sample, however, we face an issue: we do not observe labor income at $a^* - 1$.
- We assume that $a^* = a^* - 1$ and then proceed in an identical way as in **Specification 1**, the estimation procedure and the GMM framework remain the same.

THE UNIVERSITY OF CHICAGO

Specification 4

- We write:

$$
\begin{aligned}
Y_{k,a} &= \lambda_0 + \lambda_1 Y_{k,a-1} + \varepsilon_a && (38) \\
\varepsilon_{k,a} &= f + U_{k,a}, && (39)
\end{aligned}
$$

- $\mathbb{E}[U_a] = \mathbb{E}[U_a, U_{a'}] = 0$.
- We follow Arellano and Bond (1991) and note that two-lagged age values of $Y_{k,a}$ are valid instruments in the first-difference version of Equation 39.
- This allow us to estimate obtain consistent estimates of λ_0, λ_y by minimizing a weighted function (as in the previous specifications) of the empirical counterparts of the following set of moments:

$$\mathbb{E}\left[\left(\Delta Y_{k,a} - \lambda_1 \Delta Y_{k,a-1}\right) Y_{k,a-j}\right] \quad j = 2, \ldots, a-1; a = a^* + 2, \ldots, A. \tag{40}$$

1. Use the auxiliary sample (n) to estimate the coefficients in Equation (38) based on the set of moments in (40).

2. At age $a^* + 1$, use these coefficients to form the (out-of-sample) prediction in the experimental sample (e):

$$\hat{Y}_{e,a^*+1} = \hat{\lambda}_0 + \hat{\lambda}_1 Y_{e,a^*}, \tag{41}$$

noting that we observe Y_{k,a^*}.

3. At age $a^* + 2$, use the same coefficients to form the (out-of-sample) prediction, based on the $a^* + 1$ prediction. That is:

$$\hat{Y}_{e,a^*+2} = \hat{\lambda}_0 + \hat{\lambda}_1 \hat{Y}_{e,a^*+1}. \tag{42}$$

4. Iterate this procedure of to age A and stack the vector of predictions (without accounting for prediction error) as $\hat{Y}_e := \left[\hat{Y}_{e,a^*+1}, \ldots, \hat{Y}_{e,A}\right]$.

5. To account for prediction error we need an individual level estimate of $f + U_a$. We proceed as follows: (i) we observe labor income at two ages, 21 and 30. We use the estimates of the coefficients characterizing Equation (38) from the auxiliary sample (n) to predict labor income from ages 22 to 29. Then, we estimate the coefficients in Equation (38) in the experimental sample (e). This allows us to recover an estimate for $f + U_a$. In fact, we recover one estimate of $f + U_a$ for each $a \in [22, \ldots, 30]$. Each of these estimates is a valid estimate for $f + U_a$ because U_a is i.i.d. To form our prediction error, at each age, we randomly draw one element out of these available estimates. We add it to $\hat{Y}_{e,a}$ for $a \geq a^* + 1$ to form a prediction that accounts for error.

Return to main text

Inference

Predictions

1. Resample the experimental sample with replacement at the individual level. This gives us a new (re-sampled) panel dataset. Information on the entire history of each individual is obtained in each re-sample.
   - We re-sample individuals independently of their treatment status.

   Call this resampled sample (e, s). Separate this sample by treatment and control group into $(e, s, 1)$ and $(e, s, 0)$, respectively.

2. Perform the same resampling procedure on the auxiliary sample. Call this sample (n, s').

3. Form synthetic treatment and control groups by using Algorithm 1 to weight the individuals in sample (n, s'). We do not do this age by age due to problems of data availability. We use the algorithm once to match (e, s) to the CNLSY and once to match (e, s) to the PSID and NLSY79. We use the synthetic groups obtained from each of these samples to form predictions at different ages, as we explain in Appendix C.3.2. We identify synthetic control and treatment groups $(n, s', 0)$ and $(n, s', 1)$, respectively. That is, (n, s', d) for $d = 0, 1$.

4. Fit the dynamic relationship in Equation (22), using predictors as detailed in Appendix C.3.4. We fit two parameterizations of the dynamic relationships. One for the synthetic treatment, and one for the synthetic control. When providing estimates by gender, we also produce different predictions by gender.

5. To account for prediction error we need an individual level estimate of $\phi + \eta_a$. We proceed as follows: (i) we observe labor income at two ages, 21 and 30. We use the estimates of the coefficients characterizing Equation (38) from the auxiliary sample (n) to predict labor income from ages 22 to 29. Then, we estimate the coefficients in Equation (38) in the experimental sample (e). This allows us to recover an estimate for $\phi + \eta_a$. In fact, we recover one estimate of $\phi + \eta_a$ for each $a \in [22, \ldots, 30]$. Each of these estimates is a valid estimate for $\phi + \eta_a$ because η_a is i.i.d. To form our prediction error, at each age, we randomly draw one element out of these available estimates. We add it to $\hat{Y}_{e,a}$ for $a \geq a^* + 1$ to form a prediction that accounts for error.

6. In step 4., we compute an individual-level vector of residuals in each of the samples $(n, s', 0)$ and $(n, s', 1)$. That is, each individual has a vector containing the residuals of each of her predicted variable (for example, labor income). Call this vector of residuals $\boldsymbol{\mathcal{E}}^{d}_{i',n,s'}$: the vector of residuals for individual i' in the auxiliary bootstrap sample s', in the synthetic group d.

7. Randomly pair individual i' in s' with individual i in s. The prediction accounting for prediction error is $Y^d_{i,e,s} + \mathcal{E}^d_{i',e,s'}$. As described in Appendix C.7, this step changes. We estimate the prediction error from the experimental sample (and we account for this when bootstrapping as well).

THE UNIVERSITY OF CHICAGO

8. Repeat this for all pairs of samples $(n, s'), (e, s)$. We resample the experimental sample and auxiliary sample 100 times each. This gives us the empirical bootstrap distribution, with 100*100 points.

9. Compute the standard error as the sample standard deviation of the 100*100 re-samples. Compute the *p*-value's as the proportion of times that we reject the null hypothesis, after centering the empirical bootstrap distribution according to the null hypothesis.

THE UNIVERSITY OF CHICAGO

Treatment Effects:

1. Resample the experimental sample with replacement at the individual level. This gives us a new (re-samples)panel dataset. Information on the whole story about each individual is obtained in each re-sample.

THE UNIVERSITY OF CHICAGO

2. For a partially complete outcome Y_j, run K regressions of Y_j on the set of explanatory variables $k = 1, ..., K$.
   - We perform this procedure at any age, and re-sample individuals independently of their treatment status so we drop the respective indices.

   K is determined by the number of possible control sets we can construct with 1, 2 and, 3 baseline variables. We document this procedure and describe the possible control sets in Appendix G.1.

3. Choose the control set that best predicts Y_j, as we describe in Appendix G.1. Call this control set k_j^*. There is one control set per each of the partially complete outcomes Y_j.

4. Construct the IPW using the inverse of the prediction of a logistic regression of an indicator of "observed or not" on control set k_j^*.

5. For an incomplete outcome (an outcome after age a^*), we construct a prediction for treatment- and control-group individuals following the steps in Appendix C.8.1.

6. If we estimate our parameter of interest using matching (treatment vs. stay at home or treatment vs. alternative preschool —see Algorithm 1 to weight the treatment group as to make it comparable in observed characteristics to the control group individuals who either stay at home or attend alternative preschools. We use the procedure in 3. to choose the variables used to weight.

7. Repeat this procedure 1,000 times to obtain the empirical bootstrap distribution. Compute the standard error as the sample standard deviation of these resamples. Compute the *p*-value's as the proportion of times that we reject the null hypothesis, after centering the empirical bootstrap distribution according to the null hypothesis.

Combining Functions

1. Use the same procedure as before to re-sample the experimental data.
2. Calculate treatment effects as described in Appendix C.8.2.
3. If counting the number of positive effects, compute this number and generate standard errors and p-value's as before.
4. If counting the number of positive and at significant treatment effects, compute the number of positive and significant treatment effects (at the desired significance level). Re-sample the non-experimental sample a second time. The second re-sample creates an empirical bootstrap distribution for this count. Generate standard errors and p-value's as before.

Cost-benefit ratio or Internal Rate of Return

1. Use the same sampling procedure as when computing the standard error for the predictions. In this case, compute the predictions for all outcomes.
2. Discount the predictions to age of birth.
3. Compute cost-benefit ratios and internal rates of return.
4. Discard internal rate of returns not satisfying the single crossing property (see Appendix C.4).
5. Compute standard errors and p-value's as before.

THE UNIVERSITY OF CHICAGO

Return to main text

Using Matching to Construct Virtual Treatment and Comparison Groups

- Under exogeneity Assumption A–3 and invariance condition A–4 we can use matching to construct counterparts to treatment and control groups in the auxiliary sample.
- Doing so compresses the two stages of constructing a comparison group and creating predictions into one stage.
- Matching in this fashion creates direct auxiliary counterparts for each member of the experimental samples.
- It is an intuitively appealing estimator.
- Matching is a non-parametric estimator of the conditional mean functions.

- For each treatment group member i in treatment status d, we find counterparts following the analysis of Heckman et al. (1998).
- We do not construct different synthetic treatment and control groups for each age and for each outcome that we predict.
- We find one synthetic treatment and one synthetic control group in each auxiliary sample and use these samples to predict each outcome at each age.
- Match an individual $l(i)$ in the auxiliary sample to person i in the treatment samples to find synthetic treatment and control groups by following Algorithm 1.

Algorithm 1

For individual i in experimental sample ($k = e$), an individual $l(i)$ in the auxiliary sample ($k = n$) is a potential counterpart if

$$\sqrt{(X^d_{i,e} - X^d_{i,l(i),n})'(\Sigma^d_e)^{-1}(X^d_{i,e} - X^d_{i,l(i),n})} \leq \epsilon \tag{43}$$

where $X^d_{i,l(i),n}$ represents the observed characteristics of the matched potential counterpart in the non-experimental sample for $d \in \{0, 1\}$, where Σ^d_e is the covariance matrix in the experimental sample for fixed to treatment status d. We construct a synthetic control group ($d = 0$) and a synthetic treatment group ($d = 1$) by weighting the potential counterparts according to the inverse value of the left-hand-side of (43).

- The matching variables to construct the synthetic treatment and control groups are the following:
  - B_k: year of birth, male, number of siblings at birth
  - X_k: years of education, number of children, overall health index, labor income (all at age 30)

THE UNIVERSITY OF CHICAGO

Table 55: Net Present Value of Labor Income: Parametric and Non-Parametric Approaches

Pooled	Male	Female
a. Parametric (Main Paper)		
71,345	300,896	50,390
(86,343)	(241,588)	(63,060)
b. Non-Parametric (This Section)		
62,080	289,471	59,163
(75,030)	(232,414)	(74,039)

Note: this table compares the net-present value of labor and transfer income (treatment - control) using the parametric approach of the main text and the approach that we use in this section. All values are discounted to birth and reported in 2014 dollars.

- Close agreement between non-parametric estimates based on matching and parametric model-based approaches like the one we use.

Return to main text

Determinants of High Risk Index

Table 56: High-risk Index for ABC

	Item	Response	Weight
1	Maternal education (years of education)	6	8
		7	7
		8	6
		9	3
		10	2
		11	1
		12	0
2	Paternal education (years of education)	same as maternal education	
3	Year family income (2014 USD)	$5,663.54 or less	8
		$5,663.54-$11,327.08	7
		$11,327.08-$16,990.62	6
		$16,990.62-$22,654.16	5
		$22,654.16-$28,317.70	4
		$28,317.70-$33,981.24	0
4	Father's absence from the household for reason other than health or death	Yes	3
5	Lack of maternal relatives in the area	Yes	3
6	Siblings in school age one or more grades behind age-appropriate level or low scores on school-administered achievement tests	Yes	3
7	Received payments from welfare agencies within the past 3 years	Yes	3
8	Father's work unstable or unskilled and semi-skilled labor	Yes	3
9	Maternal or paternal IQ 90 or below	Yes	3
10	Sibling with an IQ score 90 or below	Yes	3
11	Relevant social agencies indicate that family is in need of assistance	Yes	3
12	One or more family members has sought professional help in the past 3 years	Yes	1
13	Special circumstances not included in any of the above that are likely contributors to cultural or social disadvantage	Yes	1

Note: This table shows the High-risk Index (HRI) for ABC. A score of 11 or more determined eligibility (Ramey and Campbell, 1984, 1991; Ramey et al., 2000; Ramey and Smith, 1977). The weighting scale aimed to establish the relative importance of each item in the index (Ramey and Smith, 1977). Race was not considered for eligibility; however, 98% of the families who agreed to participate were African-American(Ramey and Campbell, 1979; Ramey and Smith, 1977).

Return to main text

Discussion of Program Costs

- The basic costs of center-based care 50 weeks a year and 30-45 hours a week includes the costs facilities and supplies and the salaries of teaching and administrative personnel.
- The nutrition provided to the children was planned by a nutritionist prepared on site, adding to the costs of the center-based care.
- There were also daily health screenings and transportation to the center, which added additional capital and personnel costs.
- Diapers and formula were given to the subjects in both the treatment and control groups.

Return to main text

Table 57: Summary of Prediction Methodology to Construct Life-cycle Costs and Benefits

Component	Subject's Age at Prediction	Baseline Prediction Method	Variables Used to Construct Synthetic Experimental Groups
Program Costs	0 to 5	Observed (source documents)	N/A
Alternative Preschools Costs	0 to 5	Imputed from Location & Time Relevant Documents	N/A
Education Costs (includes special education and grade retention)	up to 30	Level is Observed (Per Level Cost taken from NCES)	N/A
Labor Income or Transfer Income	21 to 30	Based on Prediction Model in the Auxiliary Sample	Birth-year; Gender; Siblings at Birth
Labor Income or Transfer Income	30 to 67	Based on Prediction Model in the Auxiliary Sample	Birth-year; Gender; Siblings at Birth
Parental Labor Income	0 to 21	Linear Interpolation (Observed Values at Ages 1.5, 2.5, 3.5, 8, 12, 15, 21)	N/A
Crime (Arrests and Sentences)	up to Mid-30's	Observed* (Combines Administrative and Self-reported Data)	N/A

Table 57: Summary of Prediction Methodology to Construct Life-cycle Costs and Benefits, Cont'd

Component	Subject's Age at Prediction	Variables Used to Predict Used	Auxiliary Samples
Program Costs	0 to 5	N/A	N/A
Alternative Preschools Costs	0 to 5	N/A	N/A
Education Costs (includes special education and grade retention)	up to 30	N/A	N/A
Labor Income or Transfer Income	21 to 30 Siblings at Birth	Gender; Mother's Education; at Birth; PIAT Math (5 to 7); Education (30) Labor Income (21) Lagged Outcome	CNLSY
Labor Income or Transfer Income	30 to 67 Siblings at Birth	Gender; Education (30); Labor Income (30); Lagged Outcome	Pooled NLSY79 and PSID
Parental Labor Income	0 to 21	N/A	N/A
Crime (Arrests and Sentences)	up to Mid-30's	N/A	N/A

Table 57: Summary of Prediction Methodology to Construct Life-cycle Costs and Benefits, Cont'd

Component	**Subject's Age at Prediction**	**Baseline Prediction Method**	**Variables Used to Construct Synthetic Experimental Groups**
Crime (Arrests and Sentences)	Mid-30's to 50	Based on Prediction Model in the Auxiliary Sample (One Prediction per Arrest or Sentence)	Use Full Auxiliary Sample to Predict Control and Treatment Outcomes
Victimization Inflation	up to Age 50	Impute national victims-arrests ratio	Use Full Auxiliary Samples to Impute
Health Costs	before Age 30	Based on Prediction Model in the Auxiliary Sample	Use Full Auxiliary Sample to Predict
Health Transitions (includes disability claims)	30 to Death	Based on Prediction Model in the Auxiliary Sample	Use Full Auxiliary Samples to Predict
Health Costs	30 to Death	Based on Prediction Model in the Auxiliary Sample	Use Full Auxiliary Samples to Predict
QALYs	30 to Death	Based on Prediction Model in the Auxiliary Sample	Use Full Auxiliary Samples to Predict
Deadweight-loss	0 to Death	.50 cents per each government-spent dollar	N/A

Table 57: Summary of Prediction Methodology to Construct Life-cycle Costs and Benefits, Cont'd

Component	**Subject's Age at Prediction**	**Variables Used to Predict**	**Auxiliary Samples Used**
Crime (Arrests and Sentences)	Mid-30's to 50	Lagged Crime Outcomes	NCDPS
Victimization Inflation	up to Age 50	N/A	NCVS; NJRP; UCRS (vary by crime)
Health Costs	before Age 30	Age-specific (four follow-ups	MEPS
Health Transitions (includes disability claims)	30 to Death	Gender; Education (30); Lagged Health Outcomes	PSID and HRS (only for
Health Costs	30 to Death	Age; Gender; Race; Education (30); Marital Status (30); Disease Conditions; Labor Income (30)	MEPS MCBS (if Medicaid eligible)
QALYs	30 to Death	ADL and IADL counts; Disease Conditions	PSID and MEPS
Deadweight-loss	0 to Death	N/A	N/A

Return to main text

Table 58: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Process

	Specification 1: $\lambda_1 \neq 0$ $\rho = 0$ $f = 0$			Specification 2: $\lambda_1 = 0$ $\rho \neq 0$ $f = 0$		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	71,345	0.13	6.29	154,547	0.26	12.39
	(86,343)	(.05)	(2.11)	(187,036)	(0.11)	(5.16)
Males	300,896	0.13	11.1	200,509	0.09	7.62
	(241,588)	(0.06)	(6.35)	(160,988)	(0.04)	(3.73)
Females	59,390	0.10	2.45	79,441	0.15	3.61
	(63,060)	(0.07)	(0.79)	(99,416)	(0.11)	(1.56)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the four different specifications for prediction that are explained below. Specification 1: prediction based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: prediction based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: prediction based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: prediction based on lagged outcome; no serial autocorrelation; and fixed effect.

Table 58: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Process, Cont.

	Specification 3: $\boldsymbol{\lambda}_1 \neq 0$, $\boldsymbol{\rho} \neq 0$, $f = 0$			**Specification 4:** $\boldsymbol{\lambda}_1 = 0$, $\boldsymbol{\rho} = 0$, $f \neq 0$		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	268,179	0.49	23.64	46,953	0.09	4.14
	(211,089)	(0.12)	(5.16)	(25,323)	(0.01)	(0.62)
Males	456,078	0.2	16.82	74,775	0.03	2.76
	(358,534)	(0.09)	(9.42)	(54,752)	(0.01)	(1.44)
Females	31,303	0.05	1.29	19,959	0.03	0.82
	(168,160)	(0.19)	(2.11)	(34,142)	(0.04)	(0.43)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the four different specifications for prediction that are explained below. Specification 1: prediction based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: prediction based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: prediction based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: prediction based on lagged outcome; no serial autocorrelation; and fixed effect.

THE UNIVERSITY OF CHICAGO

Table 58: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Process, Cont.

	Specification 5: Non-parametric (under exogeneity)			**Baseline Estimates**		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	62,080	0.10	4.98	636,674	0.13	6.29
	(75,030)	(0.03)	(2.07)	(183,224)	(0.05)	(2.11)
Males	289,471	0.13	11.01	919,049	0.13	11.10
	(232,471)	(0.06)	(5.39)	(287,442)	(0.06)	(6.35)
Females	59,163	0.11	2.69	161,759	0.10	2.45
	(74,039)	(0.08)	(1.16)	(72,355)	(0.08)	(0.79)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the four different specifications for prediction that are explained below. Specification 1: prediction based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: prediction based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: prediction based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: prediction based on lagged outcome; no serial autocorrelation; and fixed effect.

THE UNIVERSITY OF CHICAGO

Return to main text

Table 59: Treatment Effects on Selected Outcomes, Males

Category	Variable	Age	$\bar{Y}_C$	(1)	(2)	(3)	(4)	(5)	(6)
Parental Income	Parental Labor Income	3.5	13,505	1,036	494	73.862	1,462	123	690
				(0.374)	(0.411)	(0.474)	(0.390)	(0.479)	(0.417)
				[0.892]	[0.927]	[1.000]	[0.857]	[0.955]	[0.891]
		12	23,868	7,085	9,625	18,050	12,639	6,620	5,383
				(0.092)	**(0.020)**	**(0.038)**	**(0.074)**	**(0.098)**	(0.139)
				[0.318]	[0.192]	[0.206]	[0.425]	[0.472]	[0.564]
		15	22,985	8,488	4,495	5,540	4,805	2,885	4,345
				(0.071)	(0.221)	(0.243)	(0.264)	(0.354)	(0.296)
				[0.288]	[0.778]	[0.825]	[0.855]	[0.911]	[0.839]
		21	21,585	12,732	8,809	122	-933	10,784	10,283
				(0.005)	**(0.098)**	(0.448)	(0.456)	**(0.056)**	**(0.041)**
				[0.068]	[0.456]	[1.000]	[0.857]	[0.367]	[0.240]
Education	Graduated High School	30	0.600	0.073	0.044	0.116	0.083	0.040	0.063
				(0.262)	(0.375)	**(0.001)**	(0.346)	(0.407)	(0.317)
				[0.653]	[0.582]	[1.000]	[0.908]	[0.597]	[0.565]
	Graduated 4-year College	30	0.120	0.170	0.138	0.149	0.099	0.135	0.143
				(0.055)	(0.128)	(0.216)	(0.338)	(0.154)	(0.130)
				[0.280]	[0.505]	[0.219]	[0.896]	[0.538]	[0.473]
	Years of Education	30	12.867	0.525	0.541	1.010	0.777	0.351	0.344
				(0.151)	(0.163)	(0.998)	(0.136)	(0.280)	(0.256)
				[0.564]	[0.505]	[1.000]	[0.638]	[0.597]	[0.565]
Labor Income	Employed	30	0.700	0.119	0.196	0.108	0.040	0.237	0.261
				(0.128)	**(0.025)**	**(0.001)**	(0.383)	**(0.025)**	**(0.013)**
				[0.456]	[0.136]	[1.000]	[0.567]	[0.113]	**[0.080]**
	Labor Income	30	30,079	19,810	24,365	25,220	20,611	23,072	21,836
				(0.091)	**(0.092)**	(0.998)	(0.122)	(0.107)	**(0.094)**
				[0.357]	[0.293]	[1.000]	[0.390]	[0.339]	[0.321]

Table 60: Treatment Effects on Selected Outcomes, Males, Cont'd

Category	Variable	Age	$\bar{Y}_C$	(1)	(2)	(3)	(4)	(5)	(6)
Crime	Total Felony Arrests	Mid-30s	1.370	0.196	0.685	1.523	1.340	0.481	0.188
				(0.368)	(0.183)	**(0.064)**	**(0.026)**	(0.284)	(0.410)
				[0.396]	[0.429]	[0.120]	**[0.081]**	[0.562]	[0.434]
	Total Misdemeanor Arrests	Mid-30s	1.296	-0.501	-0.244	-0.298	-0.034	-0.246	-0.507
				(0.171)	(0.289)	(0.314)	(0.422)	(0.329)	(0.168)
				[0.395]	[0.429]	[0.314]	[0.422]	[0.562]	[0.411]
Health	Self-reported drug user	Mid-30s	0.500	-0.333	-0.438	-0.673	-0.557	-0.326	-0.330
				(0.019)	**(0.002)**	**(0.000)**	**(0.000)**	**(0.039)**	**(0.023)**
				[0.092]	**[0.014]**	**[0.010]**	**[0.067]**	[0.102]	[0.112]
	Systolic Blood Pressure (mm Hg)	Mid-30s	138.071	-9.791	-13.275	14.196	14.976	-24.166	-18.559
				(0.113)	**(0.049)**	**(0.013)**	**(0.000)**	**(0.000)**	**(0.011)**
				[0.196]	**[0.086]**	[0.116]	**[0.001]**	**[0.010]**	**[0.014]**
	Diastolic Blood Pressure (mm Hg)	Mid-30s	89.214	-10.854	-14.134	-9.709	-8.741	-18.387	-13.987
				(0.032)	**(0.004)**	**(0.049)**	**(0.032)**	**(0.000)**	**(0.007)**
				[0.089]	**[0.012]**	[0.200]	[0.168]	**[0.011]**	**[0.014]**
	Hypertension	Mid-30s	0.571	-0.291	-0.377	-0.120	-0.074	-0.492	-0.434
				(0.042)	**(0.009)**	(0.302)	(0.353)	**(0.006)**	**(0.006)**
				[0.115]	**[0.036]**	[0.610]	[0.771]	**[0.018]**	**[0.014]**

Table 61: Treatment Effects on Selected Outcomes, Females*

Category	Variable	Age	$\bar{Y}_C$	(1)	(2)	(3)	(4)	(5)	(6)
Parental Income	Parental Labor Income	3.5	11,465	2,756	2,986	6,864	8,584	1,521	3,773
				(0.189)	(0.213)	(0.122)	**(0.045)**	(0.332)	(0.154)
				[0.447]	[0.519]	[0.249]	[0.143]	[0.624]	[0.496]
		12	20,917	13,633	19,592	28,328	26,489	15,343	18,678
				(0.054)	**(0.027)**	**(0.027)**	**(0.009)**	**(0.064)**	**(0.019)**
				[0.310]	[0.179]	[0.124]	**[0.035]**	[0.358]	[0.128]
		15	13,772	8,565	7,159	2,713	8,441	7,465	10,487
				(0.060)	(0.137)	(0.480)	(0.345)	(0.134)	**(0.064)**
				[0.310]	[0.519]	[0.709]	[0.547]	[0.504]	[0.262]
		21	20,804	5,708	8,670	45,697	25,142	6,251	3,943
				(0.136)	(0.140)	**(0.000)**	**(0.000)**	(0.224)	(0.261)
				[0.402]	[0.519]	**[0.009]**	**[0.001]**	[0.589]	[0.510]
Education	Graduated High School	30	0.529	0.253	0.131	0.553	0.595	-0.026	0.066
				(0.009)	(0.152)	**(0.003)**	**(0.000)**	(0.413)	(0.320)
				[0.072]	[0.513]	**[0.021]**	**[0.001]**	[0.698]	[0.701]
	Years of Education	30	11.794	2.143	1.843	3.861	3.923	1.163	1.409
				(0.001)	**(0.002)**	**(0.000)**	**(0.000)**	**(0.054)**	**(0.017)**
				[0.003]	**[0.033]**	**[0.010]**	**[0.001]**	[0.365]	[0.111]
Labor Income	Employed	30	0.706	0.131	0.081	0.381	0.340	-0.010	0.070
				(0.096)	(0.206)	**(0.039)**	**(0.057)**	(0.465)	(0.264)
				[0.261]	[0.501]	[0.108]	[0.131]	[0.809]	[0.660]
	Labor Income	30	23,267	2,548	1,884	15,094	13,096	-2,677	-2,122
				(0.335)	(0.382)	**(0.056)**	**(0.022)**	(0.330)	(0.363)
				[0.407]	[0.586]	[0.133]	**[0.086]**	[0.781]	[0.660]

THE UNIVERSITY OF CHICAGO

Table 62: Treatment Effects on Selected Outcomes, Females*, Cont'd

Category	Variable	Age	$\bar{Y}_C$	(1)	(2)	(3)	(4)	(5)	(6)
Crime	Total Felony Arrests	Mid-30s	0.419	-0.328	-0.351	-0.944	-0.965	-0.059	0.004
				(0.077)	**(0.087)**	**(0.095)**	**(0.095)**	(0.287)	(0.500)
				[0.134]	[0.215]	[0.167]	[0.186]	[0.432]	[0.610]
	Total Misdemeanor Arrests	Mid-30s	1.161	-0.973	-0.737	-2.010	-2.451	-0.269	-0.201
				(0.057)	(0.134)	(0.134)	(0.120)	(0.273)	(0.289)
				[0.134]	[0.238]	[0.167]	[0.186]	[0.432]	[0.610]
Health	Self-reported drug user	Mid-30s	0.259	-0.033	0.004	-0.114	-0.101	0.020	0.033
				(0.381)	(0.478)	(0.273)	(0.323)	(0.443)	(0.406)
				[0.844]	[0.857]	[0.745]	[0.559]	[0.611]	[0.943]
	Systolic Blood Pressure (mm Hg)	Mid-30s	133.963	-2.899	-5.407	-0.488	-0.822	-6.239	-6.784
				(0.307)	(0.241)	(0.488)	(0.457)	(0.249)	(0.170)
				[0.418]	[0.569]	[0.832]	[0.664]	[0.578]	[0.433]
	Diastolic Blood Pressure (mm Hg)	Mid-30s	87.556	-0.002	-0.179	4.091	4.122	-1.347	-2.160
				(0.483)	(0.438)	(0.245)	(0.222)	(0.392)	(0.339)
				[0.486]	[0.643]	[0.679]	[0.659]	[0.611]	[0.569]
	Hypertension	Mid-30s	0.407	0.172	0.085	0.077	0.162	0.102	0.107
				(0.111)	(0.293)	(0.331)	(0.245)	(0.299)	(0.255)
				[0.288]	[0.643]	[0.800]	[0.664]	[0.611]	[0.569]

Table 63: Maternal Education, ABC/CARE, Females

	N		Mean		σ		Two-sided
	Treat.	Control	Treat.	Control	Treat.	Control	t -test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Years of Education							
Birth	37	51	10.65	10.49	1.69	1.99	0.69
1.5 years	33	47	11.36	11.15	1.43	1.89	0.56
2.5 years	33	47	11.42	11.21	1.44	1.90	0.57
3.5 years	32	47	11.53	11.38	1.32	1.97	0.69
4.5 years	30	45	12.13	11.82	1.14	2.00	0.40
5.5 years	24	24	12.46	11.38	0.98	2.37	0.05
8 years	25	31	12.64	12.16	1.08	2.48	0.34
12 years	28	37	14.82	13.49	1.66	2.61	0.01
15 years	23	25	14.96	13.88	2.03	2.15	0.08
21 years	22	26	15.14	14.00	2.03	2.26	0.07
Education Level							
12 years	28	37	4.46	3.95	0.79	1.27	0.05
15 years	23	25	4.13	3.88	1.36	1.09	0.49
Graduated High School							
12 years	37	51	1.00	0.90	0.00	0.30	0.02
15 years	30	32	1.00	0.91	0.00	0.30	0.08

Note: This table shows raw descriptives of the education variables. Columns (1) through (6) give the sample size, means, and standard deviations of the variables by experimental group. In this table, "Treat." indicates the treatment group that received ABC/CARE center-based childcare and "Control" is the control group. Column (7) gives the two-sided p-value for a t-test of means between the two groups accounting for different variances in the groups. Education level is a categorical variable with higher values corresponding to more education. The ages in parentheses are the ages of the *subjects* when the measure was collected.

Table 64: Maternal Education, ABC/CARE, Males

	N		Mean		σ		Two-sided
	Treat.	Control	Treat.	Control	Treat.	Control	t -test
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Years of Education							
Birth	38	56	10.53	10.25	1.74	1.64	0.44
1.5 years	35	50	11.14	10.86	1.70	1.47	0.43
2.5 years	35	50	11.14	10.94	1.70	1.45	0.57
3.5 years	35	48	11.23	11.02	1.73	1.49	0.57
4.5 years	34	47	11.44	11.19	1.76	1.58	0.51
5.5 years	24	20	11.75	11.05	1.98	1.79	0.23
8 years	31	29	12.32	11.72	2.07	1.46	0.20
12 years	28	37	14.57	13.22	2.63	1.90	0.03
15 years	24	21	14.88	14.00	2.35	2.14	0.20
21 years	25	20	14.20	14.35	2.57	2.28	0.84
Education Level							
12 years	28	37	4.36	3.81	1.22	1.05	0.06
15 years	24	21	4.46	4.00	1.25	1.00	0.18
Graduated High School							
12 years	38	56	0.92	0.93	0.27	0.26	0.89
15 years	29	25	0.97	0.92	0.19	0.28	0.49

Note: This table shows raw descriptives of the education variables. Columns (1) through (6) give the sample size, means, and standard deviations of the variables by experimental group. In this table, "Treat." indicates the treatment group that received ABC/CARE center-based childcare and "Control" is the control group. Column (7) gives the two-sided p-value for a t-test of means between the two groups accounting for different variances in the groups. Education level is a categorical variable with higher values corresponding to more education. The ages in parentheses are the ages of the *subjects* when the measure was collected.

Return to main text

Table 65: Second-phase Treatment vs. Control Groups, ABC

Variable	Age	Control Obs.	Treated Obs.	Control Mean	Treated Mean	p-value Single H_0	p-value 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	**(0.035)**	**(0.075)**
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.

Table 66: Observed vs. Attritted Children, ABC

Variable	Age	Obs.	Att.	Observed Mean	Attritted Mean	p-value Single H_0	p-value Multiple H_0
Male	0	103	13	0.488	0.248	**(0.085)**	(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	**(0.045)**	**(0.095)**
Mother's Education	0	103	13	10.302	9.192	**(0.100)**	(0.165)
Mother's Age	0	103	13	20.016	18.178	**(0.080)**	(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	**(0.070)**	(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 attrited 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.

Table 67: First-phase Treatment vs. Control Groups, Dropping Attrited Children, ABC

Variable	Age	Control Obs.	Treated Obs.	Control Mean	Treated Mean	*p*-value Single H_0	*p*-value 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	**(0.085)**	(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.

Table 68: CARE Baseline Characteristics, Control vs. Family Education, Cohort 6

Variable	Age	Control Obs.	Treated Obs.	Control Mean	Treated Mean	p-value Single H_0	p-value 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	**(0.075)**	(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 UNIVERSITY OF CHICAGO

Return to main text

Table 69: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Process

	Specification 1: ("Baseline") $\lambda_1 \neq 0$, $\rho = 0$, $f = 0$			Specification 2: $\lambda_1 = 0$, $\rho \neq 0$, $f = 0$		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	636,674	0.14	7.33	154,547	0.15	7.31
	(183,224)	(0.03)	(1.84)	(187,036)	(0.12)	(5.15)
Males	919,049	0.15	10.19	200,509	0.11	9.35
	(287,442)	(0.04)	(2.93)	(160,988)	(0.05)	(5.51)
Females	161,759	0.10	2.61	79,441	0.19	4.64
	(72,355)	(0.06)	(0.73)	(99,416)	(0.28)	(3.19)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the five different specifications for forecasts that are explained below. Specification 1 is our baseline estimate. It also presents the calculation of the internal rate of return and the benefit/cost ratio of the program using these different net present values. Specification 1: forecast based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: forecast based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: forecast based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: forecast based on lagged outcome; no serial autocorrelation; and fixed effect. Model: The different specifications are particular cases of the following model.

Table 69: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Process, Cont'd

	Specification 3: $\lambda_1 \neq 0$, $\rho \neq 0$, $f = 0$			**Specification 4:** $\lambda_1 \neq 0$, $\rho = 0$, $f \neq 0$		
	NPV	IRR	B/C	NPV	IRR	B/C
Pooled	268,179	0.26	12.68	46,953	0.05	2.22
	(211,089)	(0.14)	(5.81)	(25,323)	(0.02)	(0.7)
Males	456,078	0.25	21.26	74,775	0.04	3.49
	(358,534)	(0.12)	(12.28)	(54,752)	(0.02)	(1.88)
Females	31,303	0.07	1.83	19,959	0.05	1.17
	(168,160)	(0.48)	(5.4)	(34,142)	(0.1)	(1.1)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the five different specifications for forecasts that are explained below. Specification 1 is our baseline estimate. It also presents the calculation of the internal rate of return and the benefit/cost ratio of the program using these different net present values. Specification 1: forecast based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: forecast based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: forecast based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: forecast based on lagged outcome; no serial autocorrelation; and fixed effect. Model: The different specifications are particular cases of the following model.

Table 70: Net Present Value of Labor Income and Cost/Benefit Analysis Under Different Specifications for Labor Income Process, Cont'd

	Specification 5: Non-parametric matching		
	NPV	IRR	B/C
Pooled	132,924	0.13	6.28
	(11,253)	(0.01)	(0.31)
Males	196,530	0.11	9.16
	(20,210)	(0.01)	(0.69)
Females	69,317	0.17	4.05
	(4,350)	(0.0 1)	(0.14)

Note: This table displays the net present value of labor income in 2014 USD (treatment - control) using the five different specifications for forecasts that are explained below. Specification 1 is our baseline estimate. It also presents the calculation of the internal rate of return and the benefit/cost ratio of the program using these different net present values. Specification 1: forecast based on lagged outcome; no serial autocorrelation; and no fixed effect. Specification 2: forecast based on lagged outcome; arbitrary serial autocorrelation; and no fixed effect. Specification 3: forecast based on lagged outcome; first-order serial autocorrelation; and no fixed effects. Specification 4: forecast based on lagged outcome; no serial autocorrelation; and fixed effect. Model: The different specifications are particular cases of the following model.

Return to main text

Table 71: Sensitivity Analysis for Benefit/Cost Ratios

	Pooled	Males	Females
Baseline	**6.29** (s.e. 2.11)	**11.10** (s.e. 6.35)	**2.45** (s.e. 0.79)

Baseline: IPW and Controls, Life-span up to predicted death, Treatment vs. Next Best, 50% Marginal tax 50% (deadweight loss), Discount rate 3%, Parental income 0 to 21 (child's age), Labor Income predicted from 21 to 65, All crimes (full costs), Value of life 150,000.

	Pooled		Males		Females	
Specification	*No IPW* **6.69** (2.09)	*and No Controls* **5.96** (1.65)	*No IPW* **11.10** (6.69)	*and No Controls* **9.67** (3.76)	*No IPW* **2.82** (0.87)	*and No Controls* **2.56** (0.76)
Prediction Span	*to Age 21* **1.56** (0.38)	*to Age 30* 2.01 (0.86)	*to Age 21* 2.02 (0.76)	*to Age 30* 2.31 (1.86)	*to Age 21* 1.31 (0.44)	*to Age 30* 1.69 (0.50)
Counter-factuals	*vs. Stay at Home* **4.58** (1.63)	*vs. Alt. Presch.* **6.69** (2.02)	*vs. Stay at Home* 4.44 (2.90)	*vs. Alt. Presch.* **10.40** (4.18)	*vs. Stay at Home* **5.04** (1.22)	*vs. Alt. Presch.* **2.48** (0.75)
Deadweight-loss	*0%* **10.02** (3.30)	*100%* **5.05** (1.66)	*0%* **17.39** (9.61)	*100%* **8.65** (4.83)	*0%* **4.27** (1.39)	*100%* **2.37** (0.75)
Discount Rate	*0%* **16.49** (6.97)	*7%* **2.67** (0.76)	*0%* **29.66** (18.12)	*7%* 4.10 (2.07)	*0%* **5.77** (3.55)	*7%* **1.66** (0.41)

Table 71: Sensitivity Analysis for Benefit/Cost Ratios

	Pooled	Males	Females
Baseline	**6.29** (s.e. 2.11)	**11.10** (s.e. 6.35)	**2.45** (s.e. 0.79)

Baseline: IPW and Controls, Life-span up to predicted death, Treatment vs. Next Best, 50% Marginal tax 50% (deadweight loss), Discount rate 3%, Parental income 0 to 21 (child's age), Labor Income predicted from 21 to 65, All crimes (full costs), Value of life 150,000.

	Pooled		Males		Females	
Parental Income	*Mincer Life-cycle* **6.52** (2.11)	*Life-cycle Prediction* **6.71** (2.20)	*Mincer Life-cycle* **11.29** (6.35)	*Life-cycle Prediction* **11.56** (6.42)	*Mincer Life-cycle* **2.73** (0.84)	*Life-cycle Prediction* **3.00** (0.94)
Labor Income	*.5% Annual Decay* **5.93** (2.09)	*.5% Annual Growth* **6.64** (2.15)	*.5% Annual Decay* **10.39** (6.02)	*.5% Annual Growth* **11.81** (6.71)	*.5% Annual Decay* **2.38** (0.76)	*.5% Annual Growth* **2.53** (0.83)
Crime	*Drop Major Crimes* **5.68** (1.97)	*Halve Costs* **4.68** (1.37)	*Drop Major Crimes* **11.31** (5.42)	*Halve Costs* **7.72** (4.16)	*Drop Major Crimes* **2.56** (0.83)	*Halve Costs* **2.09** (0.73)
Health (QALYs)	*Drop All* **5.38** (2.04)	*Double Value of Life* **7.19** (2.42)	*Drop All* 9.90 (6.13)	*Double Value of Life* **12.27** (6.75)	*Drop All* **2.32** (0.76)	*Double Value of Life* **2.59** (1.15)

Table 72: Sensitivity Analysis for Internal Rate of Return, ABC/CARE

	Pooled	Males	Females
Baseline	**0.13** (s.e. 0.05)	**0.13** (s.e. 0.06)	0.10 (s.e. 0.08)

Baseline: IPW and Controls, Life-span up to predicted death, Treatment vs. Next Best, 50% Marginal tax 50% (deadweight loss), Discount rate 3%, Parental income 0 to 21 (child's age), Labor Income predicted from 21 to 65, All crimes (full costs), Value of life 150,000.

	Pooled		Males		Females	
Specification	*No IPW* **0.14** (0.05)	*and No Controls* **0.13** (0.04)	*No IPW* **0.15** (0.10)	*and No Controls* **0.14** (0.05)	*No IPW* 0.12 (0.08)	*and No Controls* 0.09 (0.05)
Prediction Span	*to Age 21* **0.11** (0.05)	*to Age 30* **0.11** (0.04)	*to Age 21* **0.12** (0.05)	*to Age 30* **0.13** (0.05)	*to Age 21* **0.12** (0.08)	*to Age 30* **0.13** (0.07)
Counter-factuals	*vs. Stay at Home* **0.09** (0.04)	*vs. Alt. Presch.* **0.14** (0.04)	*vs. Stay at Home* **0.07** (0.04)	*vs. Alt. Presch.* **0.15** (0.06)	*vs. Stay at Home* **0.14** (0.07)	*vs. Alt. Presch.* 0.10 (0.08)
Deadweight-loss	*0%* **0.20** (0.08)	*100%* **0.11** (0.07)	*0%* **0.18** (0.08)	*100%* **0.11** (0.05)	*0%* 0.21 (0.19)	*100%* **0.09** (0.10)

Table 72: Sensitivity Analysis for Internal Rate of Return, ABC/CARE

	Pooled	Males	Females
Baseline	**0.13** (s.e. 0.05)	**0.13** (s.e. 0.06)	0.10 (s.e. 0.08)

Baseline: IPW and Controls, Life-span up to predicted death, Treatment vs. Next Best, 50% Marginal tax 50% (deadweight loss), Discount rate 3%, Parental income 0 to 21 (child's age), Labor Income predicted from 21 to 65, All crimes (full costs), Value of life 150,000.

	Pooled		Males		Females	
Parental Income	*Mincer Life-cycle*	*Life-cycle Prediction*	*Mincer Life-cycle*	*Life-cycle Prediction*	*Mincer Life-cycle*	*Life-cycle Prediction*
	0.15	**0.14**	**0.14**	**0.14**	0.14	**0.13**
	(0.06)	(0.06)	(0.07)	(0.06)	(0.12)	(0.11)
Labor Income	*.5% Annual Decay*	*.5% Annual Growth*	*.5% Annual Decay*	*.5% Annual Growth*	*.5% Annual Decay*	*.5% Annual Growth*
	0.13	**0.13**	**0.13**	**0.13**	0.10	0.10
	(0.05)	(0.04)	(0.06)	(0.06)	(0.08)	(0.08)
Crime	*Drop Major Crimes*	*Halve Costs*	*Drop Major Crimes*	*Halve Costs*	*Drop Major Crimes*	*Halve Costs*
	0.13	**0.11**	**0.14**	**0.11**	0.11	0.10
	(0.05)	(0.05)	(0.05)	(0.05)	(0.08)	(0.08)
Health (QALYs)	*Drop All*	*Double Value of Life*	*Drop All*	*Double Value of Life*	*Drop All*	*Double Value of Life*
	0.12	**0.13**	**0.12**	**0.13**	0.10	0.09
	(0.06)	(0.05)	(0.07)	(0.06)	(0.08)	(0.07)

Return to main text

Table 73: Summarizing Results Across Methodologies, Females

	A. Treatment Effects (Step Down)		
Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income		$\times$	
Employment		$\times^*$	
Crime	$\times$	$\times$	
Health			

Note: This table summarizes the results across three methodologies for our four main benefit categories, by counterfactual comparison. Panel a. marks with a cross categories of outcomes for which the outcomes are mostly positive and significant. A starred cross indicates that these results tend to persist after accounting for multiple hypotheses testing (with blocks formed across outcomes within that category).

Table 73: B. Factors (Across Blocks), Females

Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	0.286	0.506*	0.298
Employment	0.434*	0.999*	0.244
Crime	0.239*	0.725	0.070
Health	0.172	0.177	0.177

This table presents treatment effects on latent factor scores constructed using all measures of parental income, employment and labor income, crime, and health (hypertension- and cholesterol-related measures). The in-sample mean of the latent factor scores is zero and the standard deviation is 1. Starred point estimates are statistically significant at the 10% level.

Table 73: Summarizing Results Across Methodologies, Females, Cont.

	C. Combining Functions		
Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	87*	93*	80*
Employment	80*	80*	80*
Crime	100*	100*	75*
Health	53	56	63

Note: This table summarizes the results across three methodologies for our four main benefit categories, by counterfactual comparison. Panel a. marks with a cross categories of outcomes for which the outcomes are mostly positive and significant. A starred cross indicates that these results tend to persist after accounting for multiple hypotheses testing (with blocks formed across outcomes within that category). Panel b. displays the proportion of outcomes displaying a positive treatment effect. For Panels b. and c. we star point estimates that are significant at the 10% level.

Table 73: Summarizing Results Across Methodologies, Females, Cont.

	D. Benefit NPV		
Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	148,854*	199,653*	133,595*
Employment	41,908*	188,241*	16,307
Crime	31,668*	46,684	5,041
Health	42,102	240,947*	51,064

Note: This table summarizes the results across three methodologies for our four main benefit categories, by counterfactual comparison. Panel c. presents the life-cycle net-present value of the benefit due to this category in the calculation of the cost benefit analysis (2014 USD). For Panels b. and c. we star point estimates that are significant at the 10% level.

Table 74: Summarizing Results Across Methodologies, Males

	A. Treatment Effects		
Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	$\times$		$\times$
Employment	$\times$		$\times$
Crime	$\times$	$\times^*$	
Health	$\times^*$		$\times^*$

Note: This table summarizes the results across three methodologies for our four main benefit categories, by counterfactual comparison. Panel a. marks with a cross categories of outcomes for which the outcomes are mostly positive and significant. A starred cross indicates that these results tend to persist after accounting for multiple hypotheses testing (with blocks formed across outcomes within that category).

Table 74: B. Factors (Across Blocks), Males

Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	0.108	0.363	0.124
Employment	0.635*	0.251*	0.693*
Crime	0.396	0.649*	0.226
Health	0.875*	0.062*	1.140*

This table presents treatment effects on latent factor scores constructed using all measures of parental income, employment and labor income, crime, and health (hypertension- and cholesterol-related measures). The in-sample mean of the latent factor scores is zero and the standard deviation is 1. Starred point estimates are statistically significant at the 10% level.

Table 74: Summarizing Results Across Methodologies, Males, Cont.

	C. Combining Functions		
Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	73*	67*	73*
Employment	80*	60	80*
Crime	25	25	25
Health	63*	32	74*

Note: This table summarizes the results across three methodologies for our four main benefit categories, by counterfactual comparison. A starred cross indicates that these results tend to persist after accounting for multiple hypotheses testing (with blocks formed across outcomes within that category). Panel b. displays the proportion of outcomes displaying a positive treatment effect. For Panels b. and c. we star point estimates that are significant at the 10% level.

THE UNIVERSITY OF CHICAGO

Table 74: Summarizing Results Across Methodologies, Males, Cont.

	D. Benefit NPV		
Counterfactual:	Next Best	Stay at Home	Alternatives
Outcome:			
Parental Income	107,907*	100,450*	95,437*
Employment	238,105*	265,680*	227,883*
Crime	638,923*	-123,498	796,147*
Health	106,218	201,383	51,793

Note: This table summarizes the results across three methodologies for our four main benefit categories, by counterfactual comparison. A starred cross indicates that these results tend to persist after accounting for multiple hypotheses testing (with blocks formed across outcomes within that category). Panel c. presents the life-cycle net-present value of the benefit due to this category in the calculation of the cost benefit analysis (2014 USD). For Panels b. and c. we star point estimates that are significant at the 10% level.

Return to main text

The description of each of the estimates is as follows:
Estimate 1 —.58 (s.e. .28):

- obtain the return of a one-standard gain in cognition from Chetty et al (2011), in terms of labor income at age 27 (this is how Chetty et al report their return to cognition)
- Use the fraction of std found to project
- obtain the standardized gain in cognition in ABC/CARE
- assign a labor income gain to each individual based on the return of Chetty et al (2011).
- compute the average gain (netting out the control).
- divide the average gain by the average cost of the program.
- standard errors come from bootstrapping sample, although the return remains constant in all the calculation.
- This is what Kline and Walters do.

Estimate 2 —.09 (s.e. .04)
Exactly analogous to Estimate 1, but instead of using Chetty's return to cognition, use our own return to cognition in terms of labor income at age 30.

- in this case, the standard errors do account for variation in the return, as we calculate the return in every bootstrap re-sample. In that sense this is a "better" estimate if compared to Kline and Walters.
- The return is smaller because our sample is way more disadvantaged than that of Chetty et al., whose individuals come from a more mixed background (Project STAR).

Estimate 3:

- our CBA, but the only benefit is labor income up to age 34, and the cost is the cost of the program.

Estimate 4:

- our CBA, but stopping the accounting of benefits and costs at age 34 and the cost is the cost of the program.

Estimate 5:

- our CBA, but accounting only labor income as a benefit and the cost is the cost of the program.

Estimate 6:

- our CBA (baseline)

THE UNIVERSITY OF CHICAGO

Return to main text