



“ . . . Peter [Rossi] believed the evaluator brought one distinctive and indispensable ingredient to the mix of advocacy, ideology, politics, and entrenched interests that shape social programs: impartial facts. As many commentators have noted, Peter was someone for whom, above all, the facts were primary—the proper foundation for social programs, for politics, for theory, and very much the proper basis for any contribution program evaluators might make to any of these.” (p. 201)

Lipsey, M.W. (2007) Peter H. Rossi: Formative for program evaluation. *American Journal of Evaluation*, 28(2), 199-202.



Roiling the Waters: Controversy over the First Longitudinal Randomized Study of a State Pre-K Program

Mark W. Lipsey (with Dale C. Farran
and an outstanding research team)
Vanderbilt University

Peter H. Rossi Award Session, APPAM
November 9, 2018
Washington, DC



PART 1: THE CONTEXT



Growth of Public Pre-K

- **Head Start**
 - Federal legislation in 1966; authorized as part-day 9-month program
 - Enrollment of about 200,000 in 1967; close to 1.0 million in recent years
- **State pre-k**
 - Prior to 1980, only two states funded pre-k programs
 - Growth spurt beginning in the '80s
 - 28 states by 1991
 - 40 states by 2001
 - 43 states in 2017
 - Enrollment increase nationally from about 290,000 in 1991 to about 1.5 million currently; includes nearly 1/3 of all 4 year olds.

NIEER State Preschool Yearbooks for 2003 and 2017; Head Start Federal Funding and Funded Enrollment History, Administration for Children and Families, HHS.

Prior Research on the Effects of State Pre-K Programs

- **End of pre-k effects**
 - No randomized studies; majority are age-cutoff RDD studies modeled on the early Tulsa study
 - **Near universal positive effects on emergent literacy, language, math**
- **Longer-term effects (end of K and beyond)**
 - No randomized studies
 - Majority are post hoc matched with no true baseline measures (matched only on demographic variables), typically with state achievement test outcomes in 3rd grade, some retention and special ed
 - Most of the rest are DID studies with aggregate units (e.g., states, counties) and 4th grade NAEP outcomes
 - **Overall notable vulnerability to bias, inconclusive mixed findings**

The Long Shadow of the Early Longitudinal Research & Demonstration Projects

	Perry Preschool	Abecedarian
Preschool cohorts	1962-67	1972-77
Sample size	N=123	N=111
Sample characteristics	Highly disadvantaged African-American children	Low income, 98% African-American, 83% single parent
Program	Part day, school year, 2 years for 75%, high staff/student ratio, home visits	Full day, all year, 5 years, high staff/student ratio, health care, home visits
Academic outcomes	Achievement + Retention --- Special ed + Graduation +	Achievement + Retention + Special ed + Graduation ---
Adult outcomes	Employment + Income + Crime +	Employment + Further education + Health +

“Robust Belief” in the Long-term Effects of Scaled-up Publicly-Funded Pre-K Programs

- “Recent evaluations show quite clearly that large-scale, publicly funded pre-k programs have significant benefits for children’s school readiness skills and future academic achievement. Importantly, these results come from today’s state-funded pre-k programs, operated at scale in various states across the country.” (Pianta & Wolcott/Raise Your Hand Texas)
- “Effects [of state and local pre-k programs] vary across programs, but are overwhelmingly positive.” (Barnett/NIEER)
- “What many may not be aware of, however, is that a vast and emerging body of research continues to demonstrate the potential of publicly funded, large-scale pre-k programs as a strategy for school reform and turning around a record of underachievement.” (Wat/PEW Center on the States)
- “Effective pre-kindergarten programs reduce costly grade retention, special education, and other services right away. ... The benefits continue into early elementary school, including the critical third-grade benchmark year.” (Ready Nation)

PART 2: THE TENNESSEE PRE-K STUDY

Funded by DOE/IES (R305E090009) and NICHD (1R01HD079461-01)
with no endorsement or responsibility for the contents of this talk.

Full Randomized Sample (RCT)

- Oversubscribed schools asked to admit students in order from randomized lists of applicants until seats filled
- **79 schools, 29 districts, 111 randomized applicant lists**
- **2990 children*: 1852 admitted to VPK, 1138 not admitted**
- Tracking through the state data system to 3rd grade and beyond

Cohort 1: Pre-k 2009-10 (N=1744)

Cohort 2: Pre-k 2010-11 (N=1246)



25% with parental consent
18% eligible for analytic sample

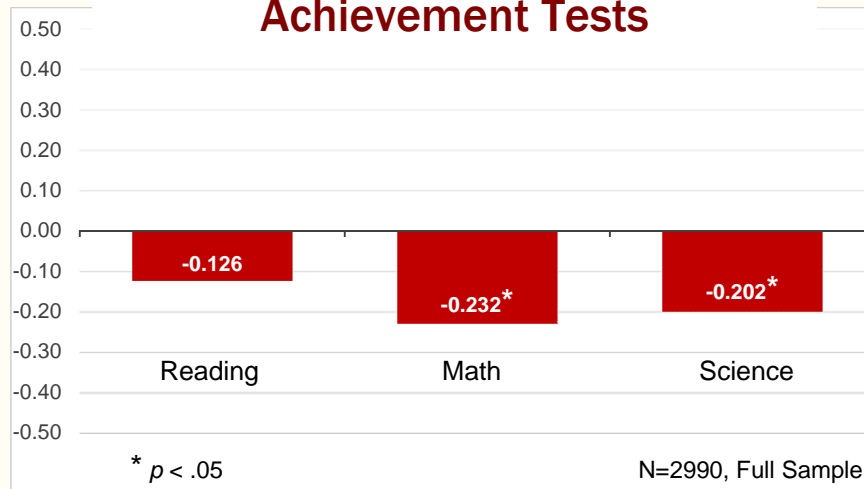
70% with parental consent
61% eligible for analytic sample

Intensive Substudy Sample (ISS)

- **47 schools, 21 districts, 76 randomized applicant lists**
- **1076 children: 773 VPK participants, 303 nonparticipants**
- Assessed by the research team before and after pre-k year and after each school year through 3rd grade
- **70% of control group not in organized preschool program; remainder divided between Head Start (12%) and private childcare (15%)**

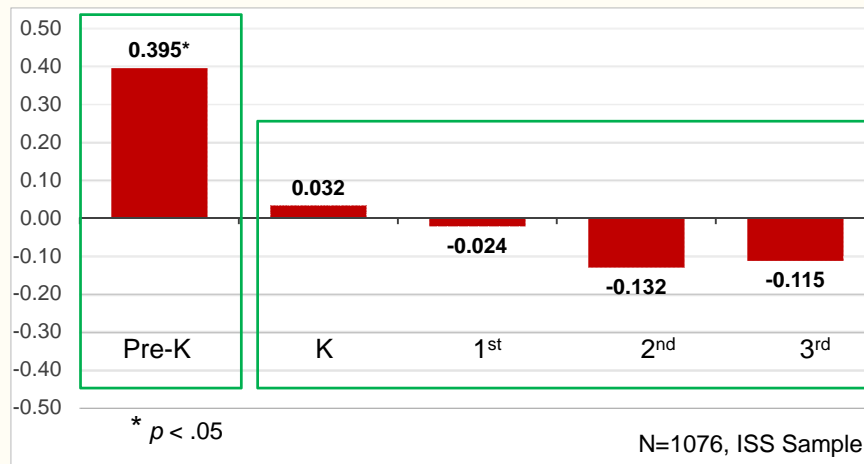
* Children from randomized lists with both treatment and control cases who appear in the state database at any time through 3rd grade.

Effect Sizes for Third Grade State Achievement Tests



2SLS IV CACE/TOT estimates, multiple imputation

Effect Sizes for Woodcock Johnson Achievement Composite



2SLS IV CACE/TOT estimates, multiple imputation

Other Results through Third Grade

Cumulative K-3rd grade retention

- **No difference:** 13.7% vs 13.7%; ES= .002

School disciplinary events, K-3rd grade

- School rule violations; **more for pre-k**
7.1% vs 4.3%; ES= .123 ($p < .10$)
- Major offenses; **no difference:** 3.4% vs 3.5%; ES= -.006
- Any offense; **no difference:** 8.4% vs 6.7%; ES= .062

Special education placements, K-3rd grade

- IEP except for physical disability or gifted:
More for pre-k 14.6% vs 9.4%; ES= .156 ($p < .05$)

Preliminary Results from Follow-up into Later Grades

State Achievement Test Scores: ITT Effect Sizes			
Grade & Cohort	Reading	Math	Science
4 th (Cohort 1)	-.11*	-.16*	-.16*
5 th (Cohort 2)	-.13*	-.14*	-.07
6 th (Cohort 1)	-.19*	-.23*	-.20*

* $p < .05$

- Retention: **No pre-k effect** through 5th grade
- School disciplinary events (5th grade):
More for pre-k; statistically significant for major offenses
- Special education placements: **Higher for pre-k** through 5th grade
- Significant **negative effects** in 5th grade on:
 - Executive function (ES= -.17)
 - Commitment to school (ES= -.15)
 - Problem behaviors (ESs -.22 to -.14)

“No good applied social research goes unpunished” (Rossi, 1987)

Rather vigorous responses to reports of these
“anomalous” pre-k results from Tennessee:

- Advocacy groups
- Researcher blogs
- Media coverage
- Peer reviewers

Rossi, P.H. (1987), No good applied research goes unpunished. *Society*, 25(1), 73-79.



PART 3: ISSUES AND EXPLORATIONS



Incompetent Analysis?

Our primary analysis approach

- Typical in educational and psychological research
- Random effects hierarchical linear models (HLM)
- Baseline covariates included in all analyses; propensity scores in some
- ITT and CACE/TOT estimates via 2SLS instrumental variable models
- Multiple imputation for missing values
- Myriad sensitivity analyses

Alternative (more appropriate?) approach

- The way economists do it
- Tyler Watts, Greg Duncan, & Mariela Rivas. *A Reanalysis of the Impacts of the Tennessee Voluntary Prekindergarten Program.*

Saturday 3:15-4:45, Marriott Balcony B, Mezzanine Level

Poor Quality of the Tennessee Program?

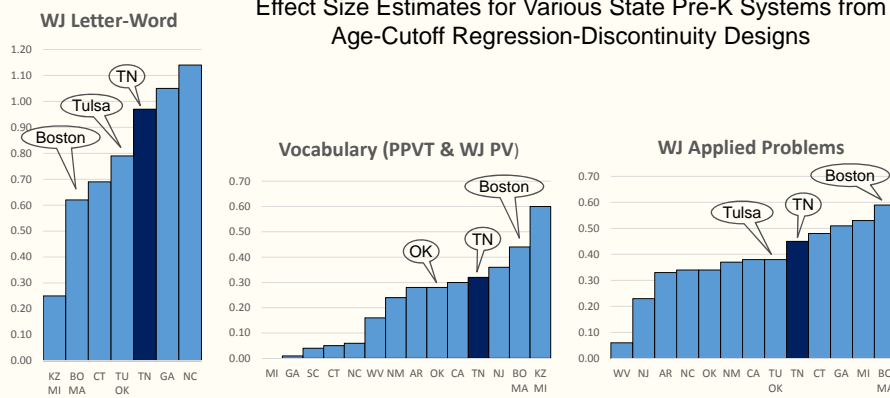
Structurally typical state pre-k program

- Organized and overseen by the state department of education
- Serves 4-year-old children from low-income families statewide
- Local programs in all but a few of the school districts in the state
- Full day, 5 days/week, school year; most classrooms in elementary schools
- State-licensed teacher with early childhood endorsement & credentialed aide
- Maximum class size of 20; 1:10 staff to student ratio
- Curriculum selected from a state-approved list

Limited basis for comparisons on presumptive quality measures

- Meets 9 of the 10 standards advocated until recently by the National Institute of Early Education Research (NIEER)
- Mean total scores from classroom observations on the *Early Childhood Environment Rating Scale (ECERS)*:
4.15 for TN vs mean of 4.34 found for 14 other state programs

Better Perhaps: Direct Comparison of Pre-K Effects on Common Outcome Measures



Little Support for Pre-K Gains in Later Grades

Redundancy in instructional content of pre-k and K

- Systematic observations in 103 pre-k and 98 K classrooms^a found:
 - For time on literacy (28% pre-k, 23% K), 43% in pre-k and 58% in K devoted to foundational skills (phonological awareness, letter and word recognition, etc.)
 - For time spent on math (12% pre-k, 9% K), 64% in pre-k and 63% in K devoted to counting and cardinality
 - Little differentiation for proportion of K children who had participated in pre-k

Between beginning of K and the end of 1st, teacher ratings of “feelings about school” changed from more positive for pre-k participants to more negative

Few of the pre-k sample attended high quality elementary schools and/or were exposed to multiple years of highly rated teachers (<20%)

- Analysis of effects on state achievement tests inconclusive because of limited variation

^a Farran et al. (2018). *Kindergarten follow-up report*. Peabody Research Institute, Vanderbilt University

Enhancing the Potential of State Pre-K

A substantial infrastructure has been established for serving at-risk children

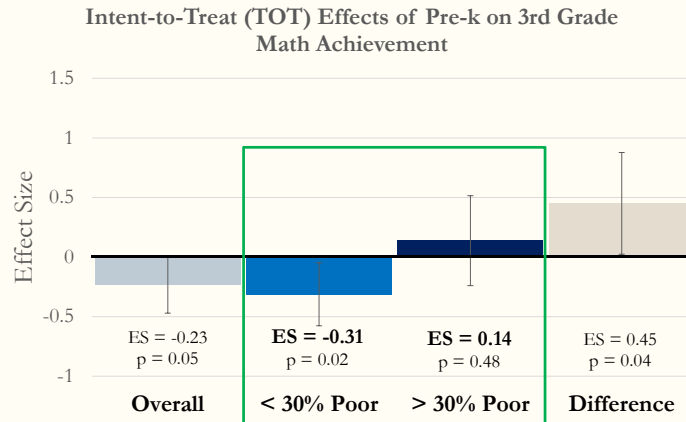
- In Tennessee (2017-18): 934 classrooms in 135 of 146 districts serving more than 18,000 children from low income families.

Analysis of pre-k classroom observations and language, literacy, math, and self regulation gains has identified features related to greater gains^a:

- Relatively little transition time (routines and wait time for children)
- High quality of instruction (open-ended questions, back-and-forth interactions)
- Positive emotional climate (fewer behavior disapproving, more behavior approving comments)
- Teachers listening to children (vs. dominance of teacher talking)
- Sequential learning activities (linked tasks and activities)
- Social learning interactions among children (associative and cooperative)
- High child involvement (active learning opportunities and engagement)
- Extended math learning opportunities.

^aFarran et al. (2017). Data-driven improvement in prekindergarten classrooms: Report from a partnership in an urban district. *Child Development*, 88(5), 1466-1479

Pre-K Effects on State Math Test for Children in High vs Lower Poverty Neighborhoods



Pearman, F.A. (2018). The moderating effect of neighborhood disadvantage on preschool effectiveness: Evidence from the Tennessee Voluntary Prekindergarten Experiment. Under review.

**More questions than answers
(no surprise there)
And so this work continues . . .**

**Contact information:
mark.lipsey@vanderbilt.edu**