

The Impact of NCLB on the Level and Distribution of Student Achievement: Evidence from the NAEP Data

August 12, 2009

Thomas S. Dee, Swarthmore College
Brian A. Jacob, University of Michigan

Our Research Questions

- Has NCLB increased overall levels of student achievement?
 - ▣ Has the policy had a differential impact on Black, White and Hispanic students?
 - ▣ Has the policy had a differential impact on high- versus low-achieving students?
 - ▣ Has the impact of the policy evolved from 2002 through 2007?
 - ▣ Is there any evidence for the mediating factors by which school accountability operates?

Key features of our study

- Evidence from an analysis of state-by-year NAEP data (1992-2007)
 - ▣ Extension of results presented as the 2008 Kershaw Lecture at APPAM
- NAEP data have a comparatively high degree of “external validity”
 - ▣ Representative student-achievement data from most U.S. states both before *and* after NCLB (Table 1)
 - ▣ NAEP measures performance on a “low-stakes” test, not one that is the focus of NCLB-inspired state assessment systems
 - ▣ Comparatively immune to concerns about “teaching to the test” & potentially confounding changes in state assessment systems (Fuller et al. 2007)
- Research design based on comparing pre/post-NCLB changes across...
 - ▣ Public & private schools
 - ▣ States that already had school accountability & those that did not.

Quick Preview of Results

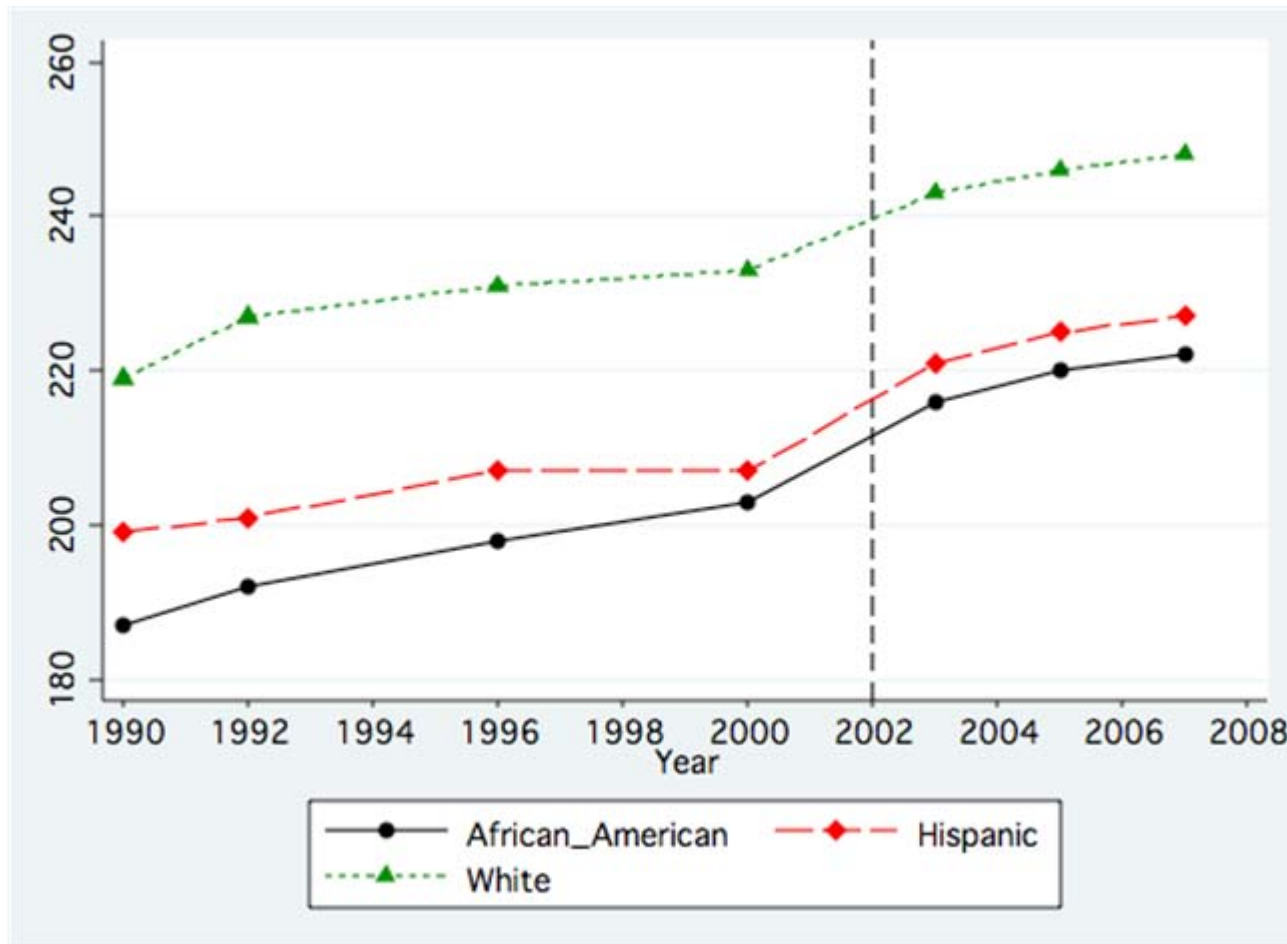
- NCLB increased mean 4th grade math achievement by 6.8 scale score points (0.22 effect size)
 - ▣ Similar effects at low and high achievement levels
 - ▣ Larger effects for Hispanic and Black students
- Less consistent evidence for similarly large and broad achievement gains in reading or among 8th graders.

Alternative views of trend data

- Achievement trends on state assessments and NAEP are “positive overall and for key subgroups” since NCLB (Stullich et al. 2006)
- Reading and math achievement has increased in most states since 2002 (Center on Education Policy 2008)
- *Growth* in grade-4 achievement has “largely faded since...NCLB” (Fuller et al. 2007)
- Reading trend is flat since NCLB and math growth merely tracks prior trends (Lee 2006)

4th Grade NAEP Math Trends

National Trends by Race



Challenges of Evaluating NCLB

- Time-series analyses lack a credible comparison group
 - ▣ Social, economic or other forces coincident with NCLB (e.g., the economic downturn)
- Recent studies make creative use of state and local data
 - ▣ Often provide a plausible comparison group
 - ▣ But the focus is largely on *relative* achievement changes (e.g., low versus high achievers)
 - ▣ Also, potential concerns about the generalizability of results from state and city-specific data

Finding a Control Group to Assess NCLB with NAEP

- Catholic Schools
 - ▣ Few Catholic school students participate in NCLB programs (DOE 2007)
 - ▣ Catholic schools *may* be subject to similar social and economic trends
- States that implemented “NCLB-like” school accountability policies *prior* to NCLB
 - ▣ Studies of pre-NCLB state accountability reforms (Carnoy and Loeb 2002, Lee and Wong 2004, Hanushek and Raymond 2005)

School accountability before NCLB

- The “consequential accountability” policies implemented in many states were similar to NCLB’s school-accountability provisions (Hanushek and Raymond 2005)
 - ▣ Routine assessment of school performance & the possibility of sanctions (e.g., restructuring) based on that performance
- In fact, several states complained that NCLB “needlessly duplicates” existing state accountability systems (Dobbs 2005)
- Our review of Hanushek and Raymond (2005) coding of state policies → 30 states with consequential accountability prior to NCLB (Table 2)

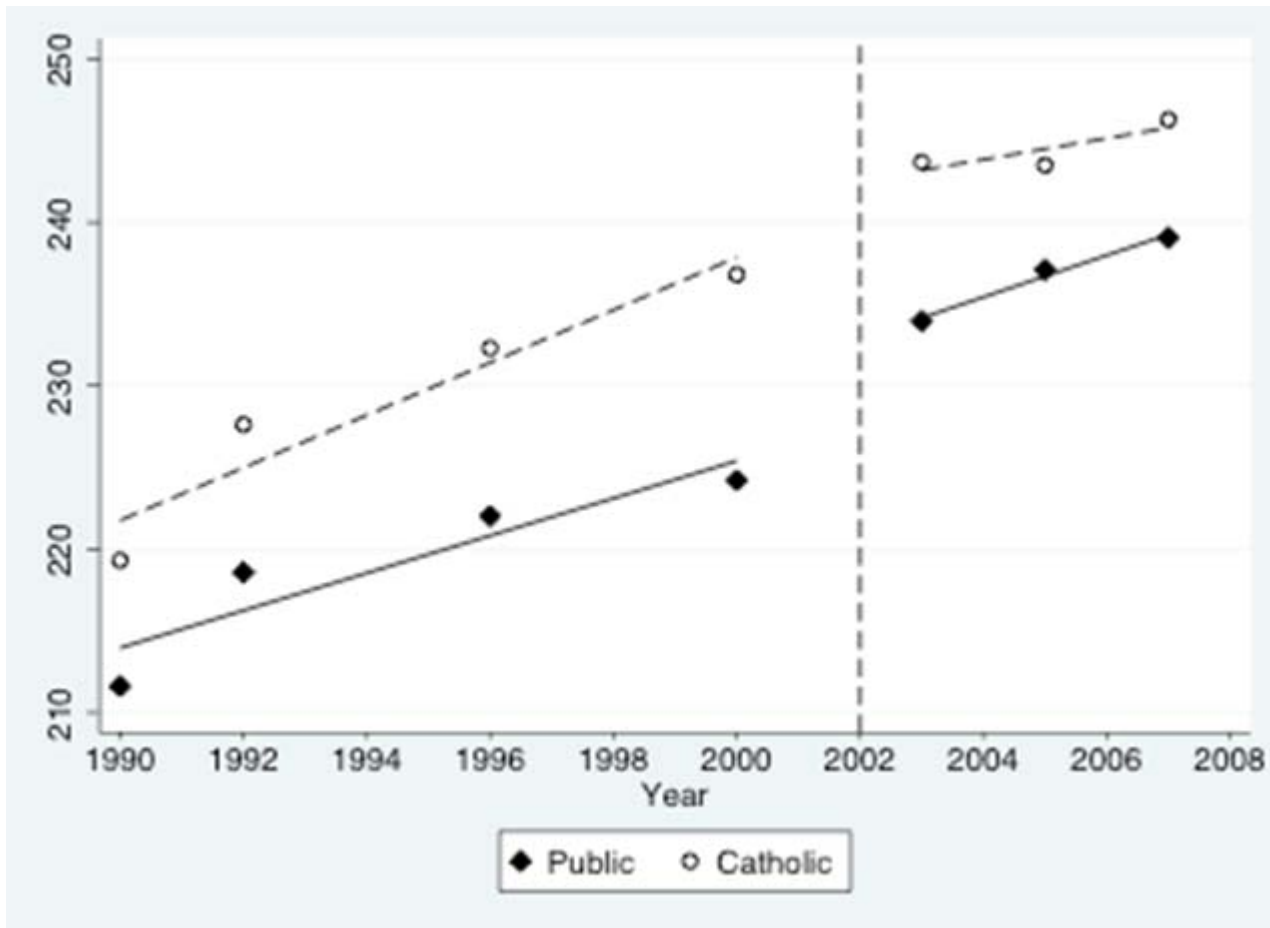
Comparative Interrupted Time Series

- Basic intuition → compare pre/post changes in states with and without “consequential accountability” prior to NCLB
- Comparative changes conditional on...
 - ▣ Pre-existing trends unique to states with and without prior consequential accountability
 - ▣ State fixed effects and a fixed effect for post-NCLB years
- Allow for flexible NCLB effects through potential changes in both achievement levels *and* trends
- Allow for treatment heterogeneity by achievement level (e.g., mean, percent proficient, various percentiles) and by racial-ethnic subgroups

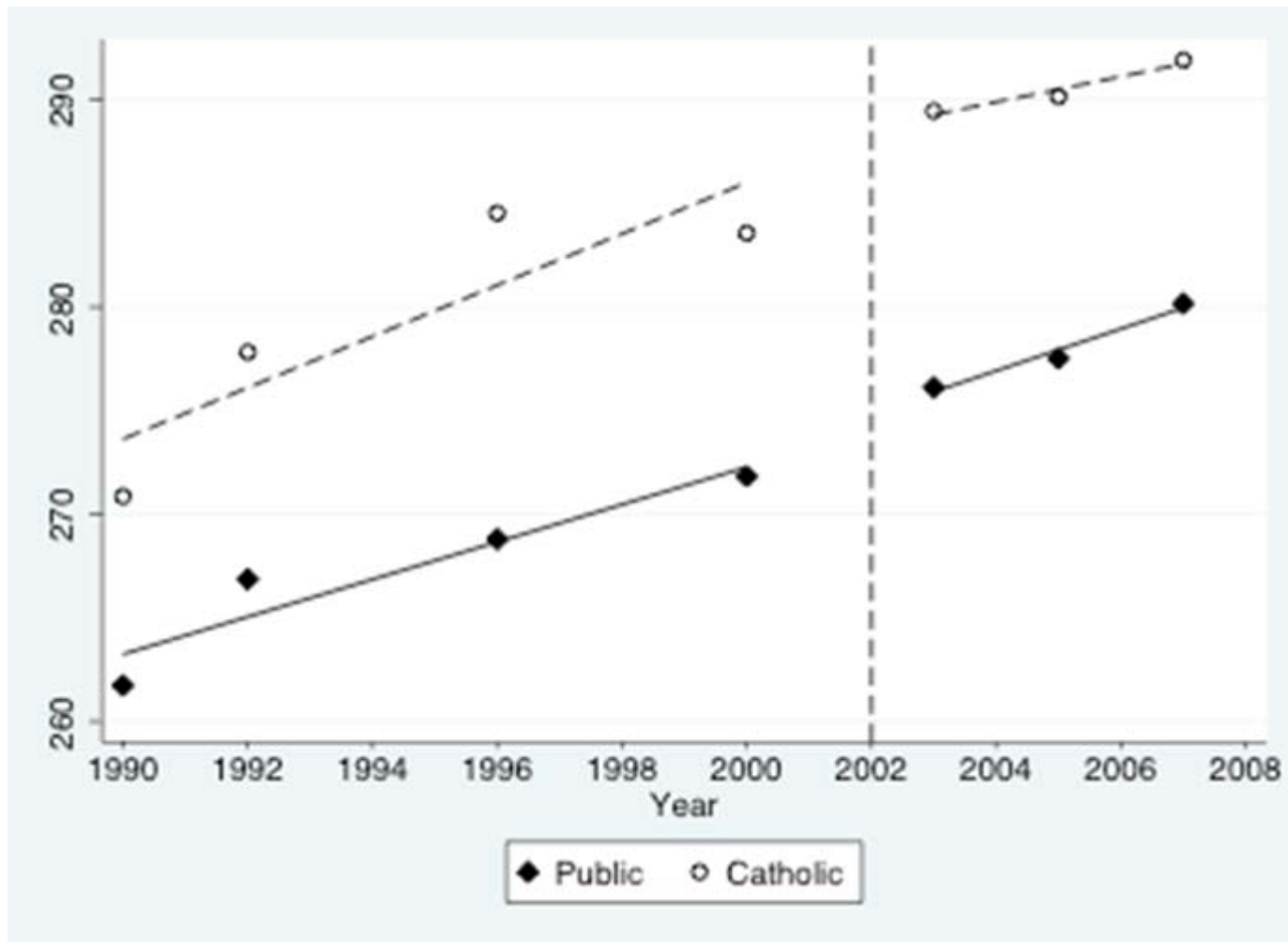
Findings

- Comparisons of public vs. Catholic schools
 - ▣ Potential concern involving compositional changes across groups
- Comparisons across states by prior school accountability
 - ▣ Results appear robust to a variety of alternative specifications (Table 4)

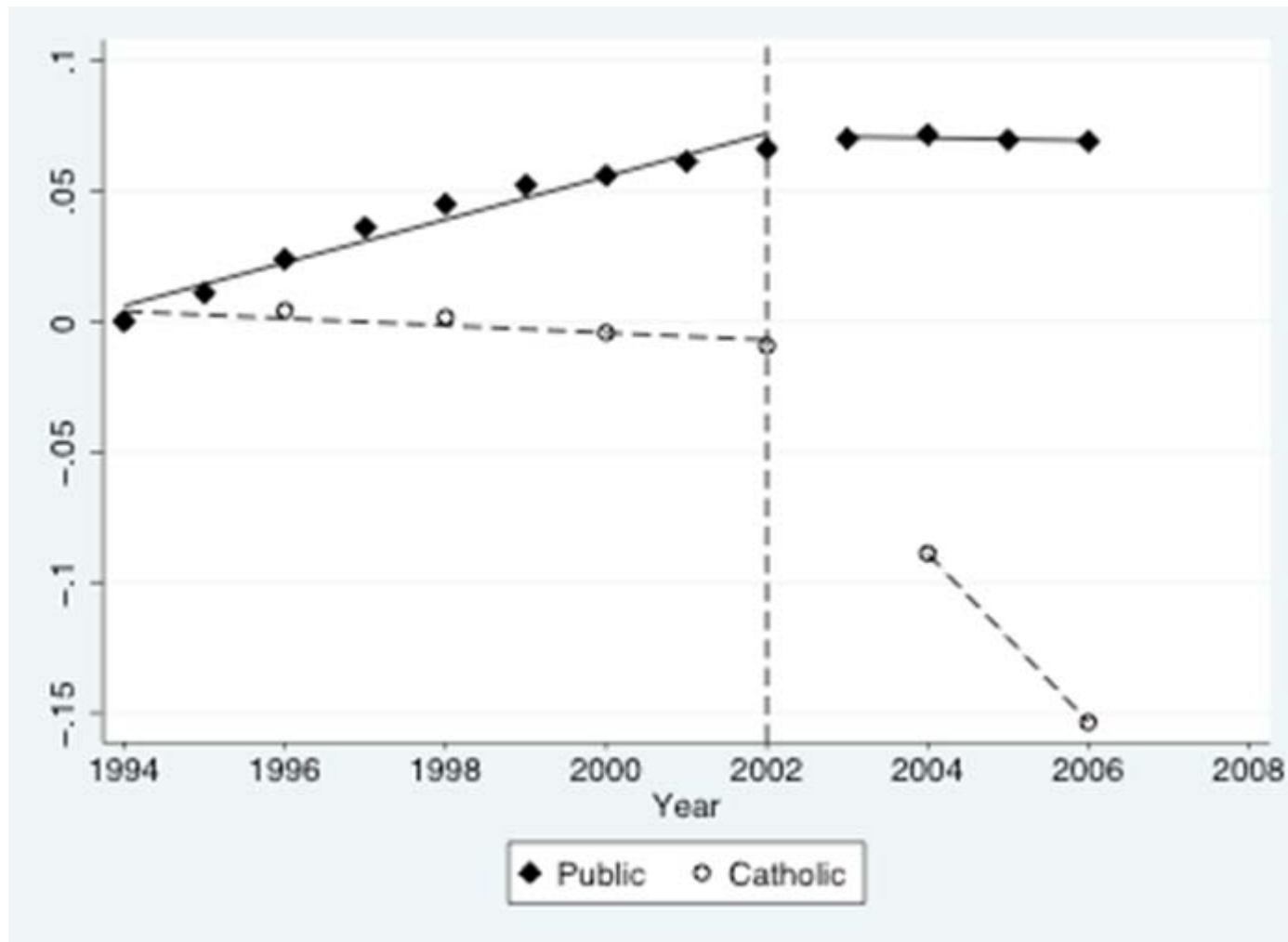
4th Grade Math Trends: Public vs. Catholic Schools



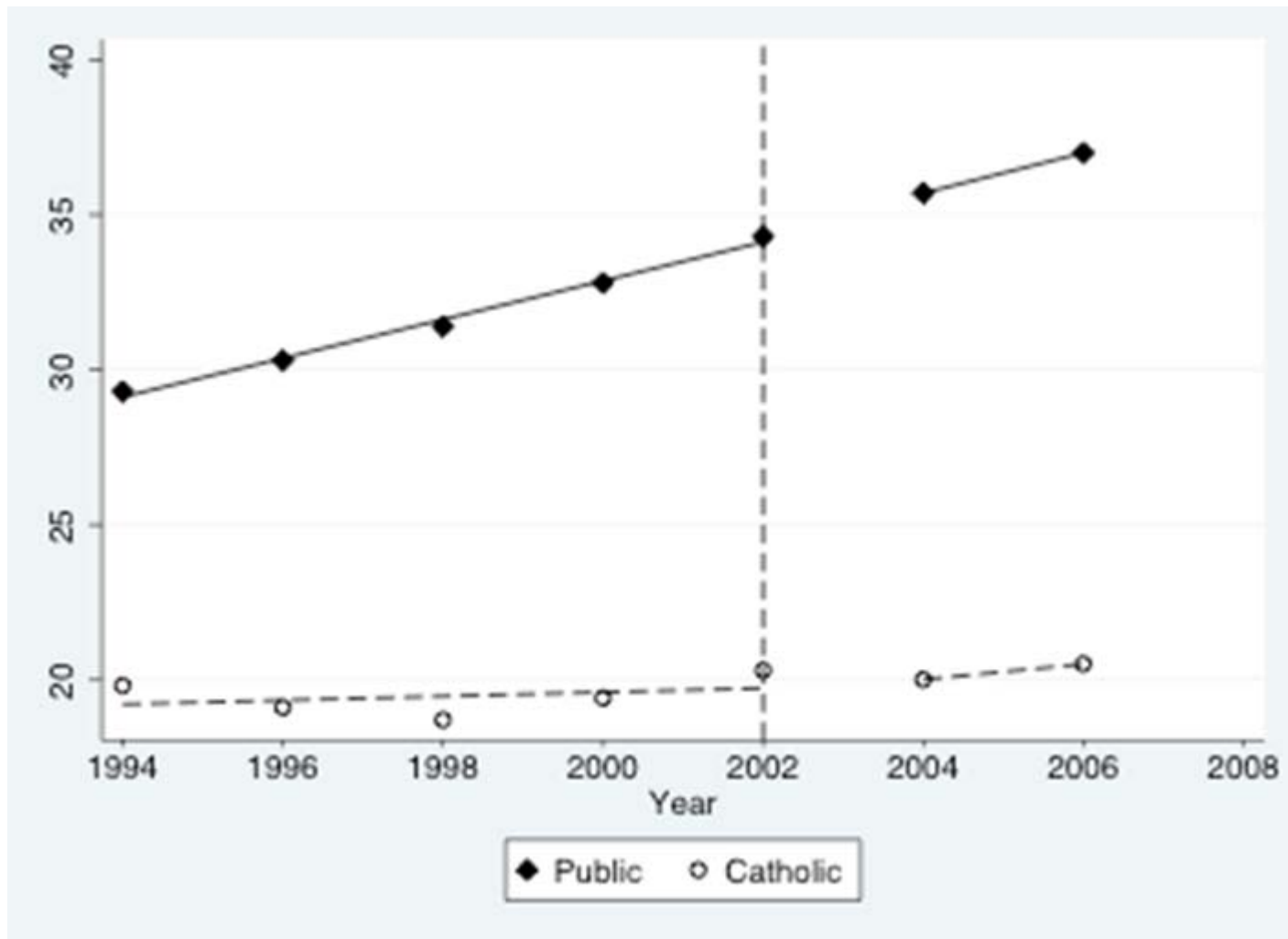
8th Grade Math Trends: Public vs. Catholic Schools



Enrollment Trends: Public vs. Catholic Schools



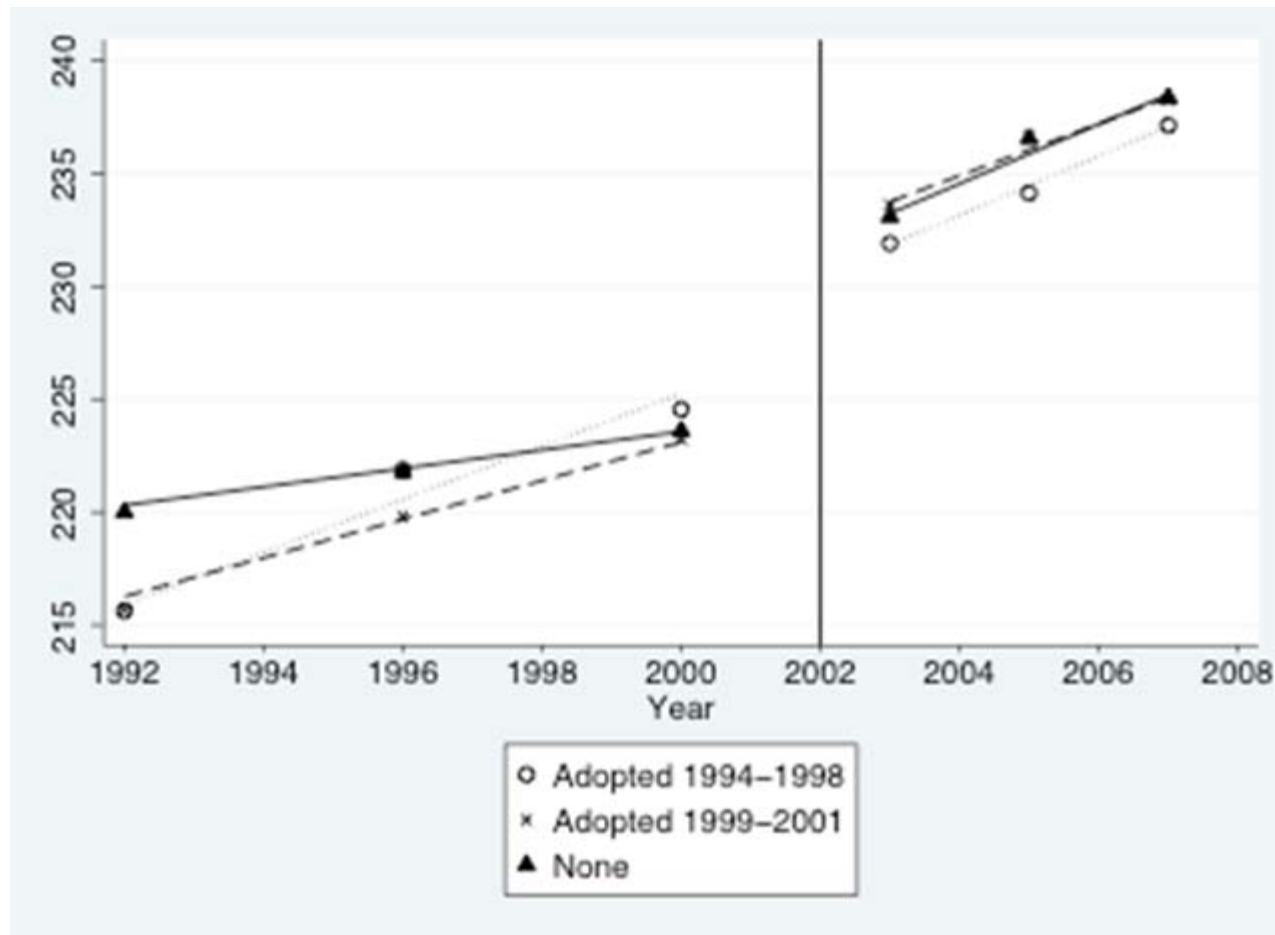
Trends in % Black or Hispanic: Public vs. Catholic Schools



Differential Achievement Trends by Pre-NCLB School Accountability

- NCLB may have weaker impact in states that had previously adopted similar policies
 - ▣ Important to account for across state differences in initial achievement and pre-NCLB achievement trends
- Assumes NCLB had same impact as pre-NCLB accountability
 - ▣ Only consider states with NCLB-like sanctions prior to 2002 (and not merely testing and/or report cards)

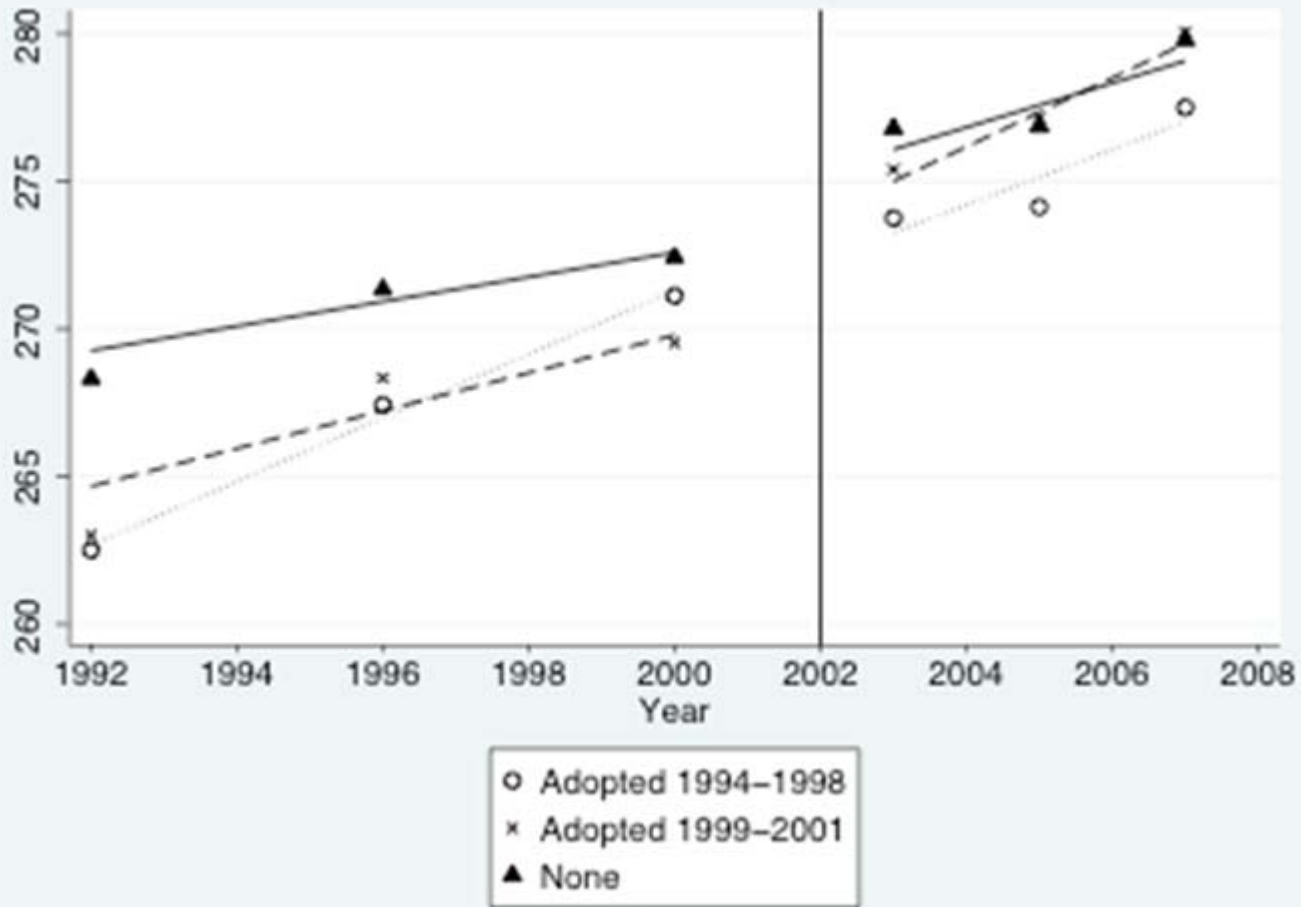
NAEP 4th Grade Math Trends by Pre-NCLB Accountability



NCLB Impact in 4th Grade Math

- Substantial positive effect overall
 - ▣ Mean: 6.8 scale score points (.22 effect size)**
- Effects largest at lower achievement levels, but significant even at higher achievement levels
 - ▣ % basic: 9.3pp (base = 64%, implies 14% increase)**
 - ▣ % proficient: 5.1 pp (base = 21% , implies 24% increase)**
- Effects positive for all major race groups
 - ▣ Larger effects for Black and Hispanic students
 - White: 6.8pp increase in % basic (base = 77%)
 - Black: 18.5pp increase in % basic (base = 33%)
 - Hispanic: 21 pp increase in % basic (base = 36%)
 - ▣ Caveats: weighting can make a difference; unweighted effects for Black students are large but not significant

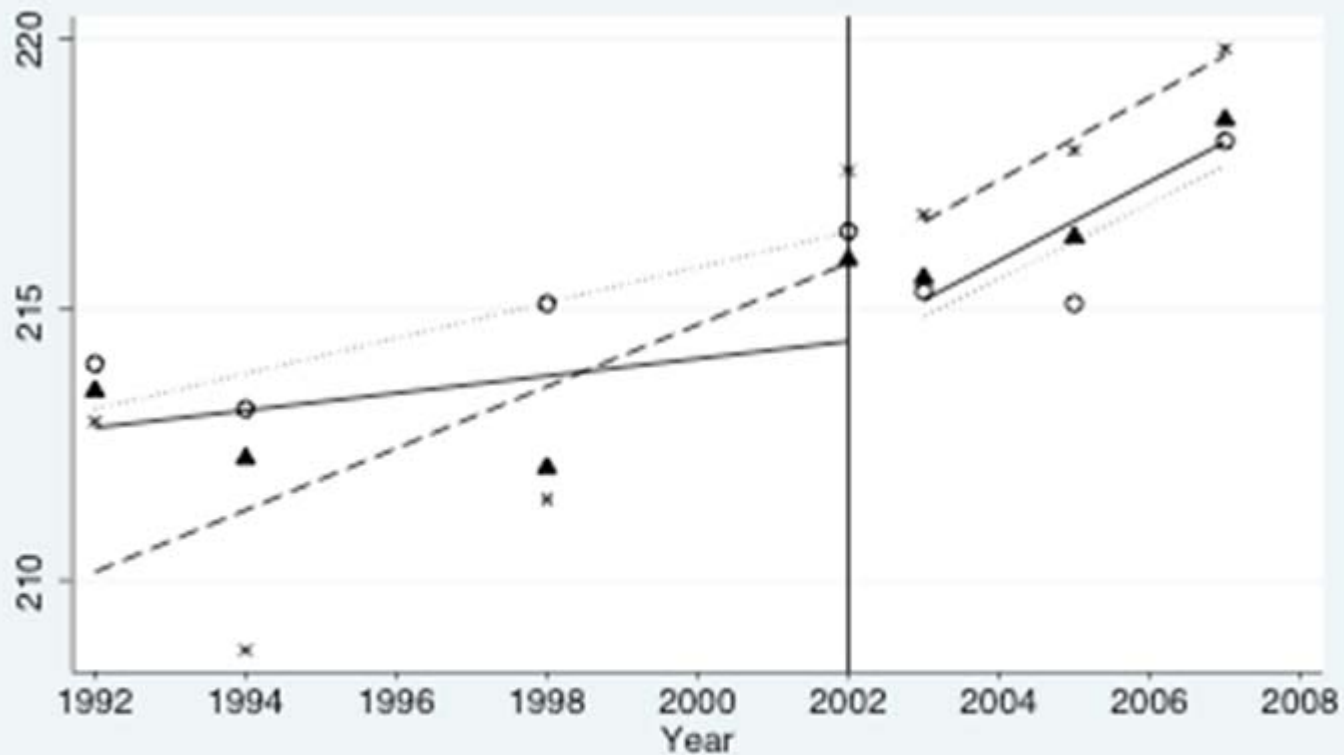
NAEP 8th Grade Math Trends by Pre-NCLB Accountability



NCLB Impact in 8th Grade Math

- Evidence of small positive effects overall
 - ▣ Mean: 3.4 scale score points (.09 effect size)
 - ▣ Driven by effects for lower-achieving students
 - % basic increased by 5.2pp (8%)**
- Considerable variation by race
 - ▣ No effects for white students
 - ▣ Large, positive effect on Hispanic students
 - 16pp increase in % basic (base = 36%)
 - ▣ Some evidence of positive effects for Black students, but estimates are very imprecise

NAEP 4th Grade Reading Trends by Pre-NCLB Accountability

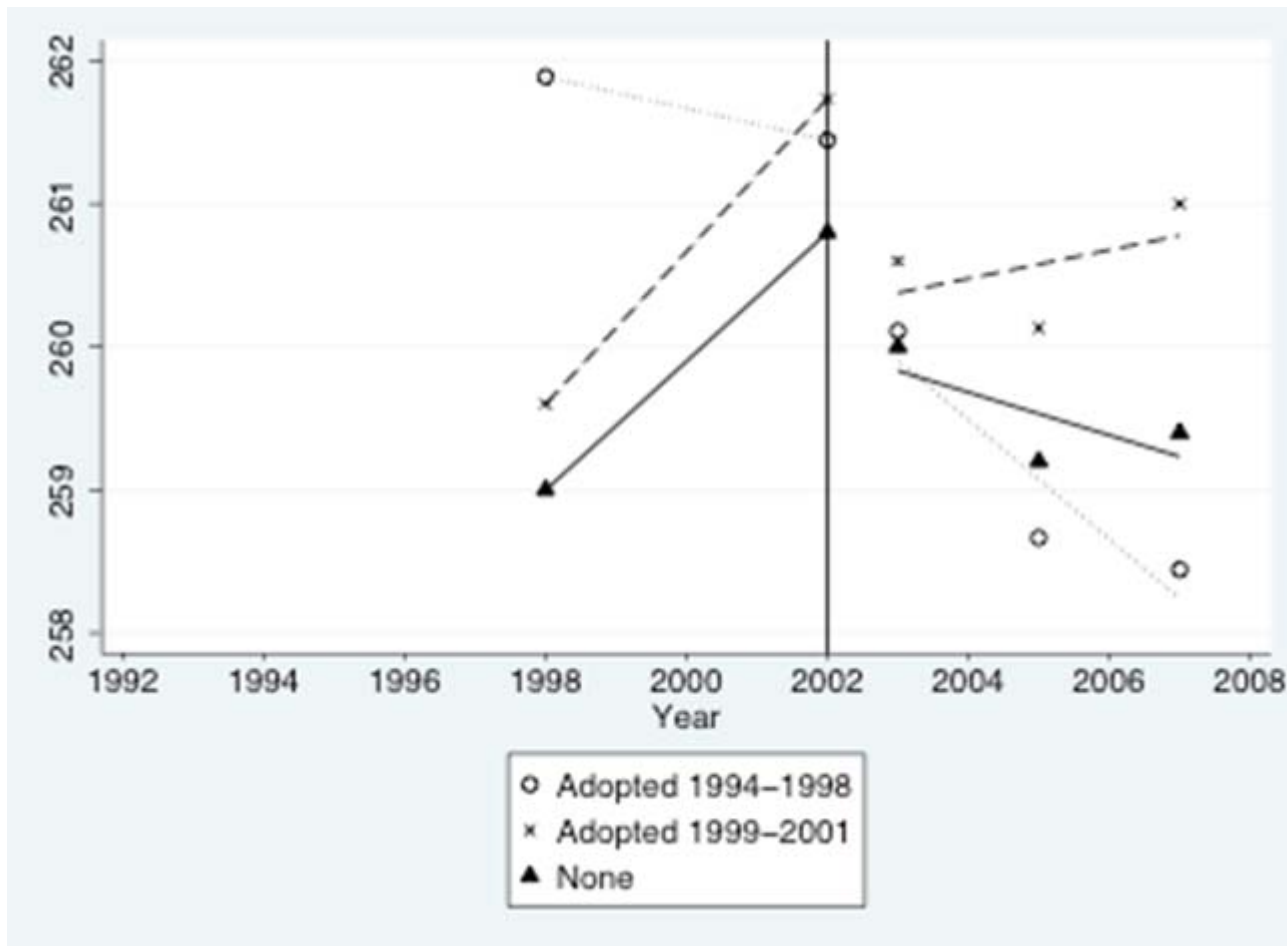


○ Adopted 1994-1998
× Adopted 1999-2001
▲ None

NCLB Impact in 4th Grade Reading

- Some evidence of small positive effects
 - ▣ Mean: 2.2 scale score points (.06 effect size)*
- Driven almost entirely by white students and students in the top percentiles (75th and 90th)
- Not robust to alternative specifications
 - ▣ Pre-NCLB achievement trend is not linear: big increase from 1998 to 2002
 - ▣ Analysis focusing on 1998 to 2007 => zero effect
- Our conclusion: no meaningful NCLB effect on 4th grade reading scores

NAEP 8th Grade Reading Trends by Pre-NCLB Accountability





NCLB Impacts in 8th Grade Reading

- No significant effect overall
- Some evidence of negative effect on Black students
 - ▣ Mean: -9.4 scale score points (.28 effect size)**
 - ▣ % Basic: -13.3 pp (base = 25%)**
- Potential compositional changes?
 - ▣ NCLB => 1-2pp increase in fraction of students who are Black in a state by 2007 (base = 13%)
 - ▣ Further research warranted

Potential Mechanisms

- NCLB associated with roughly 7% increase in per-pupil spending by 2007
 - ▣ Smaller effects for reading samples
- NCLB associated with small increase in fraction of students who are Black
 - ▣ 1-2 pp, off baseline of 13% => 8% increase

Conclusions

- NCLB appears to have generated some meaningful gains in important measures of student achievement
 - ▣ Particularly large and broad gains in 4th-grade math achievement
- However, this array of treatment effects falls short of NCLB's ambitious rhetoric
 - ▣ Little evidence for any positive effects in reading
- How to build on successes and learn from failures?
 - ▣ Understanding the sources of treatment heterogeneity by grade, subject, and race-ethnicity