

External Validity in Program Evaluation: An Introduction

Douglas J. Besharov
University of Maryland

Conference on External Validity in Program Evaluation

Co-sponsored by

The Laboratory for Interdisciplinary Evaluation of Public Policies at
Sciences Po and

The University of Maryland School of Public Policy
Paris, France

June 2023

WITHIN OUR REACH

**BREAKING
THE CYCLE OF
DISADVANTAGE**

LISBETH B. SCHORR

WITH DANIEL SCHORR

FOREWORD BY
WILLIAM JULIUS WILSON

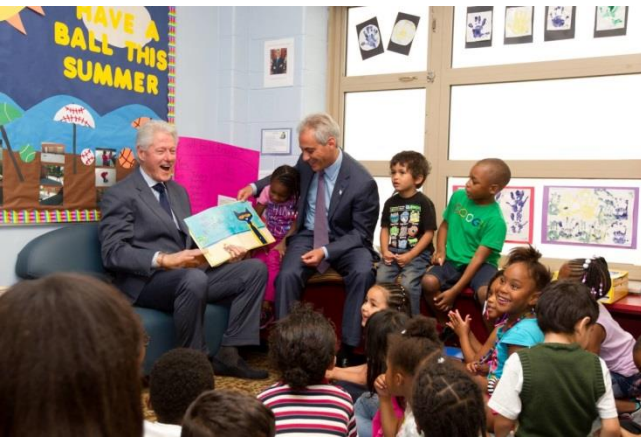
“If the superb health, education, and social services described in this book, now provided to a fraction of those who need them, were more widely available, fewer children would come into adulthood unschooled and unskilled, committing violent crimes, and bearing children as unmarried teenagers. Fewer of today’s vulnerable children would tomorrow swell the welfare rolls and the prisons. Many more would grow into responsible and productive adults, able to form stable families and contributing to, rather than depleting, America’s prosperity and sense of community.

***“Going to scale”
requires external validity***

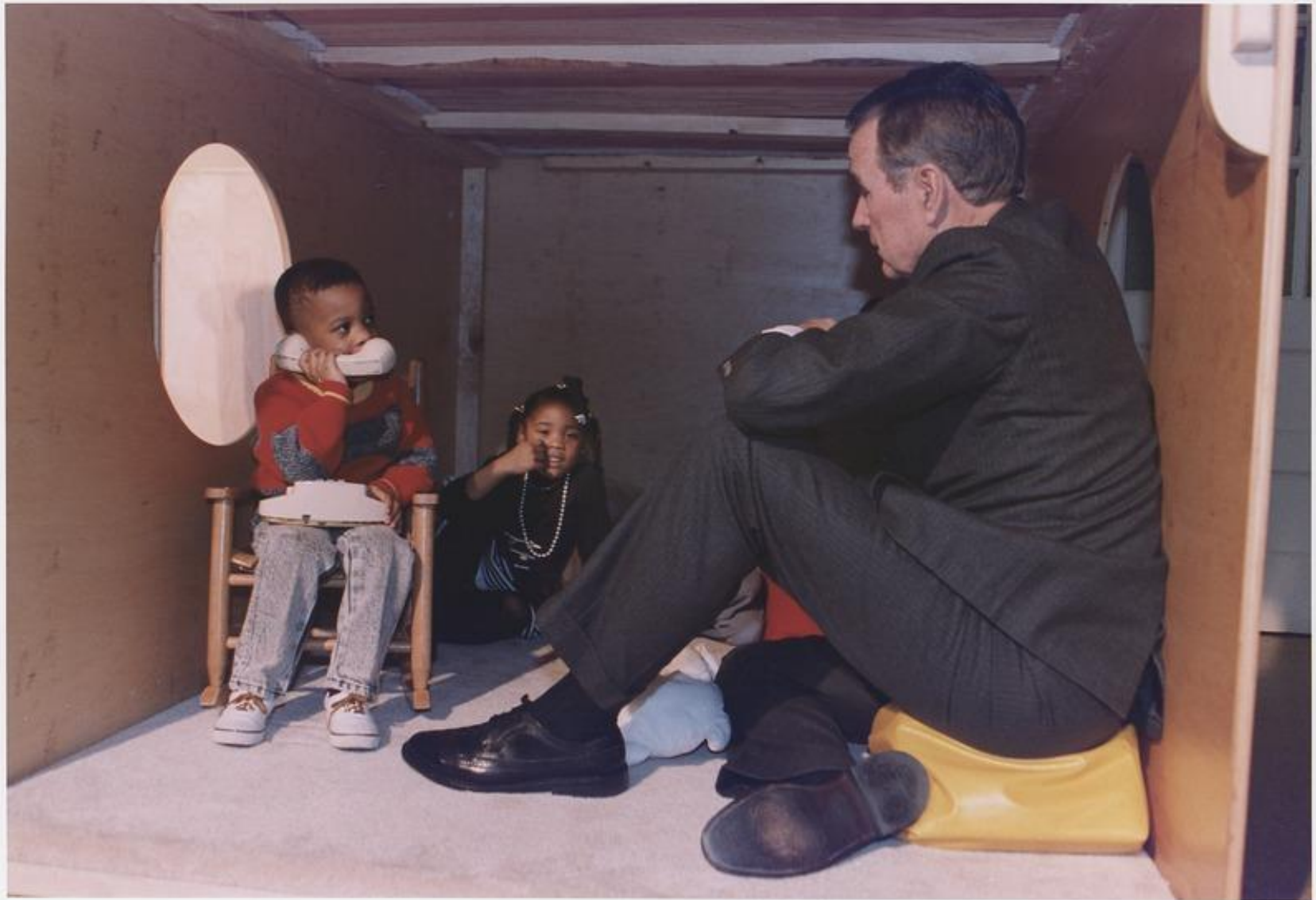
External Validity, Generalizability, and Applicability

- “*External validity*” is the extent to which a study’s results can be applied (“*generalized*”) beyond the specific situation examined, that is, to situations “external” to the study itself.
- The “*generalizability*” of results can be to all relevant elements of a program and its target population or to only some elements or populations (“*applicability*”).

Head Start: A popular program, but does it “work”



Republicans In A Box



How Children Are Raised Matters



**Of course, it did not go well for Remes
So, Parents Are Indeed Important**



Estimates of Early Childhood Benefits

With Zero Benefits Ignored

7:1
Perry (age 27)

3.8:1
Chicago

0.6:1
*NFP, Elmira
(low risk)*

7:1
Chicago

12.9:1
Perry (age 40)

3.6:1
Abecedarian

2.9:1
*Chicago**

1:0.6
*Abecedarian**

4:1
*NFP, Elmira
(high risk)*

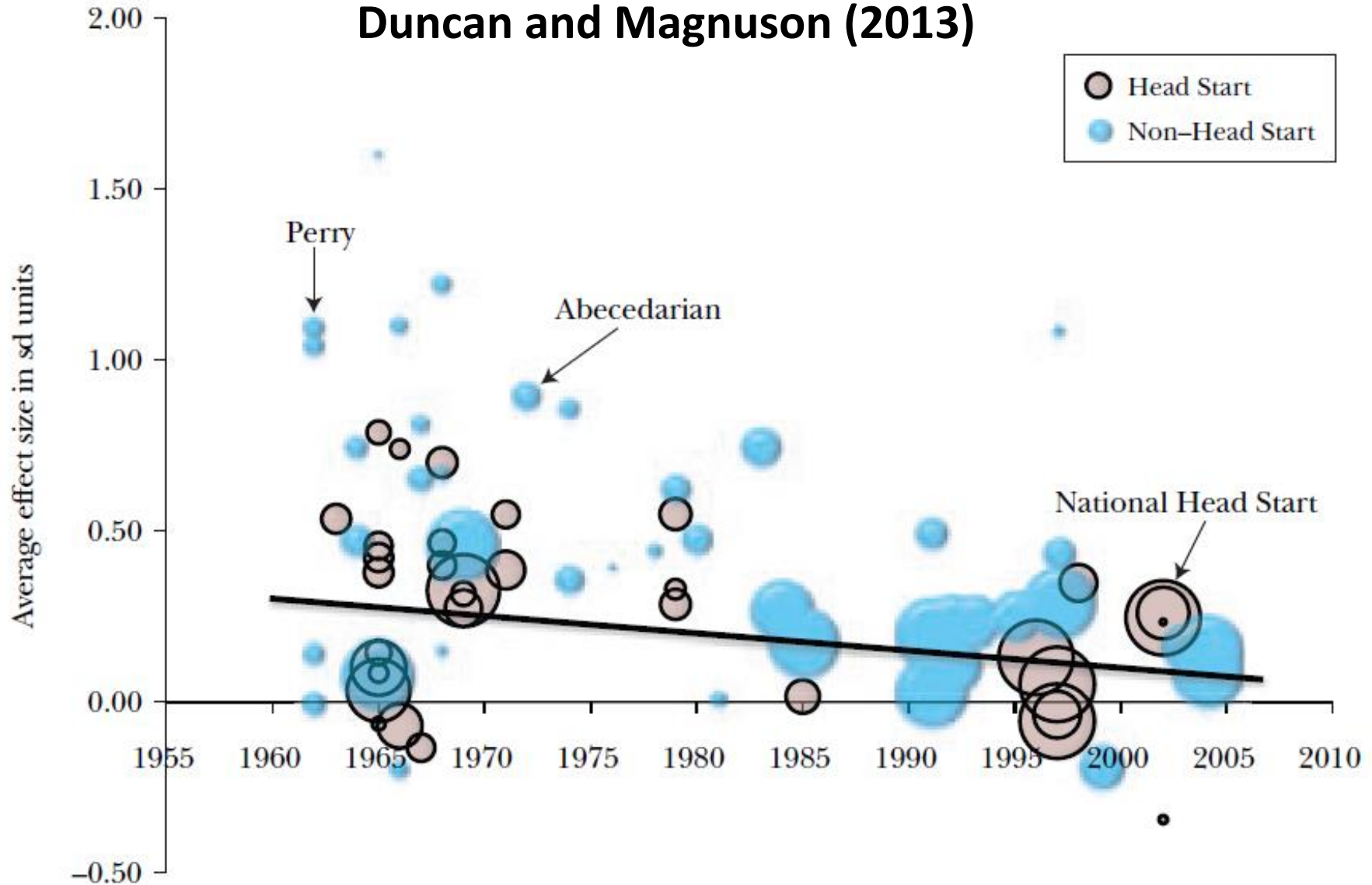
2.5:1
*Perry (age 27)**

Elements of External Validity

- Population validity
- Place validity
- Temporal validity
- Scale validity
- Treatment validity
- Cross-effects validity

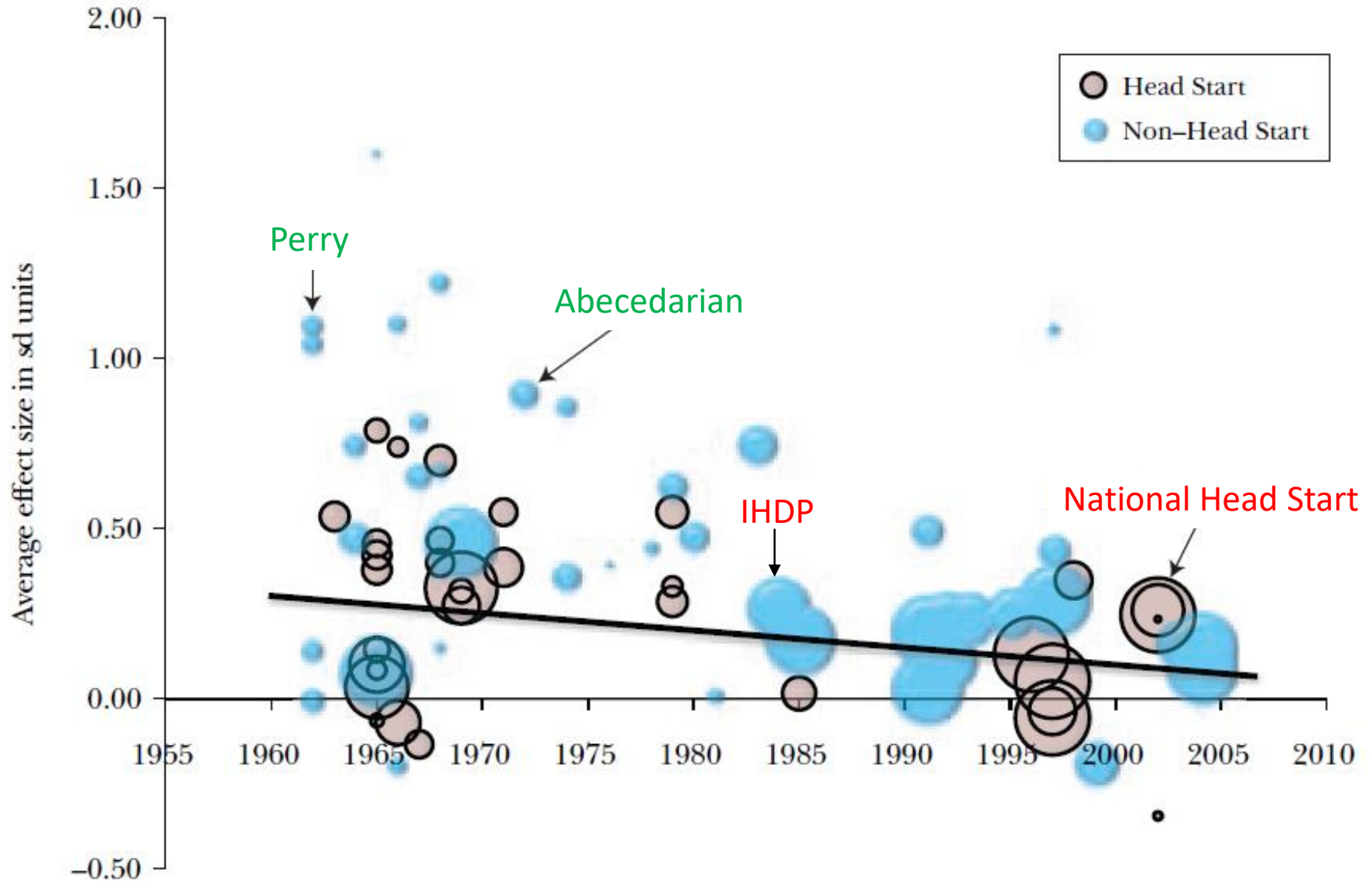
Average Impacts of Early Child Care Programs at End of Treatment

Duncan and Magnuson (2013)



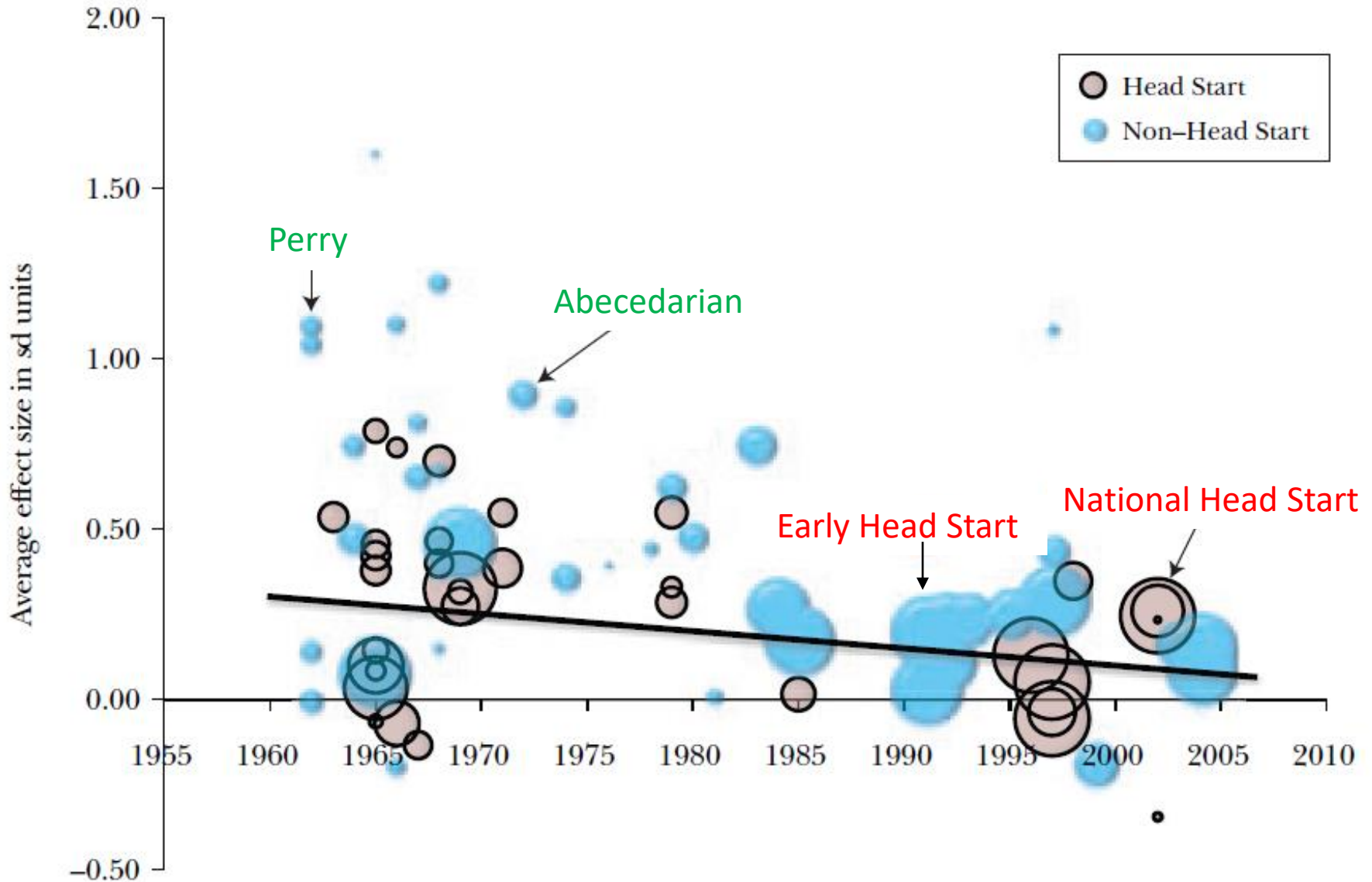
Population Validity

University Towns vs. The Nation



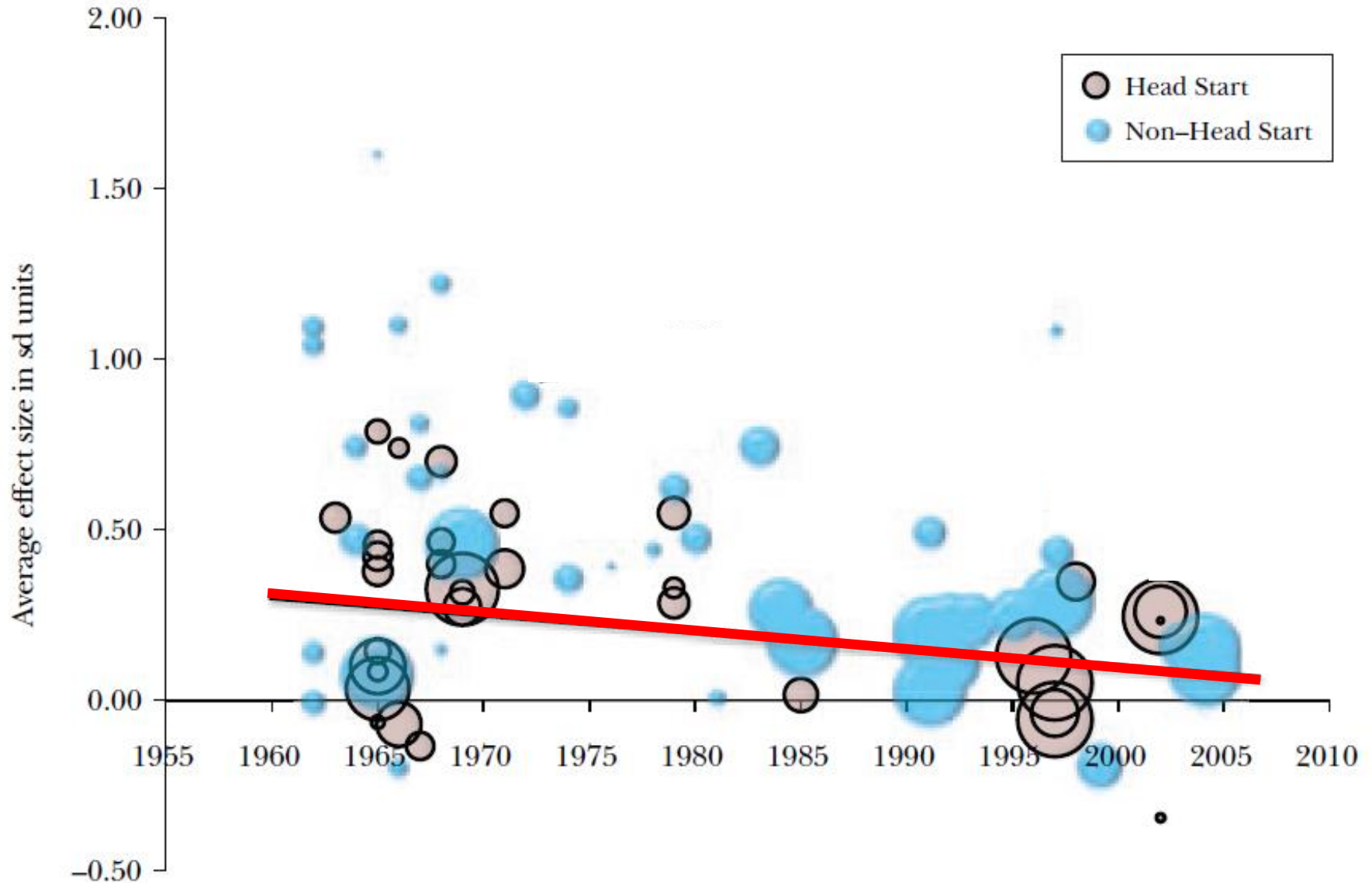
Place (Setting) Validity

University Centers vs. Community-Based Grantees



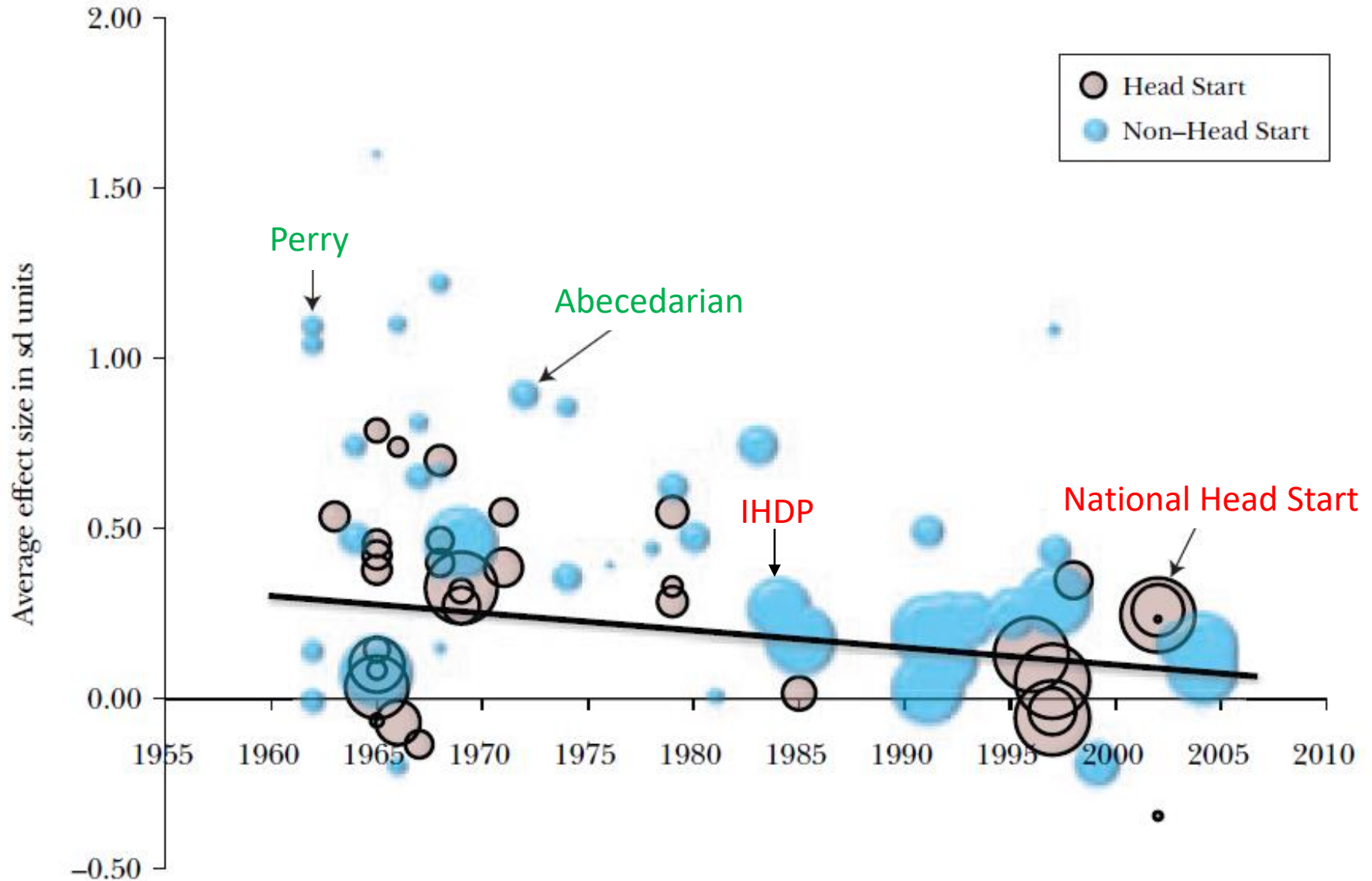
Temporal Validity

Before and After Civil Rights Progress



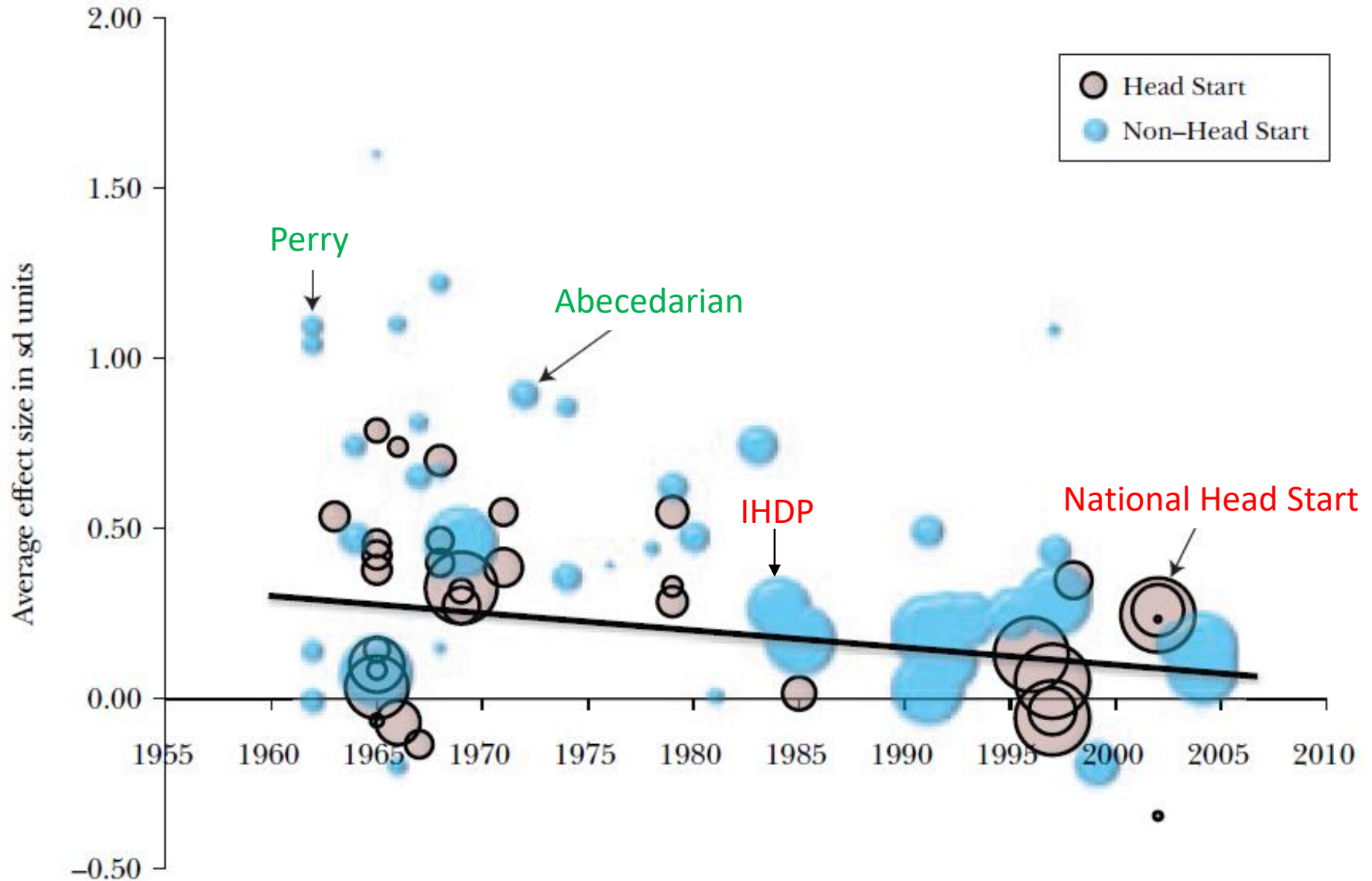
Scale Validity

Hothouse Demonstration vs. Universal



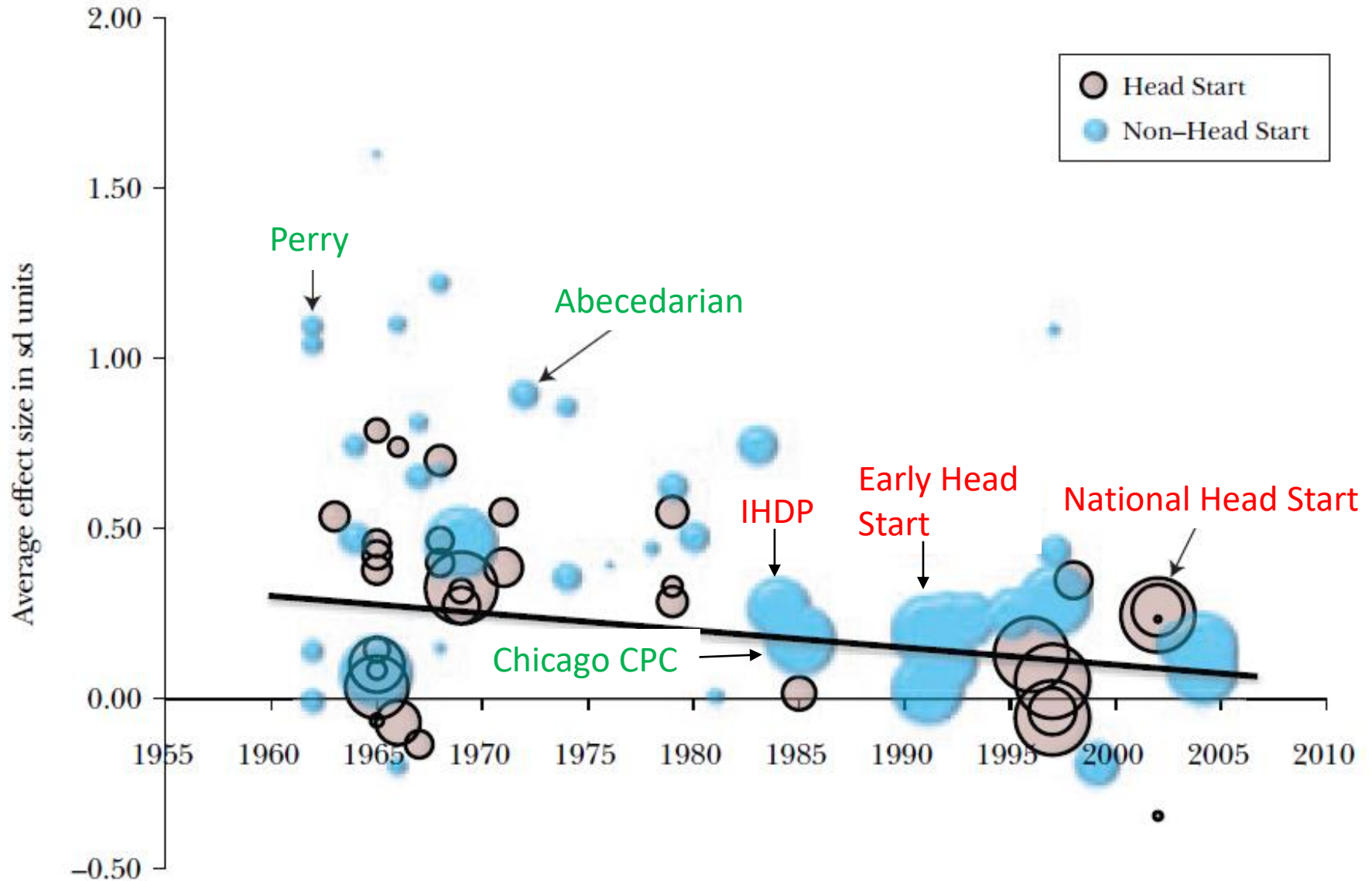
Treatment Validity

Monitored Services vs. Loose Federal Regulation



Causal Validity Always Matters

Small-Scale Studies vs. Institutional RCTs



Estimates of Early Childhood Benefits

With Zero Benefits Ignored

7:1
Perry (age 27)

3.8:1
Chicago

0.6:1
NFP, Elmira
(low risk)

7:1
Chicago

12.9:1
Perry (age 40)

3.6:1
Abecedarian

2.9:1
Chicago*

1:0.6
Abecedarian*

4:1
NFP, Elmira
(high risk)

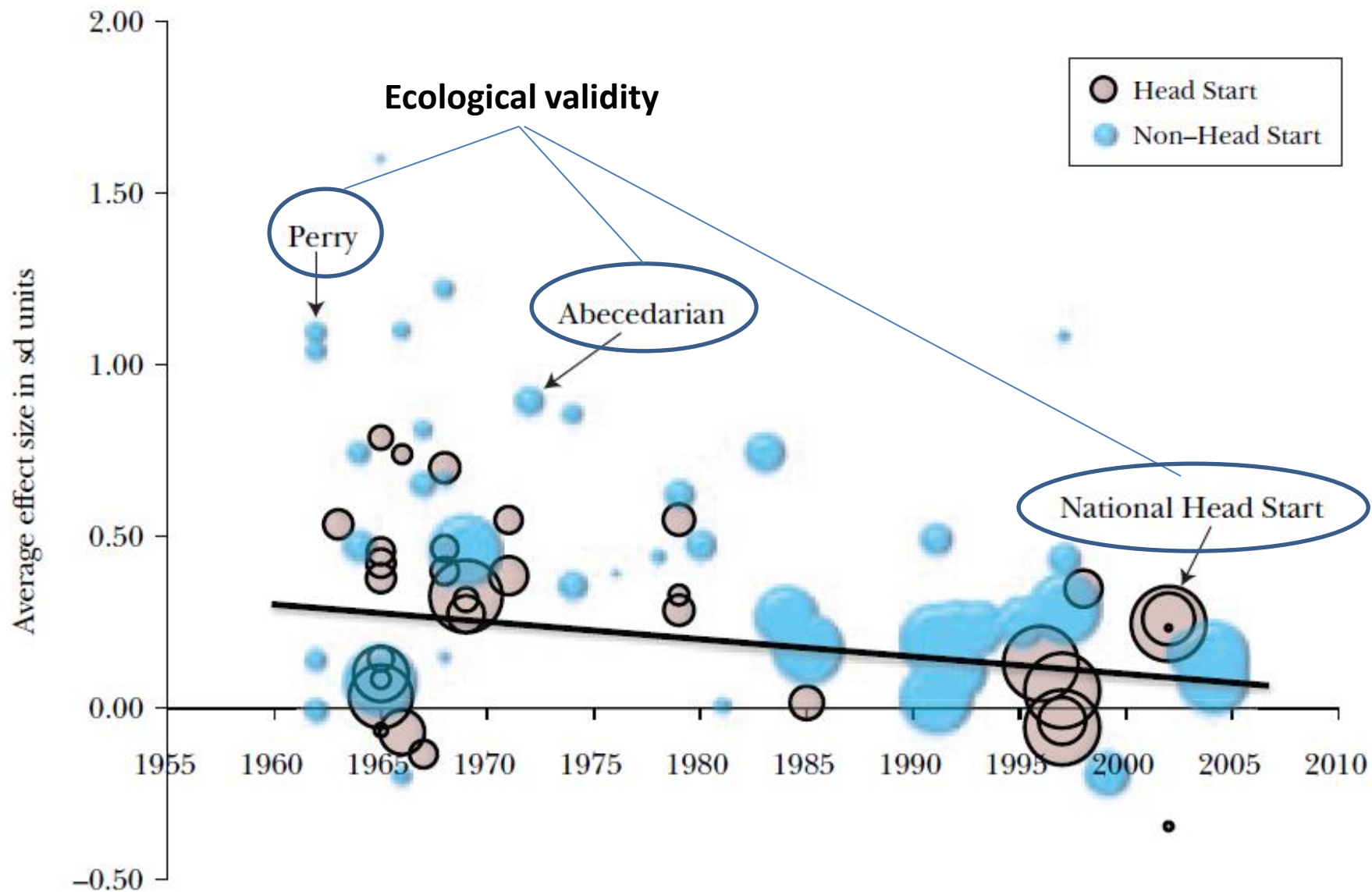
2.5:1
Perry (age 27)*

The Search For An Application That “Works”



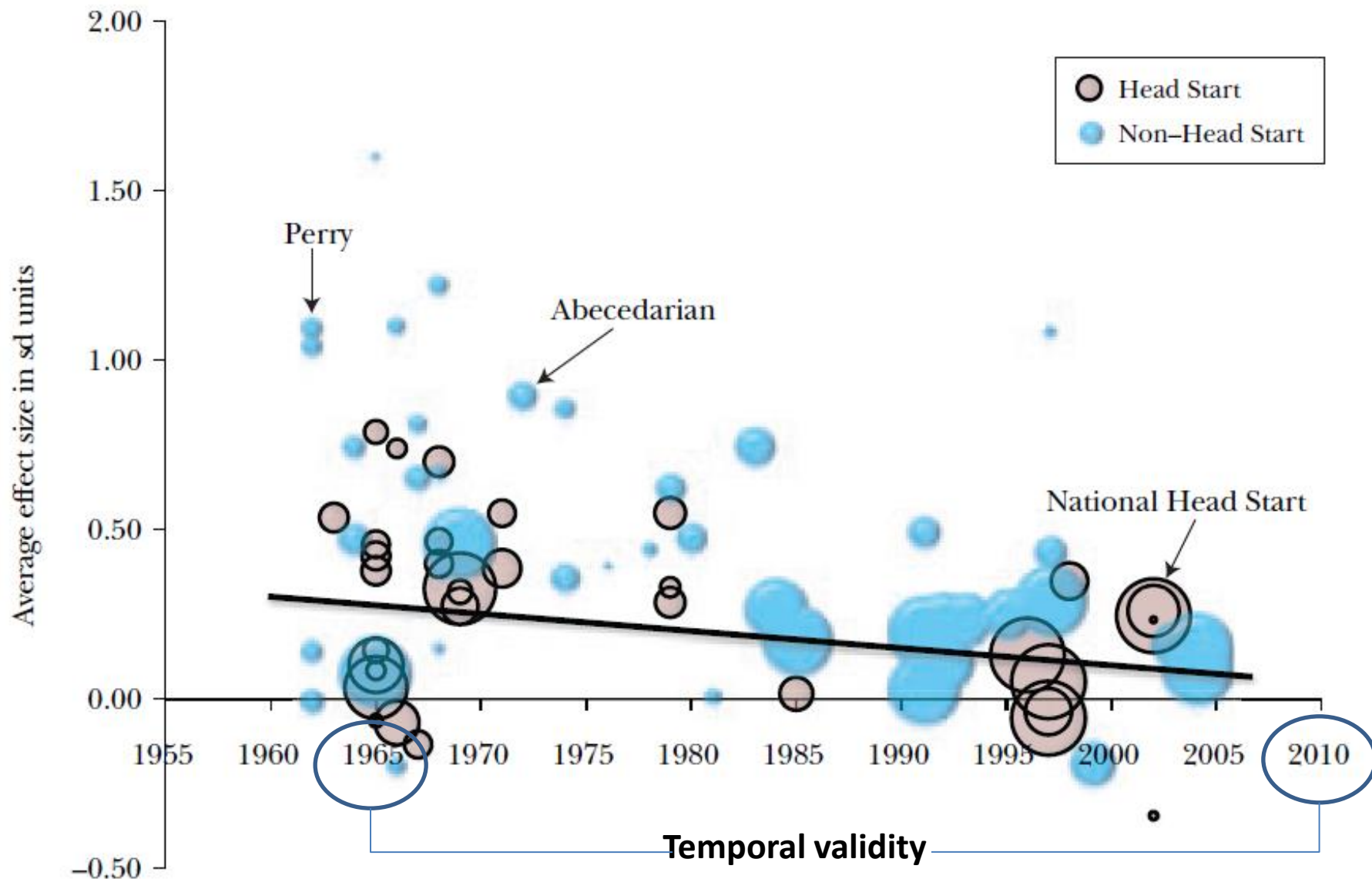
Average Impact of Early Child Care Programs at End of Treatment

(standard deviation units)



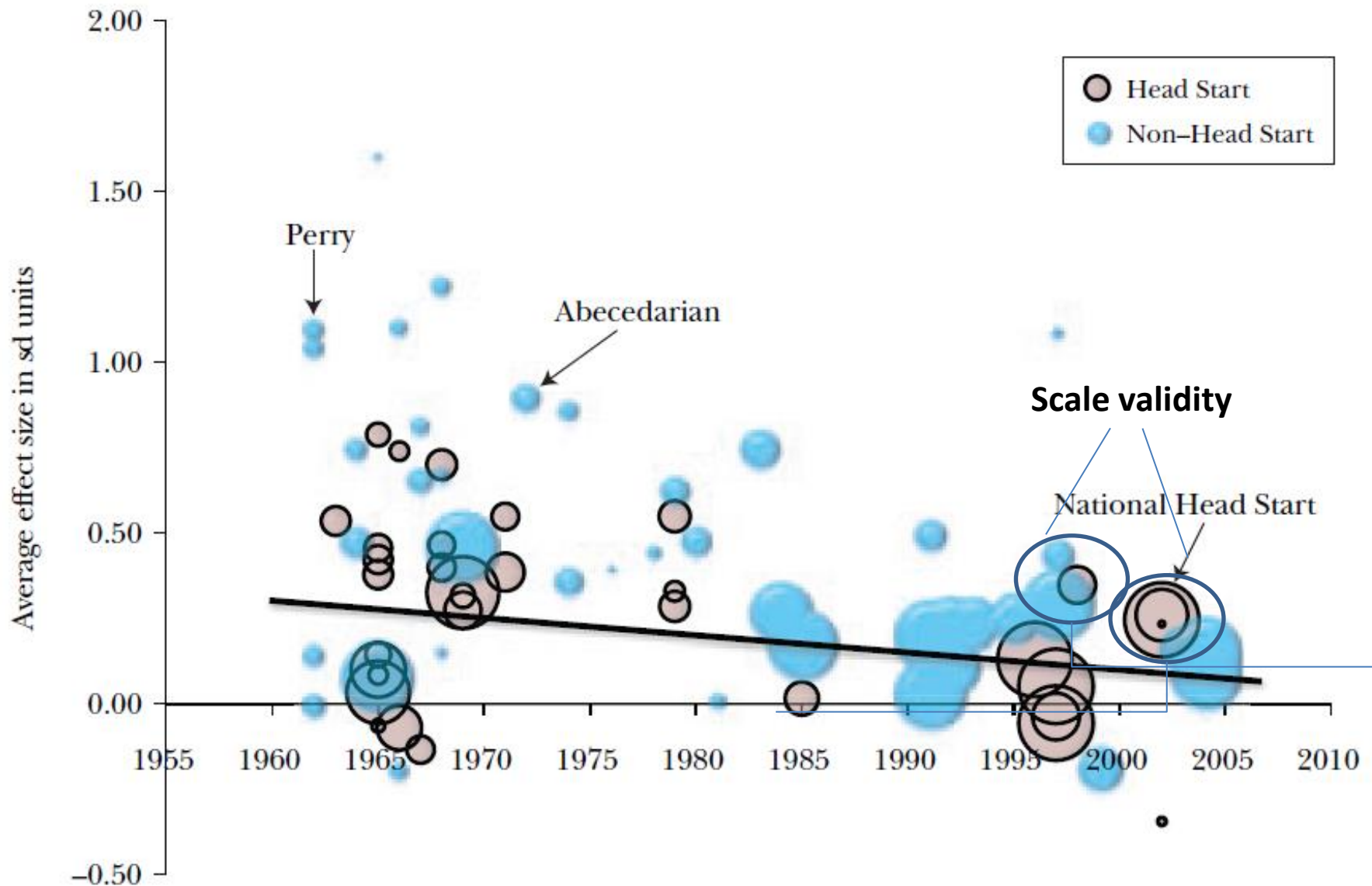
Average Impact of Early Child Care Programs at End of Treatment

(standard deviation units)



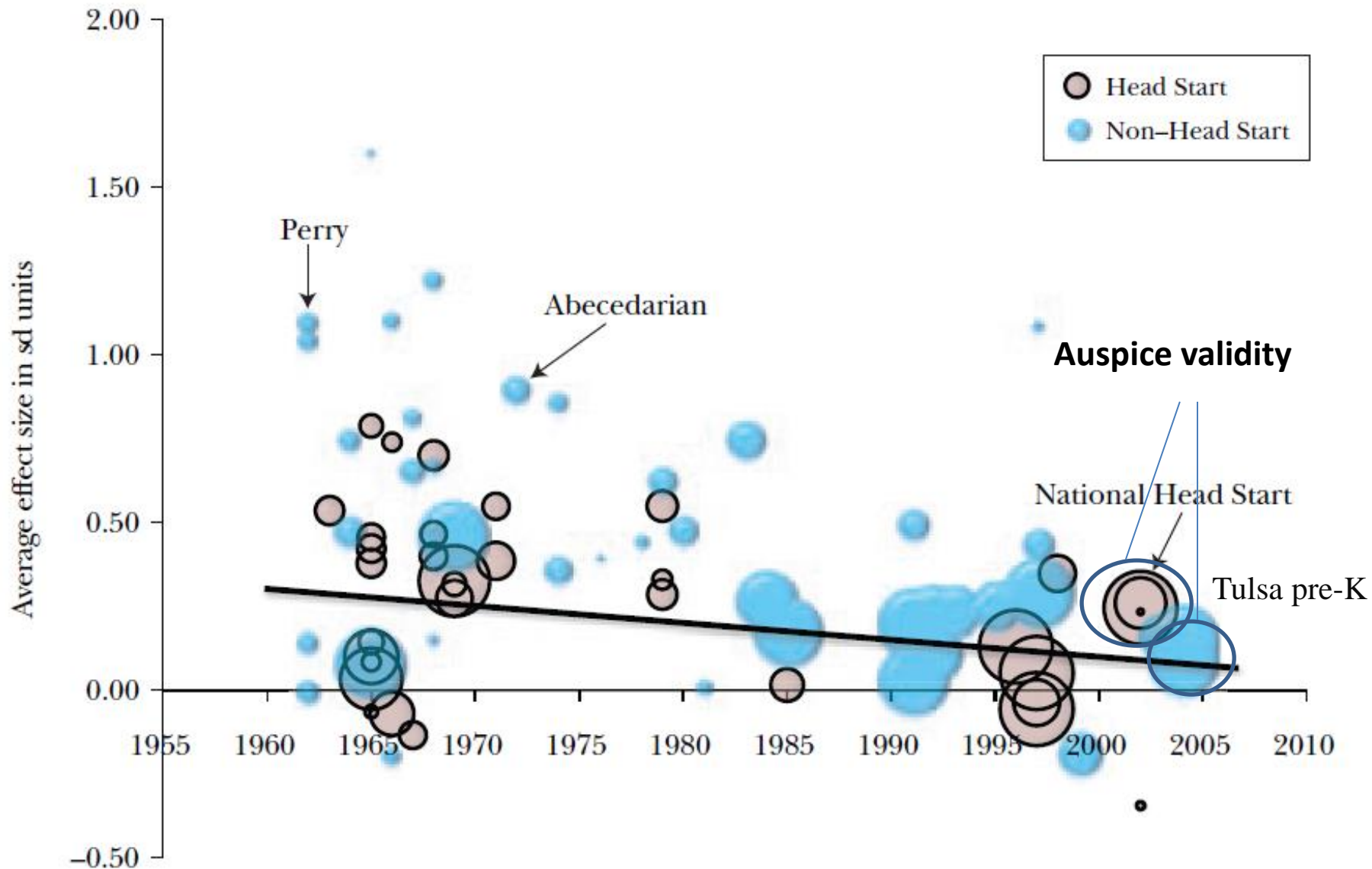
Average Impact of Early Child Care Programs at End of Treatment

(standard deviation units)



Average Impact of Early Child Care Programs at End of Treatment

(standard deviation units)



How Children Are Raised Matters

