Long Run Puzzles in Head Start Research

Doug Miller
War on Poverty conference
Center for Poverty Research
UC Davis, January 10, 2014
Long run Head Start Puzzles: This talk

• Brief history of Head Start, and history of related research debates

• What we know, and why we know so little, about long run impacts

• Advertisement / preview of ongoing work here at UC Davis
Long run Head Start Puzzles: History

• We all know and love Head Start
• Not part of LBJ’s war on poverty speech!
• Housed in Office of Economic Opportunity
• Serendipitous alignment of:
  • Excess CAP funds in first year – bad local politics – led to targeting children
  • Personal history (Eunice Kennedy Shriver, Rosemary Kennedy, president’s panel on mental retardation)
  • Legislative (Republican) & Administrative (HEW, Office of Ed) competition
• “Project Rush-Rush” (eg, $180/kid)
• Local (not state!) agencies applied directly to OEO
Long run Head Start Puzzles: History

• 1965-1972: wild West (wild South?)
• 1973-1988: relative stability
• 1989-2001: massive expansion
• 2002-2010: relative stability
• 1965-today
  • Perceived success!

Head Start’s attraction: Fairness and Efficiency

• Fairness: What a great target demographic!
Head Start’s attraction: Fairness and Efficiency

- Fairness: What a great target demographic!
Head Start’s attraction: Fairness and Efficiency

- Efficiency: long-run impacts from investment in early childhood.
- “Neuroplasticity”; “Dynamic complementarities in learning”
Long run Head Start Puzzles, part 1: Recurring debates 1965-2014

1. Does it work? And the question of “fade out” ...
2. More vs. Less
3. And if more, “quantity” vs. “quality”
4. Academic vs “Whole Child”
Long run Head Start Puzzles, part 2: What is the long-run impact?

- This is the key question. But it’s hard!
- Short-run impact is hard to measure
  - Perennial challenge of identifying causal effects from nonexperimental settings:
    - Those who don’t sign up for HS are bad comparisons to those who do
Long run Head Start Puzzles, part 2: What is the long-run impact?

• Short-run impact is hard to measure
• Economists’ approach: quasi-experiments
• Many of the confounding variables are correlated with “demand for Head Start,” so ...
• Identify a “supply shock”
  • Ideally one that’s not correlated with other determinants of long-run outcomes
Long run Head Start Puzzles, part 2: What is the long-run impact?

- Short-run impact is hard to measure
- Long-run impact is even harder!
  - Same problems as SR. AND ...
  - Difficult to find data that links “LR outcomes” to “Head Start Exposure”
  - ... and also enables quasi-experimental variation!
- Also, “external validity” issues
  - Any valid estimate speaks only to
    - The (population / program / alternatives) of the time
Long run Head Start Puzzles, part 2: What is the long-run impact?

- Ideal situation
  - Identify LR impact from earlier cohorts
    - AND impacts on SR outcomes for those cohorts
    - Like “Intermediate Clinical Endpoints” and “Ultimate Clinical Endpoints” in medicine
  - Find stable relationship between SR and LR outcomes
  - Examine SR outcomes in today’s cohorts
Long run Head Start Puzzles, part 2: What is the long-run impact?

- Two types of “best available” direct measures of LR impacts:
  - Within-family sibling comparisons
    - Currie & Thomas (1995, NLSY)
    - Deming (2009)
    - Garces, Thomas, & Currie (2002, PSID)

- Early implementation grant-writing assistance
  - Ludwig & Miller (2007)
## Long run Head Start Puzzles, part 2: What is the long-run impact?


<table>
<thead>
<tr>
<th></th>
<th>ALL</th>
<th>AFRICAN-AMERICAN</th>
<th>WHITE</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School Grad</td>
<td>0.037</td>
<td>-0.025</td>
<td>0.203**</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.065)</td>
<td>(0.098)</td>
</tr>
<tr>
<td>Some college</td>
<td>0.092</td>
<td>0.023</td>
<td>0.281**</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.066)</td>
<td>(0.108)</td>
</tr>
<tr>
<td>Booked/Charged w/ Crime</td>
<td>-0.053</td>
<td>-0.116**</td>
<td>0.122</td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.045)</td>
<td>(0.077)</td>
</tr>
<tr>
<td>N</td>
<td>1,742</td>
<td>706</td>
<td>1,036</td>
</tr>
</tbody>
</table>
Long run Head Start Puzzles, part 2: What is the long-run impact?

  - (+) schooling attainment ~ one half year
  - (+) attending some college ~ 15% of the control mean.
  - (-) child mortality
Long run Head Start Puzzles, part 2: What is the long-run impact?

• Promising, in-progress: Johnson (2013)
  • PSID geo-coded to county-year funding data
  • Panel FE design
  • Beneficial impacts on Schooling, Wages, Incarceration, Health

• The Optimistic take on LR impacts
  • Johnson (2013): “Estimated long-term benefits for previous cohorts ... From 3 separate research designs, three independent datasets (sibling difference, regression discontinuity, diff-in-diff)”
Long run Head Start Puzzles, part 2: What is the long-run impact?

- Is there a consensus? No!
- NYT, Page A1, April 14, 1969
Long run Head Start Puzzles, part 2: What is the long-run impact?

- Is there a consensus? No!
- Joe Klein, Time Magazine, July 2011
- Test score fade out, NHSIS, 2010.
- Randomized intervention = “gold standard”

TIME TO AX PUBLIC PROGRAMS THAT DON’T YIELD RESULTS

“...finally there is indisputable evidence about the program’s effectiveness, provided by the Department of Health and Human Services: Head Start simply does not work.”

“[Continued funding is ] criminal, every bit as outrageous as tax breaks for oil companies.”
Long run Head Start Puzzles, part 2: What is the long-run impact?

• Optimism: LR impacts
• Pessimism: test score fade out
• Optimism rejoinder 1: There was “fade out” for cohorts w/ LR impacts!
  • Deming (2009)
  • Ludwig & Miller, Garces Thomas & Currie, Westinghouse
  • Also, Perry Preschool
  • Also, Tennessee STAR
• Optimism rejoinder 2: cognitive scores (1-2 years out) wrong “intermediate clinical endpoint”
  • Some positive impacts w/in NHSIS
    • Parent involvement (Gelber & Isen 2013)
    • Subgroup (lower tail) impacts, non-cognitive skills (Bitler et al 2013)
Long run Head Start Puzzles, part 2: What is the long-run impact?

• (1) Optimism; (2) Pessimism; (3) Optimism rejoinders

• (4) Pessimism rejoinder 1:
  • NHSIS measured non-cognitive scores (zero effects)
  • Is this a fishing expedition? We know what we want to find!

• Pessimism rejoinder 2: the LR evidence is not bullet proof
Long run Head Start Puzzles, part 2: What is the long-run impact?

• Re-assessing the LR evidence: Ludwig-Miller (2007)

• Educational gains?
  • Marginal statistical significance.
  • E.g. NELS, Yrs Schooling, +0.58, \( (T^* = 1.55) \)
  • E.g. Census, HS Grad, +0.03, \( (p \text{ value} = 0.032) \)
  • Concerns about migration

• Health gains?
  • “HS susceptible causes” = Anemias, Meningitis, Respiratory
  • Small fraction of mortality then; much smaller now.
Long run Head Start Puzzles, part 2: What is the long-run impact?

- Re-assessing the LR evidence: Garces Thomas Currie (2002)
- Well-known concerns about “sibling comparison” strategies
  - Why did one child get exposure, the other did not?
  - Back to problems w/ non-experimental research designs

- Our replication & extension of G-T-C indicates:
  - Sibling comparison estimates in PSID only suggestive, not definitive.
PSID sibling comparison analysis

• Following G-T-C (2002), we re-construct PSID sample

• Looks good for Means and (full sample) sample size, and “observational” regression.

• Then we re-estimate “sibling comparison” regression ...
### PSID sibling comparison analysis

Sibling comparison sample, mother

FE estimates

<table>
<thead>
<tr>
<th></th>
<th>GTC (2002)</th>
<th>UC Davis Econ (2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALL</td>
<td>AFRICAN-AMERICAN</td>
</tr>
<tr>
<td>High School Grad</td>
<td>0.037</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>(0.053)</td>
<td>(0.065)</td>
</tr>
<tr>
<td>Some college</td>
<td>0.092</td>
<td>0.023</td>
</tr>
<tr>
<td></td>
<td>(0.056)</td>
<td>(0.066)</td>
</tr>
<tr>
<td>Booked/Charged w/ Crime</td>
<td>-0.053</td>
<td><strong>0.116</strong></td>
</tr>
<tr>
<td></td>
<td>(0.039)</td>
<td>(0.045)</td>
</tr>
<tr>
<td>N</td>
<td>1,742</td>
<td>706</td>
</tr>
</tbody>
</table>
PSID sibling comparison analysis
Sibling comparison sample, mother
FE estimates

• Investigating the discrepancies, we learned:
  • Smaller “N” than you might think!

• Eg., African-American sibling sample, N = 627
  • 94% of which are in families with no Head Start switching
  • About 50 children in “Head Start switching” families ..
  • ... of whom, about 13 kids booked/charged with a crime.
Next, we expand the sample
- Later cohorts
- Older siblings
- More than 3x sample size

Also, we examine longer-run outcomes (through mid-40’s)
## PSID sibling comparison analysis
### Sibling comparison sample, mother
### FE estimates

<table>
<thead>
<tr>
<th></th>
<th>UCD Original Sample</th>
<th>UCD Expanded Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ALL</td>
<td>ALL</td>
</tr>
<tr>
<td></td>
<td>AFRICAN-AMERICAN</td>
<td>AFRICAN-AMERICAN</td>
</tr>
<tr>
<td></td>
<td>WHITE</td>
<td>WHITE</td>
</tr>
<tr>
<td>High School Grad</td>
<td>0.050</td>
<td>0.011</td>
</tr>
<tr>
<td></td>
<td>-0.025</td>
<td>-0.016</td>
</tr>
<tr>
<td></td>
<td>0.140</td>
<td>0.034</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.057)</td>
</tr>
<tr>
<td></td>
<td>(0.088)</td>
<td>(0.025)</td>
</tr>
<tr>
<td></td>
<td>(0.088)</td>
<td>(0.028)</td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td></td>
</tr>
<tr>
<td>Some college</td>
<td>0.097</td>
<td>0.065**</td>
</tr>
<tr>
<td></td>
<td>-0.008</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>0.230**</td>
<td>0.161***</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td>(0.031)</td>
</tr>
<tr>
<td></td>
<td>(0.054)</td>
<td>(0.032)</td>
</tr>
<tr>
<td></td>
<td>(0.098)</td>
<td>(0.057)</td>
</tr>
<tr>
<td></td>
<td>(0.059)</td>
<td></td>
</tr>
<tr>
<td>Booked/Charged w/ Crime</td>
<td>0.052</td>
<td>0.010</td>
</tr>
<tr>
<td></td>
<td>-0.050</td>
<td>-0.038</td>
</tr>
<tr>
<td></td>
<td>0.230*</td>
<td>0.068</td>
</tr>
<tr>
<td></td>
<td>(0.036)</td>
<td>(0.029)</td>
</tr>
<tr>
<td></td>
<td>(0.042)</td>
<td>(0.024)</td>
</tr>
<tr>
<td></td>
<td>(0.13)</td>
<td>(0.055)</td>
</tr>
<tr>
<td>N</td>
<td>1,554</td>
<td>5,341</td>
</tr>
<tr>
<td></td>
<td>627</td>
<td>2,347</td>
</tr>
<tr>
<td></td>
<td>924</td>
<td>2,988</td>
</tr>
</tbody>
</table>

*Significance levels: **p < 0.01, *p < 0.05*
Also, we examine longer-run outcomes (through mid-40’s)

No impacts for:
- Cigarettes, drinks, SRHS, BMI, food stamps, TANF, ln(earnings), Employment, Unemployment
Long run Head Start Puzzles, part 2: What is the long-run impact?

- (1) Optimism; (2) Pessimism; (3) Optimism rejoinders; (4) Pessimism rejoinders

- Reminder of the Ideal situation:
  - LR impact from earlier cohorts
    - AND SR outcomes for those cohorts
  - Stable relationship between SR and LR
  - SR outcomes today

- We are a long way off!
Long run Head Start Puzzles: This talk

• Brief history of Head Start, and history of related research debates

• What we know, and why we know so little, about long run impacts

• Advertisement / preview of ongoing work here at UC Davis
Preliminary Results EULA

• I acknowledge that the following results are based on extremely preliminary data analysis.

• I expect that with further data and analysis work by the researchers, they will change.

• I will not take these too seriously – they are intended as “proof of concept”

• I may need to accept cookies to view these results.
  • (The type you eat)
New work in progress: Three projects in search of titles

1. “Untitled project: Head Start long run impact, PSID analysis”
2. “Untitled project: Head Start funding data, county-year and state-year panels”
3. “Untitled project: Head Start long run impact, rapid growth in funding during the 1990s”

Joint work with: Ariel Marek, Esra Kose, Michel Grosz, Na’ama Shenhav, Natalie Ho
2: “Untitled project: Head Start funding data, state-year and county-year panels”
State-Year Panel

• Many sources of secondary data
  • OEO reports
  • Head Start Statistical Fact Sheets
  • NCES digest
  • Congressional Research Service Report
  • GPO Budget reports

• Funding and (sometimes) enrollment

• Used in two ways
  • Can validate later county-year panel
  • Direct source of information on Head Start exposure

• Also: population (3-4) and child poverty estimates
State-Year Panel

We have many years, but not all!
State-Year Panel

HS funding per 3 & 4 year old, by state

Note: RI & SD excluded. Log scale for Y-axis. Real 2009 dollars.
County-Year Panel

- Community Action Program funding data (1965-1968)

- Federal Outlay System Files (1968-1980)

- These provide information on funding at the Program-County-year level.
County-Year Panel

• These data are very messy!
• And without decent documentation

• Three examples:
  • “letters” instead of numbers in funding data.
  • Amite County, MS, 1974
  • New York and New Jersey, 1974

• Lots of cleaning work done so far
  • Lots more left to do

• So far, data quality is a “decent start”
County-Year Panel

Time series comparing county data against state-year panel and national time series.
County-Year Panel

Cross section comparing county data against state-year panel. Log scale.
County-Year Panel

Cross sections comparing county data against state-year panel. Log scale.
County-Year Panel
What does the data look like?
State-Year and County-Year Panels
Lessons learned

• These data have potential, but require deep attention to cleaning.

• Difficult to even know what to check against

• I would welcome leads and suggestions
3: “Untitled project: Head Start long run impact, rapid growth in funding during the 1990s”
Growth in HS funding 1990-2001

• Big!
• Equalizing across states
• Not uniform across states
• Left great amounts of variation
• HS $ per poor 3-4 year old:

Growth in HS funding 1990-2001

Real (2012$) HS Funds per Poor 3-4 Year Old, US States
Growth in HS funding 1990-2001

- What is behind variation in HS growth? One potential answer: legislative language.
- We are collecting this for Head Start’s history. Example, USC 42, 1994:

§ 9831 TITLE 42—THE PUBLIC HEALTH AND WELFARE

SUBCHAPTER II—HEAD START PROGRAMS
Codicification


1. Set asides
2. Each state gets its 1981 $$
3. Of the excess ... 
   1. 1/3 based on 0-18 AFDC caseload
   2. 2/3 based on 0-5 kids poverty
Growth in HS funding 1990-2001
Legislated formula and actual HS $
Growth in HS funding 1990-2001
Legislated formula and actual HS $
Growth in HS funding 1990-2001

<table>
<thead>
<tr>
<th>Percent Change in Real HS Funding per 3-4 Year Old, 1990-2001</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Predicted growth in real funding</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Change in ln 3-4 pop</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Change in fraction kids AFDC/TANF</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Change in fraction 0-5 poor</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Fraction 0-5 poor, 1989</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Constant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Adjusted R-Squared</td>
</tr>
<tr>
<td>N</td>
</tr>
</tbody>
</table>

Notes: * p < .10, ** p < .05, *** p < .01. Robust SE. Unweighted.
The promise of this research design: we only need to know state and cohort in order to get “treatment intensity”
  • Many available datasets
  • Many outcomes – including “intermediate clinical endpoints”
  • Migration less of a concern

This design extends naturally to periods outside of “the ramp up”
Long run Head Start Puzzles: Conclusion

• We all know and love Head Start
• But we don’t know as much as we should
• Stay tuned ...