



Left Behind by Design: Proficiency Counts and Test-Based Accountability

(forthcoming, *Review of Economics and Statistics*)

Derek Neal and Diane Whitmore Schanzenbach

University of Chicago and NBER

Responses to NCLB



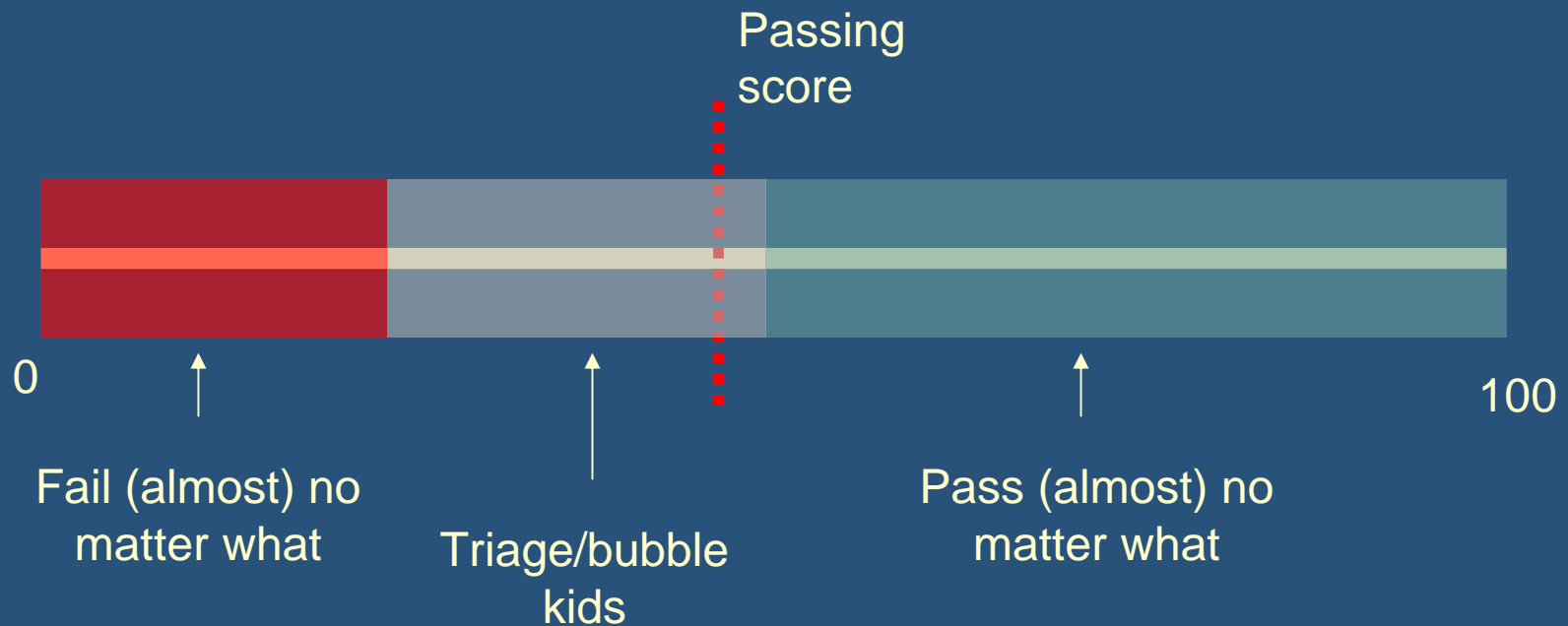
- Complex incentives & responses
- Our Paper
 - Difference-in-difference results that document changes in patterns of achievement
 - NCLB's implementation in Chicago
 - Earlier NCLB-style accountability policy
- Qualitative work (Booher-Jennings in TX)
 - Gets inside “black box” of NCLB
 - Describes how teachers change their behavior
- Popular press articles

“We were told to cross off the kids who would never pass. We were told to cross off the kids who, if we handed them the test tomorrow, they would pass. And then the kids who were left over, those were the kids we were supposed to focus on.” -- de Vise, *Washington Post*, 3/4/07

Educational Triage



- Incentives in most states: only passing or failing matters
 - No credit for moving students' test scores up unless they cross the passing threshold
 - Decide how to allocate “extra” attention



Previous Research



- Schools respond to incentives (Jacob, Cullen & Reback, Reback, Figlio, Rouse & Figlio) by
 - changing *what* they teach
 - Effort to high-stakes subjects, grades
 - changing *who* they teach
 - Strategic assignment to special ed/ELL
 - “bubble” students

Two Policy Experiments



- NCLB in Illinois
 - Reasonably high passing threshold
 - Years of failure defined retroactively
 - Schools faced high sanctions in first year
- Pre-NCLB Accountability in Chicago
 - School probation based on proficiency rates in reading on ITBS
 - Threshold pretty high: national norms
 - Student incentives: summer school, retention
 - lower passing thresholds for these

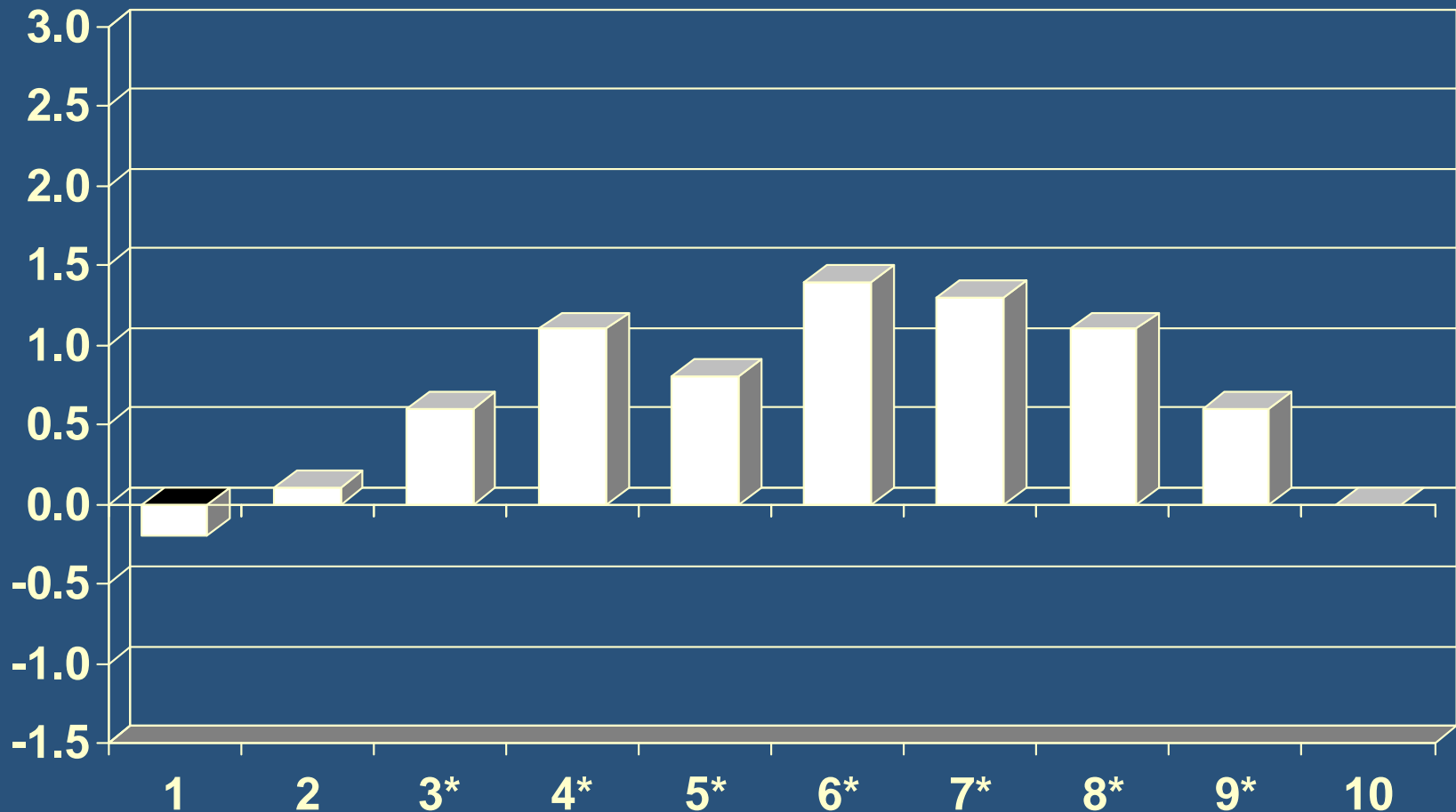
Difference-in-Difference Approach: Each Baseline Decile



Grade

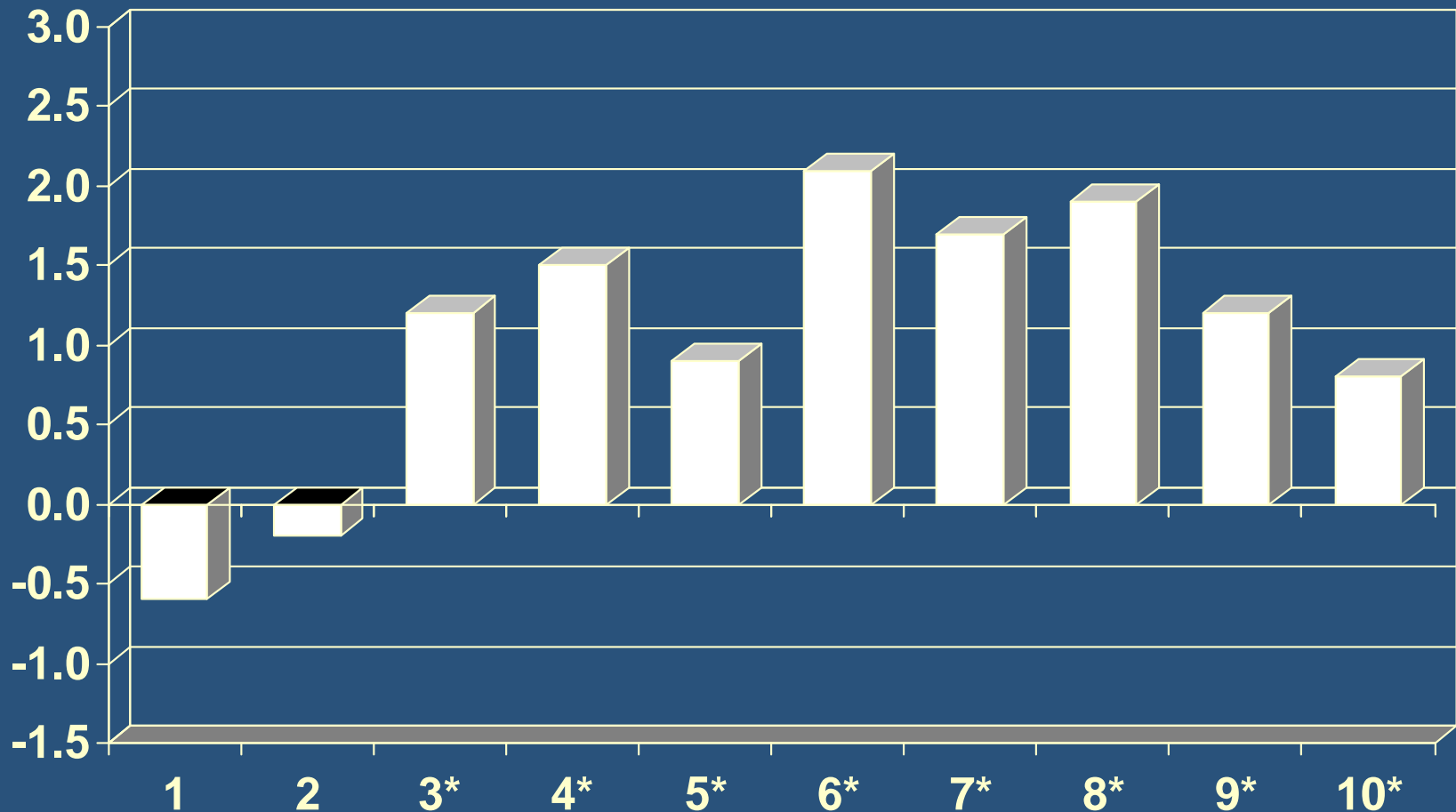
	3rd	5th	
<i>Cohort</i>			
Pre-reform	low	low	$5_{\text{pre}} - 3_{\text{pre}}$
Post-reform	low	high	$5_{\text{post}} - 3_{\text{pre}}$
			<hr/> D-D

NCLB: Change in 5th Grade Reading Scores (2003 vs. 2002)



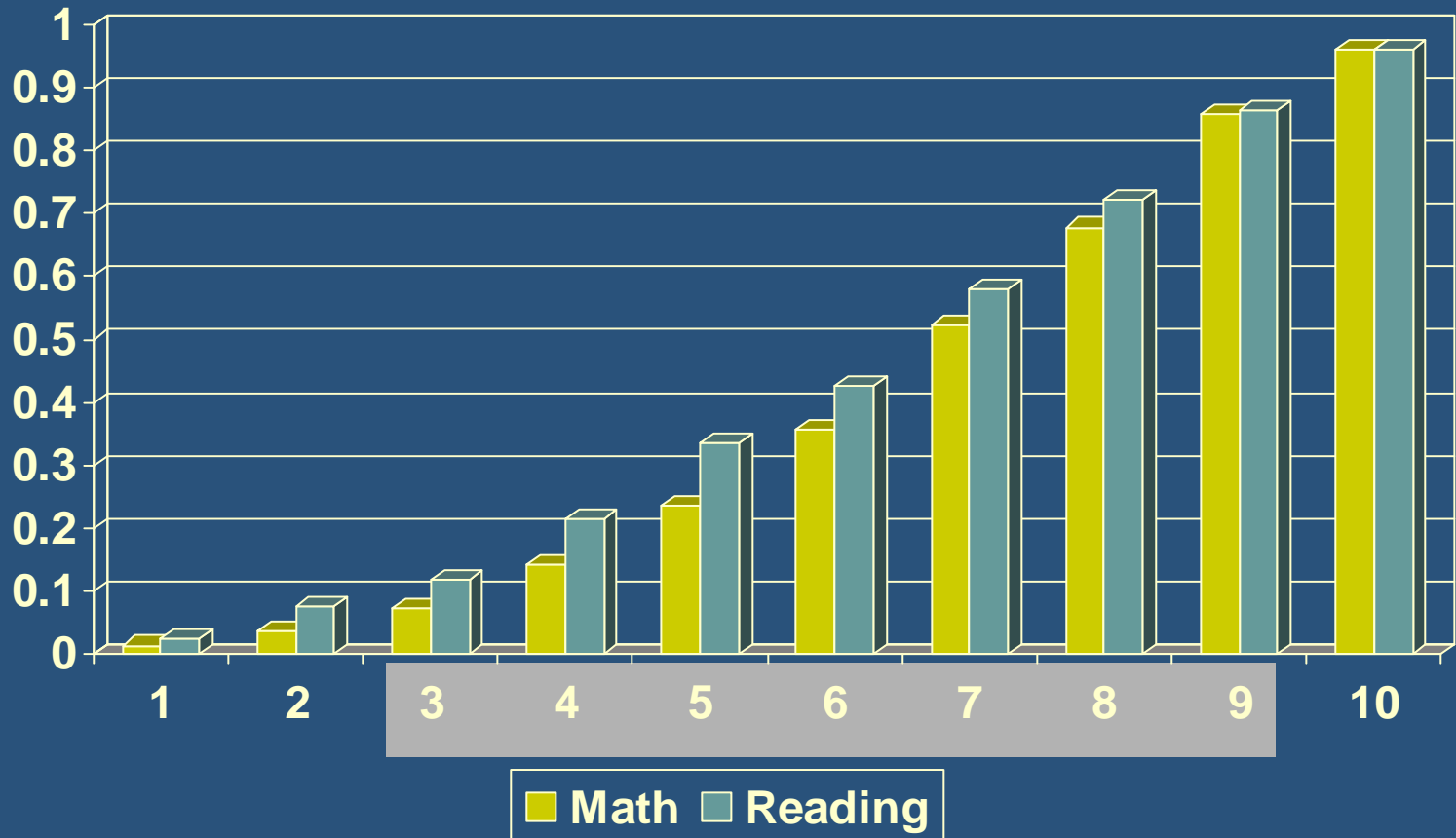
Note: * denotes that the difference is statistically significantly different from zero.

NCLB: Change in 5th Grade Math Scores (2003 vs. 2002)

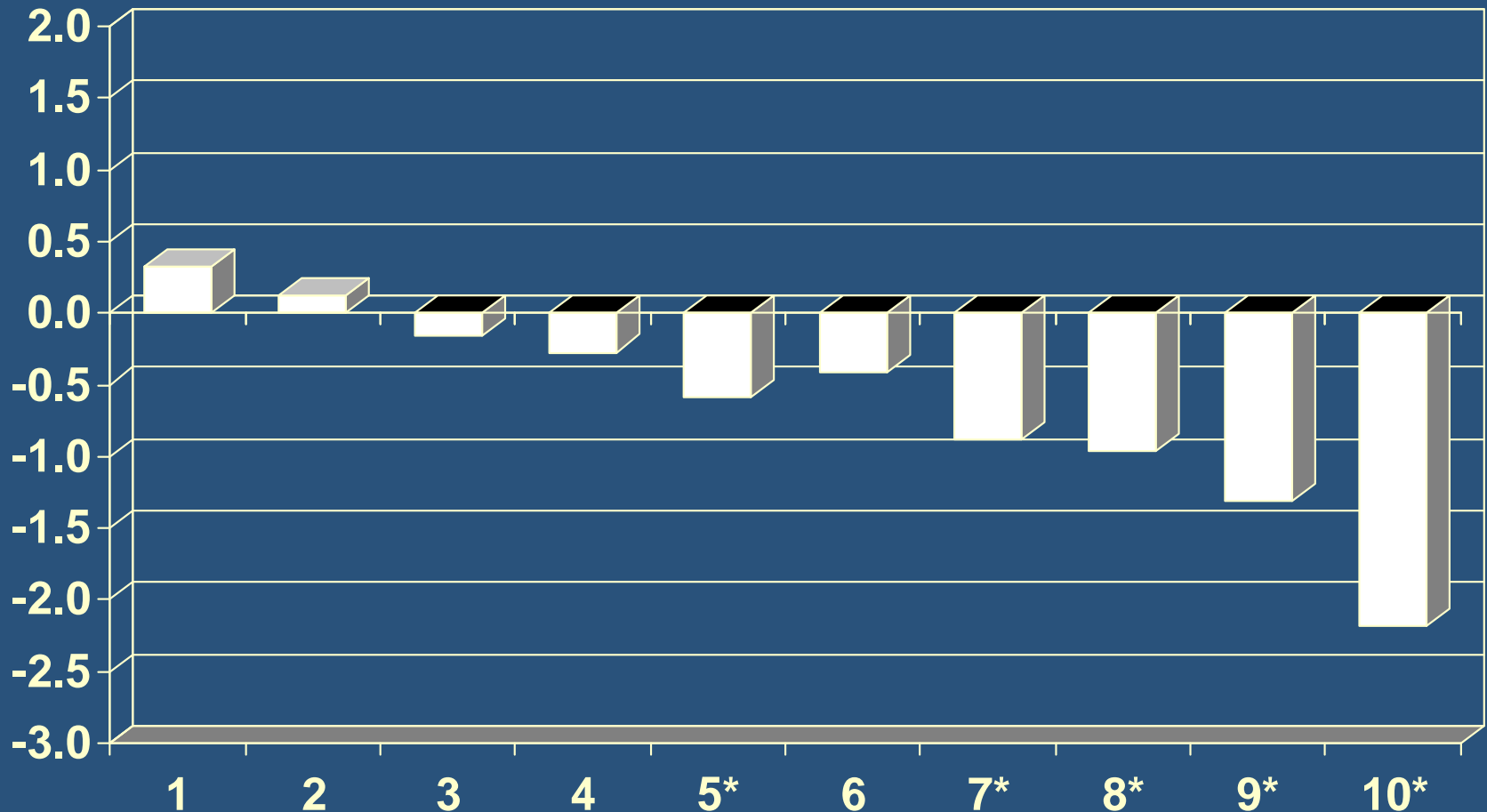


Note: * denotes that the difference is statistically significantly different from zero.

NCLB: Expected Proficiency in 5th Grade

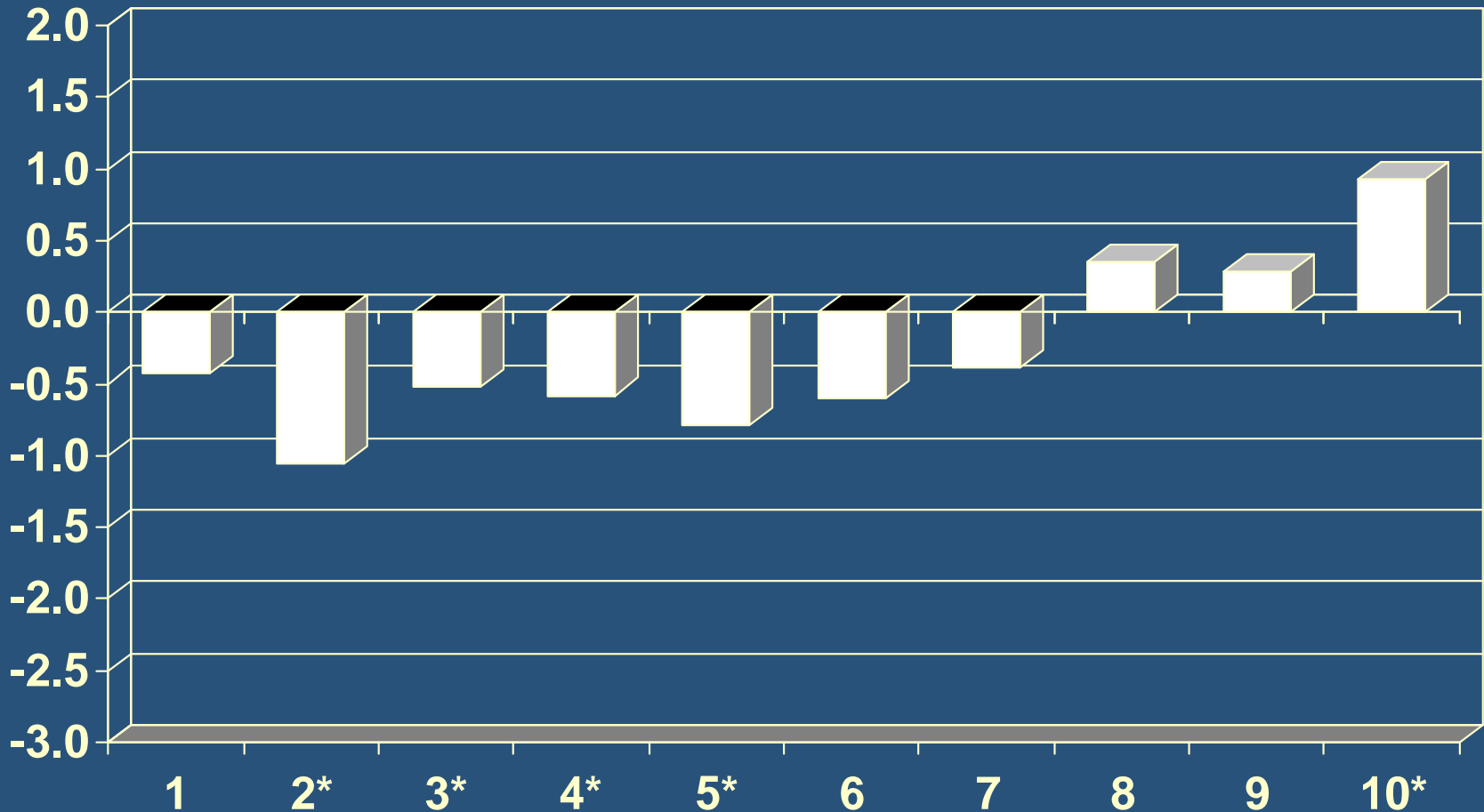


Placebo Test: Change in 5th Grade Reading Scores (2005 vs. 2004)



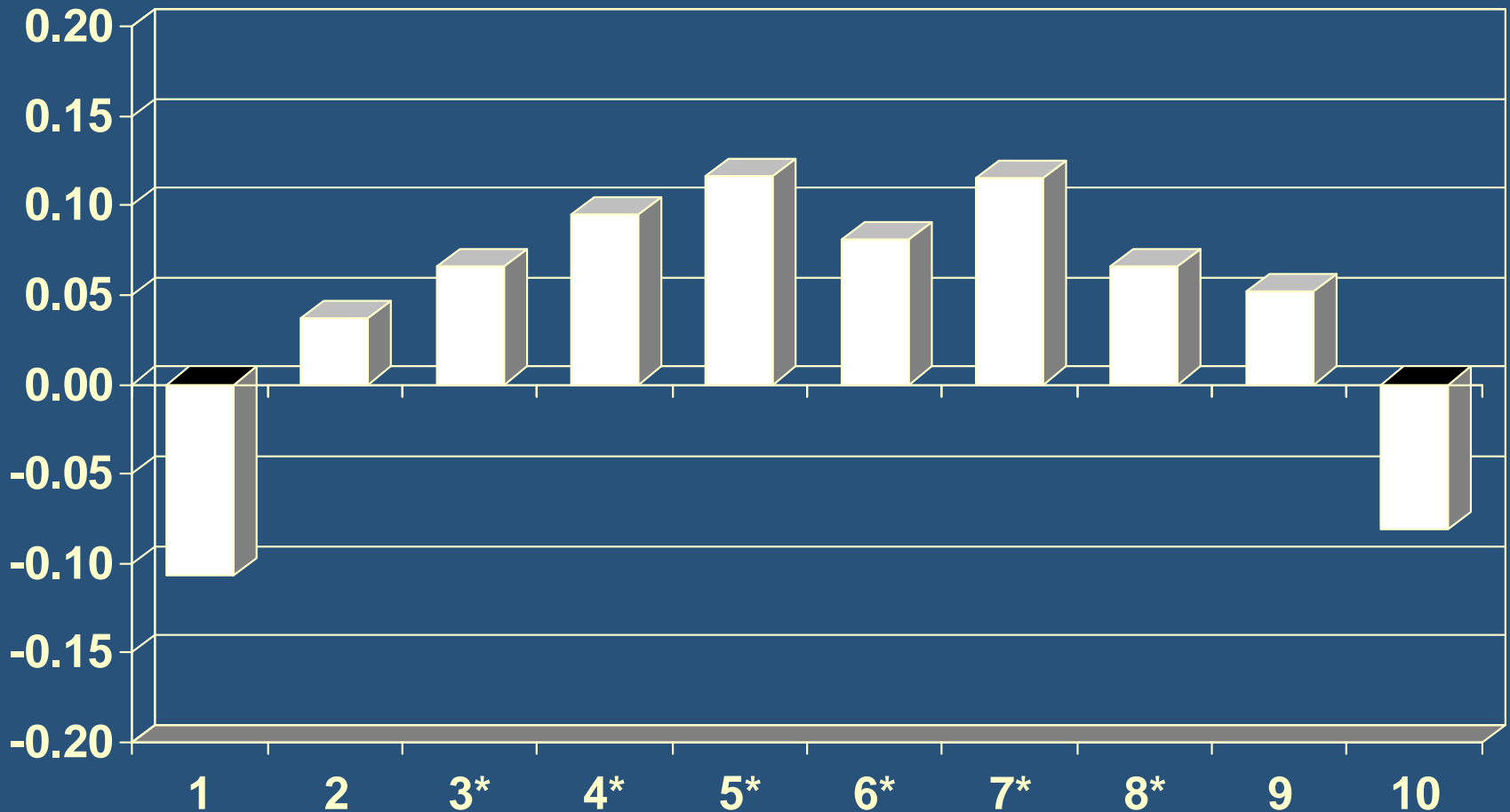
Note: * denotes that the difference is statistically significantly different from zero.

Placebo Test: Change in 5th Grade Math Scores (2005 vs. 2004)



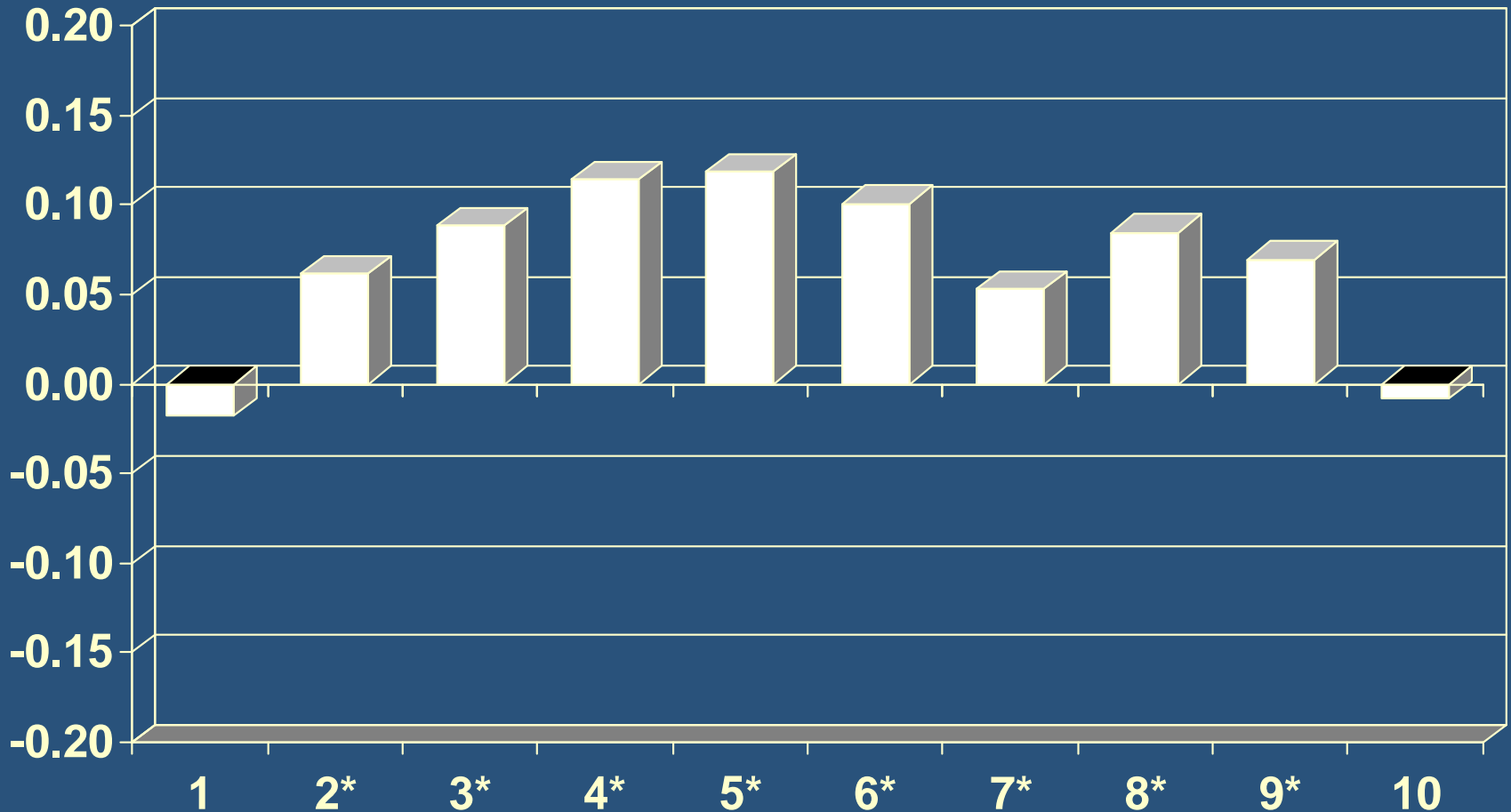
Note: * denotes that the difference is statistically significantly different from zero.

Chicago Accountability: Change in 5th Grade Reading Scores (1998 vs. 1996)



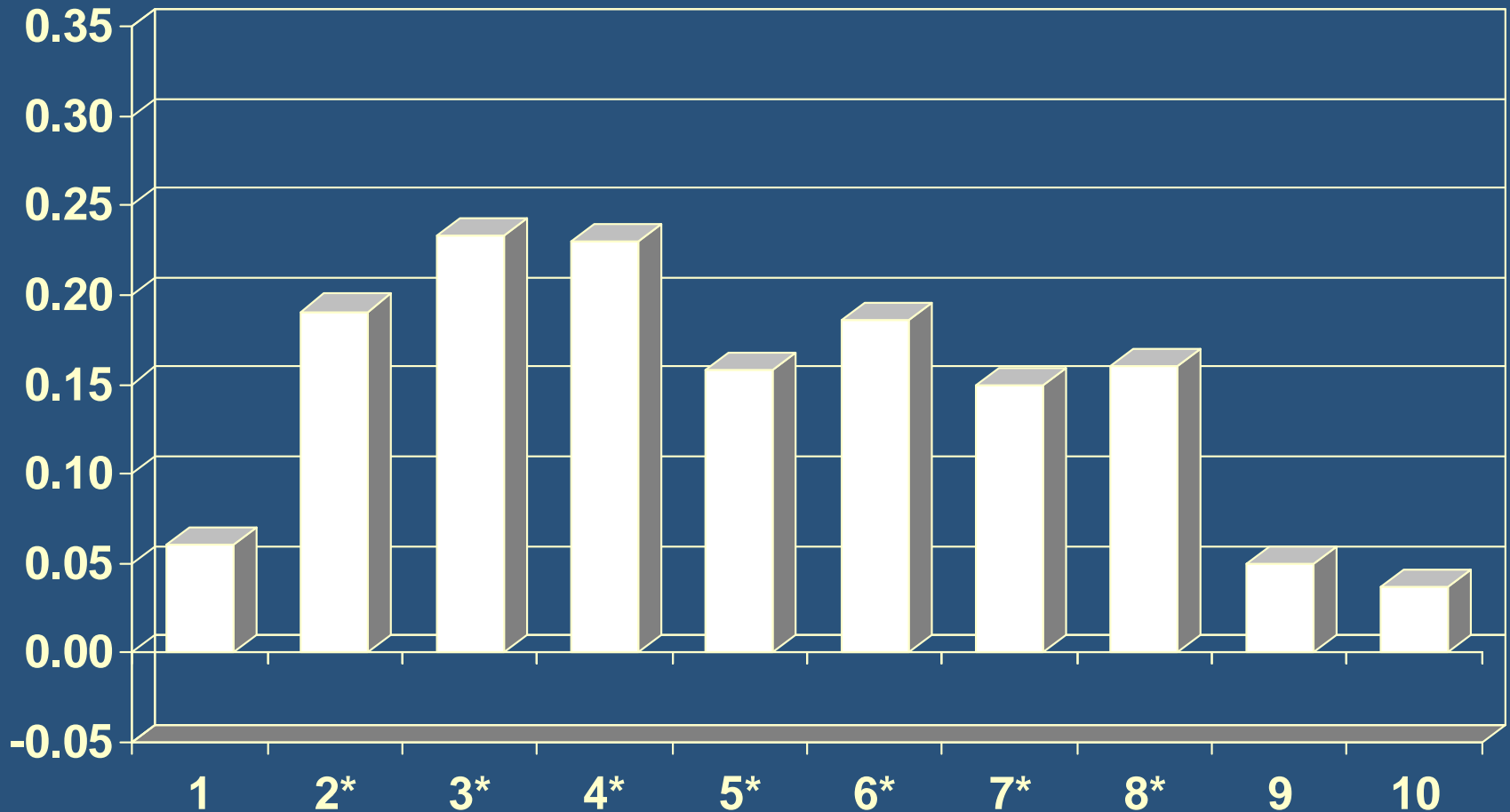
Note: * denotes that the difference is statistically significantly different from zero.

Chicago Accountability: Change in 5th Grade Math Scores (1998 vs. 1996)



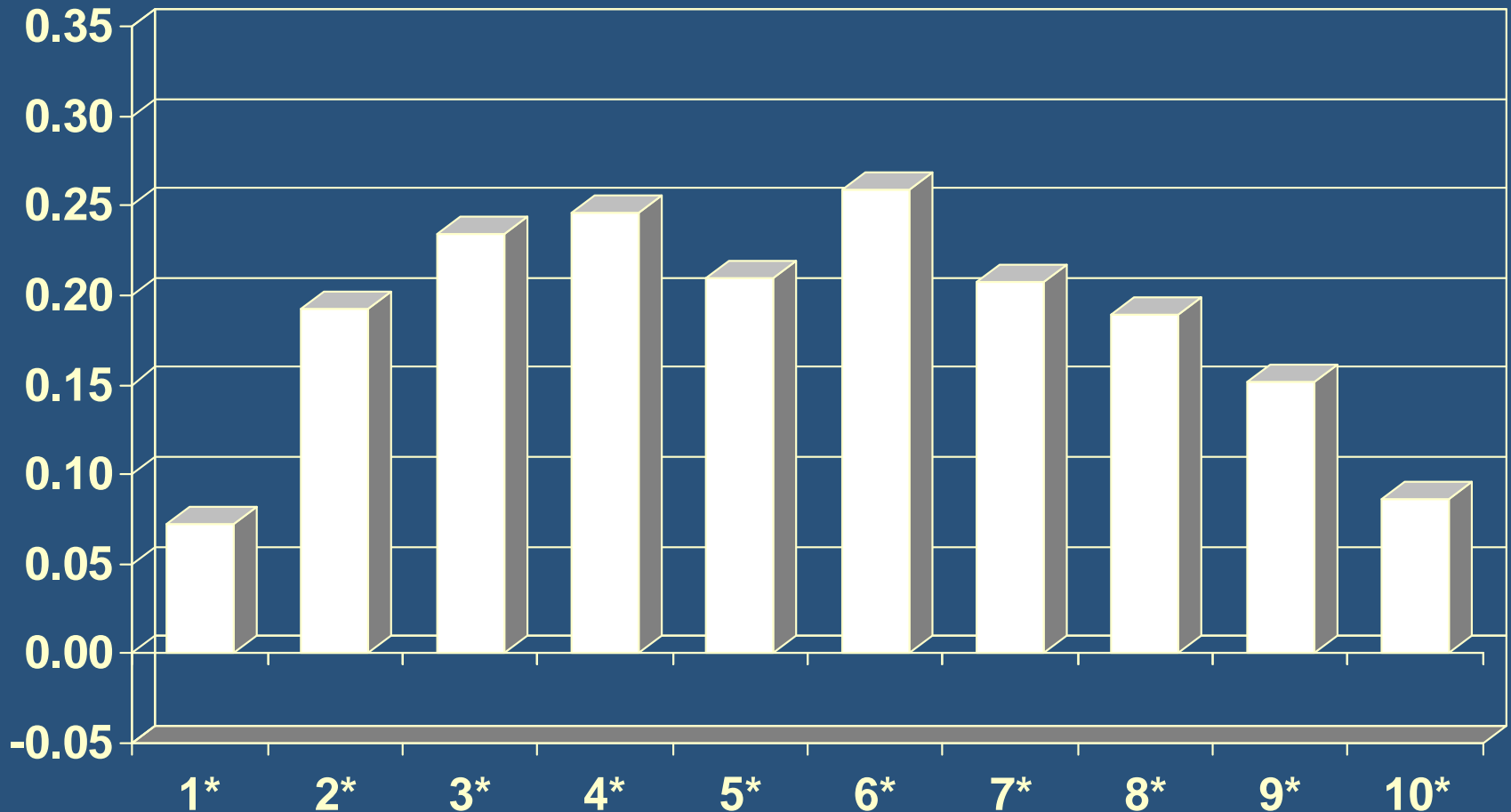
Note: * denotes that the difference is statistically significantly different from zero.

Chicago Accountability: Change in 6th Grade Reading Scores (1998 vs. 1996)



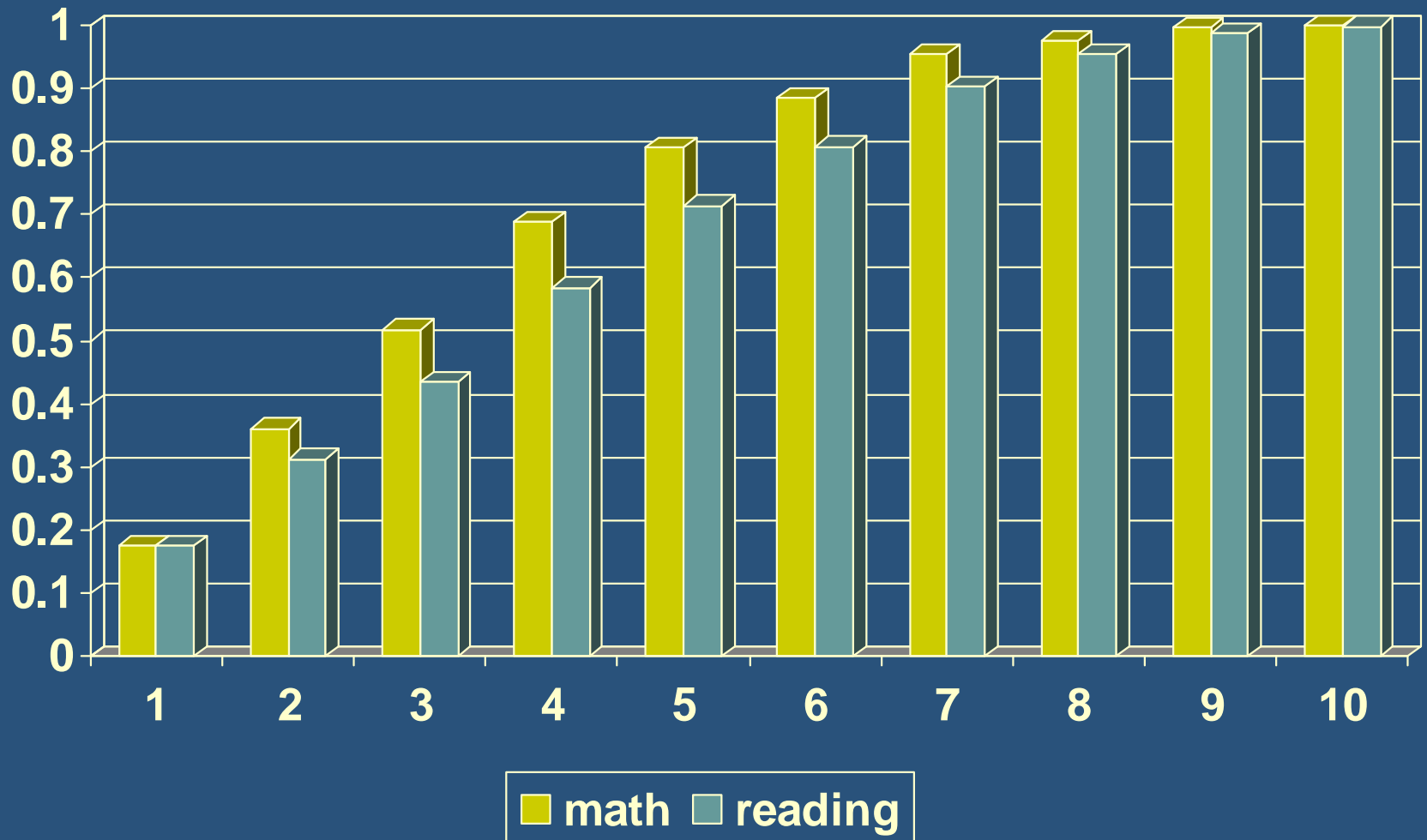
Note: * denotes that the difference is statistically significantly different from zero.

Chicago Accountability: Change in 6th Grade Math Scores (1998 vs. 1996)



Note: * denotes that the difference is statistically significantly different from zero.

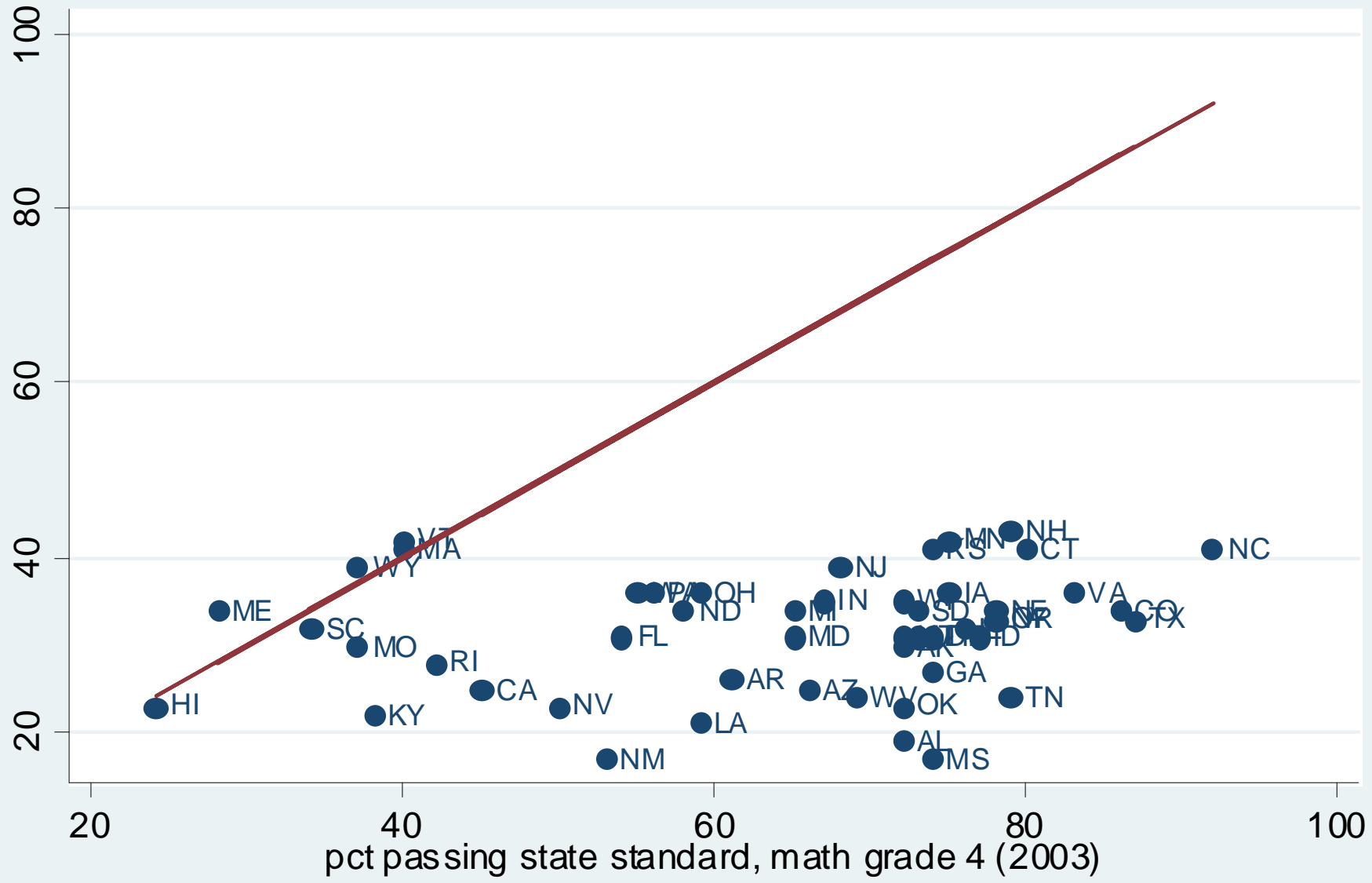
Chicago Accountability: Expected Summer School Pass Rate in 6th Grade



Policy: Race to the Top



- Rhetoric: “soft bigotry of low expectations”
- States originally encouraged to set high standards
 - Picture
- States *again* encouraged to set high standards
- Our model predicts if standards are high
 - > more high achieving kids get extra attention
 - > but can harm low achieving students



● pct proficient+ NAEP math, grade 4 (2003) — yhat

Policy: Value-Added Measures



- Need a system that gives some credit for moving children up in achievement throughout the distribution
 - Perverse incentives for best teachers now
- Value-added
 - Hard to get scales right
- Multiple thresholds (MA model)
- Previously the Dept of Ed limited # of states allowed to do this

Conclusions



- NCLB as currently structured gives incentive to leave worst (and best) scoring students behind, concentrate on middle
- Straightforward empirical evidence documenting that schools respond to this incentive by changing their behavior in such a way that students in the middle were helped
- Taken in conjunction with qualitative evidence, appears that teachers redistribute effort and attention toward “bubble” students