Leaving No Child Behind: Two Paths to School Accountability

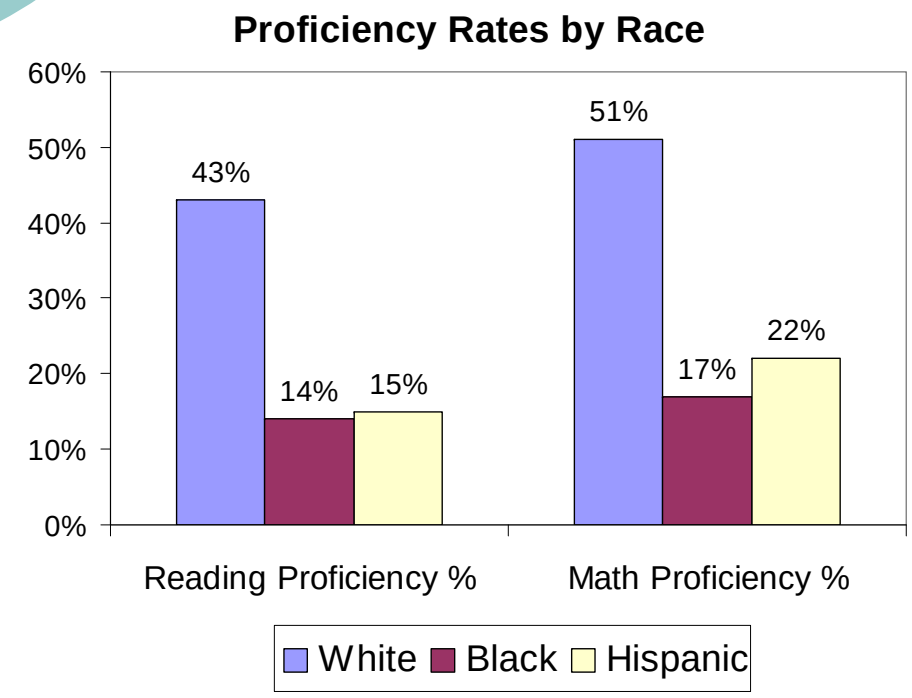
David N. Figlio, Northwestern University and NBER
Cecilia Elena Rouse, Princeton University and NBER
Analia Schlosser, Tel Aviv University

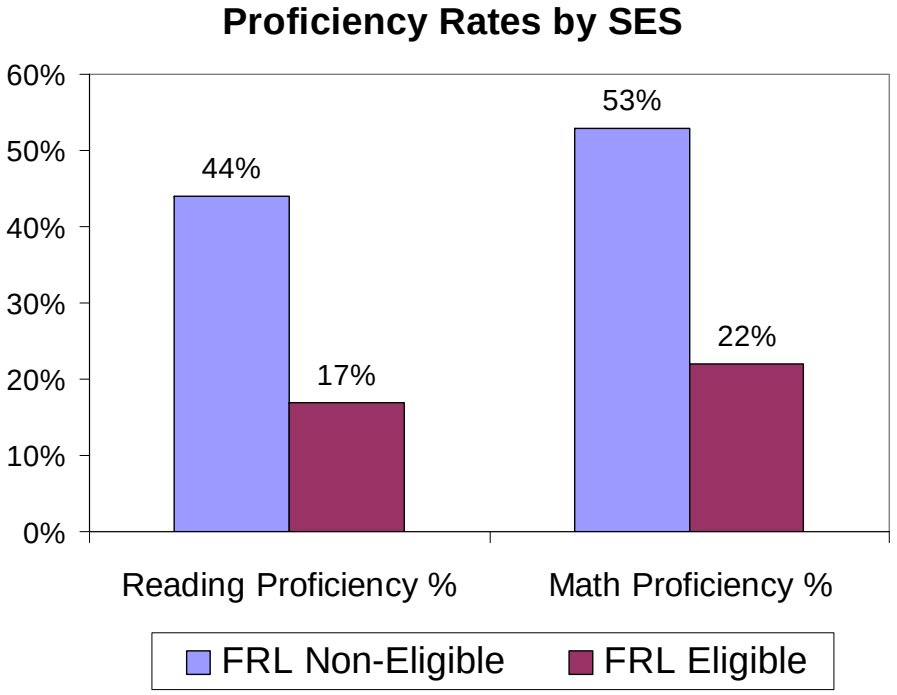
NATIONAL CENTER ON
Performance Incentives

CALDER

Introduction

- One of the top educational priorities: improving the achievement levels of the economically disadvantaged students and students of racial and ethnic minorities





Introduction

- One of the leading themes of the educational policy to close such economic, racial and ethnic gaps in student achievement is holding schools accountable for their performances.
  - Most recent example is the No Child Left Behind Act of 2001.

- Two important policy questions emerge:
  - Does the accountability pressure increase the average student achievement?
  - Which student subgroups are benefiting or losing from these systems?

Previous Literature

- Impact on average student achievement
  - Growing literature with mixed results
  - Some nationwide studies find significant improvements as a result of standards-based accountability (e.g. Carnoy and Loeb (2002), Hanushek and Raymond (2005)).
  - State-specific systems have been typically less encouraging (e.g. Koretz and Barron (1998), Clark (2003) and Haney (2000, 2002)).
- Which subgroups benefit/lose?
  - Relatively limited evidence
  - Neal and Schanzenbach (2007) and Krieg (forthcoming) reveal increased concentration on mediocre students in the expense of low and high-achievers as a result of proficiency-count-based accountability.
  - Chakrabarti (2006) indicate that schools focused on students performing just below the proficiency level in Florida.
  - Grissmer and Flanagan (1998) report narrowing achievement gaps possibly due to accountability systems.

Objectives

- Focus on the latter question:
  - What is the impact of Florida's accountability system (A+ Plan) and NCLB on two subgroups of interest?
    - Racial and ethnic minorities
    - Economically disadvantaged students
  - Results might help assess the effectiveness of accountability systems in narrowing the aforementioned achievement gaps.

Policy Background

- Florida's A+ Plan
  - In effect since 1999
  - Schools are assigned grades ("A", "B" etc.) based on the performance of their students on annual curriculum-based testing of all students in grades three through ten.
  - Introduced rewards or assistance and sanctions based on the grade of the school.
    - Assistance includes:
      - Recommendations on how to improve
      - Targeted funding for failing schools
      - Priority for a program that provides reading coaches trained in scientifically-based reading research
    - Sanctions include:
      - Students attending chronically failing schools were eligible for school vouchers, called "Opportunity Scholarships"

Policy Background

- Florida's A+ Plan
  - Between 1999 and 2001,
    - Aggregate test score levels
    - Test scores only in the grades with existing statewide curriculum-based assessments (4^{th},5^{th},8^{th}, and 10^{th} grades)
  - Starting in summer 2002,
    - Incorporate test score data from all grades from 3 through 10.
    - Year-to-year progress of individual students were also used to determine school grades.
  - New grading system places more emphasis on low-performing students
    - Yet, reduces the accountability pressure on schools performing at the highest levels.
    - 54% of the 'A' schools in 1999 received a grade of 'B' or lower in 2000 whereas only 12% of the 'A' schools in 2002 received 'B' or lower in the subsequent year.

Policy Background

- Florida's (Revised) A+ Plan
  - Might narrow the achievement gap in two ways:
  - The receipt of an 'F' grade significantly improves student achievement (Figlio and Rouse (2006), Rouse et al (2007), Chiang (2007), Chakrabarti (2006), and West and Peterson (2006)).
    - These subgroups are disproportionately represented in 'failing' or 'near-failing' schools.
    - Approximately 20% of black students, 10% of Hispanic students, yet only 3% of white students attend 'D' or 'F' schools.
  - The new grading system places more emphasis on previously low-performing students.

Policy Background

- NCLB
  - Several similarities to the A+ Plan:
    - Same grade levels in elementary school (3rd and higher) and same subjects (reading and math)
    - Evaluations based on same test: Florida Comprehensive Assessment Test (FCAT)
  - Important differences:
    - Focuses solely on the % of proficient students in a school
    - Schools must meet the proficiency requirements for each subgroup (e.g. economically disadvantaged, racial minority) with a sufficient number of students

Policy Background

- NCLB
  - Schools that meet these requirements are said to be making ‘adequate yearly progress’ (AYP)
  - Sanctions include:
    - Allowing students to transfer if the school fails to make AYP for two consecutive years.
    - Allowing parents to choose supplementary education services if the school fails to make AYP for three consecutive years.
  - However, the ‘take-up’ rates are very small:
    - Only 1% of eligible students transfer
    - Student demand for educational services leveled-off or declined after 2004 (Sunderman, 2007).

Policy Background

- NCLB
  - Expected to boost the achievement levels of minorities and the economically disadvantaged due to the specific subgroup requirements.
  - However, the high rates of failure in Florida to make AYP might induce ignorance among schools:
    - In the first year of designation, 75% of schools failed to make AYP in Florida.
    - In 2003, 55% of the ‘A’ schools and 87% of the ‘B’ schools failed to make AYP based on the federal standards.
  - Failing to make AYP has been shown not to improve student achievement in Florida (West and Peterson, 2006)

Data and Empirical Strategy

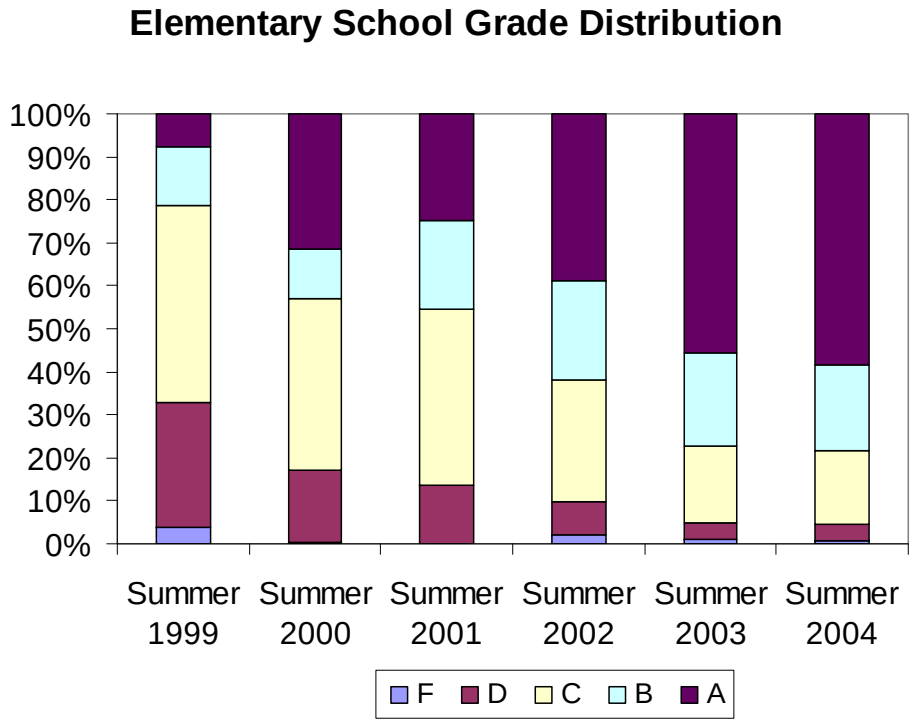
- Data
  - Longitudinal, administrative data on individual students in Florida between 1999-2000 and 2004-2005.
  - Contains FCAT scores (math and reading) and student characteristics (sex, gender, race, FRL eligibility etc.)
  - Focus on grades 3 through 5: 1,580,030 student-year observations for FRL eligible, 711,159 for black, 640,580 for Hispanic and 1,539,907 for white students.

Data and Empirical Strategy

- Empirical Strategy
  - For each racial and economic subgroup, estimate school fixed-effects models:
  - Outcome of interest: standardized, average reading and math scores
  - Variables of interest:
    - School grade in 2002 * POST: give the impact of the grading change on subgroup achievement (relative to receipt of ‘C’)
    - Identifiable subgroup in 2002 * POST: gives the impact of subgroup counting for AYP on subgroup achievement
  - Also control for student characteristics such as sex, gender, FRL status etc. and year dummies.

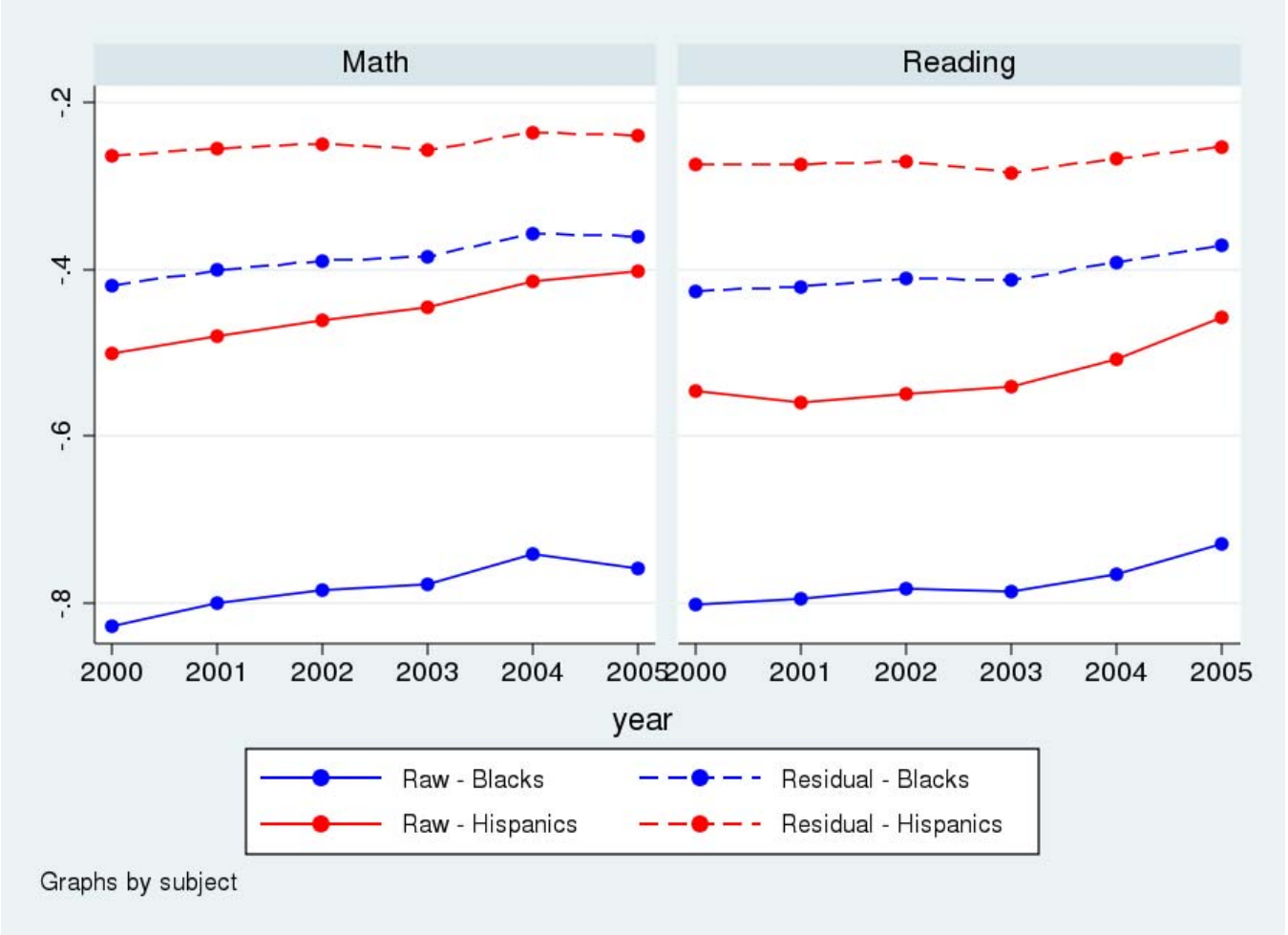
Data and Empirical Strategy

- Empirical Strategy
  - Why school grades in 2002?
  - The specifics of the new grading formula was not announced until the middle of 2001-2002 academic year
  - Left very little time for schools to adapt:
    - Change in grade distribution in 2002 was mainly due to system change (Rouse, 2007)
  - 2002 grades are more likely to reflect the 'true' hierarchy of school quality



Results:
Has the Achievement Gap Narrowed in Florida?

- Racial Achievement Gap in Florida: 2000-2005



Results:
Are the Accountability Systems Responsible?

- Regression Analysis – Impact of Grading Change

School grade in 2002	Subgroup			
	Black	Hispanic	White	Economically disadvantaged
A	-0.001 (0.012)	-0.001 (0.013)	0.014 (0.007)	-0.001 (0.009)
B	-0.006 (0.013)	0.009 (0.015)	0.001 (0.008)	-0.002 (0.010)
D	**0.034** **(0.018)**	**0.051** **(0.023)**	0.036 (0.023)	**0.044** **(0.015)**
F	**0.109** **(0.036)**	0.069 (0.047)	0.026 (0.077)	**0.101** **(0.032)**

Results:
Are the Accountability Systems Responsible?

- Regression Analysis – Impact of NCLB

	Subgroup			
	Black	Hispanic	White	Economically disadvantaged
Subgroup counted for AYP (effect starting in 2002-03)	0.010 (0.015)	**0.026** **(0.016)**	0.013 (0.020)	0.014 (0.048)
Subgroup counted for AYP (effect starting in 2003-04)	0.001 (0.015)	**0.029** **(0.016)**	0.005 (0.021)	0.020 (0.054)

Results:
Are the Accountability Systems Responsible?

- Regression Analysis – Impact of Grading Change or NCLB

School grade in 2002	Subgroup			
	Black	Hispanic	White	Economically disadvantaged
A	-0.010 (0.012)	-0.006 (0.013)	0.007 (0.008)	0.006 (0.009)
B	-0.011 (0.013)	-0.002 (0.014)	-0.003 (0.009)	-0.003 (0.010)
D	**0.036** **(0.018)**	**0.051** **(0.022)**	**0.040** **(0.022)**	**0.041** **(0.015)**
F	**0.116** **(0.036)**	**0.082** **(0.048)**	0.053 (0.080)	**0.112** **(0.033)**
Subgroup counted for AYP (effect starting in 2002-03)	0.001 (0.015)	**0.025** **(0.016)**	0.020 (0.020)	-0.016 (0.048)

Results:
Are the Accountability Systems Responsible?

- Regression Analysis – Combined Effect of Grading Change and NCLB for Schools with Different Grades

	Subgroup					
	Black students			Hispanic students		
School grade in 2002	Schools with measurable subgroup	Schools without measurable subgroup	p-value of difference	Schools with measurable subgroup	Schools without measurable subgroup	p-value of difference
"Safe" A – 430 points or higher	-0.026 (0.027)	-0.042 (0.029)	0.447	0.036 (0.027)	-0.007 (0.029)	**0.061**
"Marginal" A – 410-429 points	-0.043 (0.028)	-0.034 (0.036)	0.773	0.034 (0.028)	0.040 (0.039)	0.875
B	-0.034 (0.027)	-0.038 (0.034)	0.902	0.038 (0.027)	0.025 (0.035)	0.646
D	0.012 (0.030)	-0.002 (0.093)	0.877	0.091 (0.033)	0.083 (0.046)	0.864
F	0.092 (0.044)	n/a	n/a	0.115 (0.057)	0.156 (0.087)	0.680

Conclusions and Policy Implications

- Found evidence suggesting that the grading change in Florida has been effective:
  - Receiving ‘near-failing’ or ‘failing’ grades leads to improvements in subgroup achievement levels
- Subgroup requirements of NCLB has been less effective:
  - No improvement for black students
  - Improvement for Hispanic students – specifically at safe ‘A’ schools, which had arguably been released from accountability pressure following the change in the grading system in 2002

Conclusions and Policy Implications

- NCLB's subgroup requirements with rather ineffective sanctions are not likely to improve the achievement levels of disadvantaged students
  - Especially in states with already high standards of proficiency and heterogeneous schools
  - High rates of failure to make AYP might have contributed to ignorance among schools
- Holding schools responsible, especially, for the learning gains of low-performing students, combined with effective sanctions appears to lead to substantial gains in the progress of disadvantaged subgroups
  - Also suggest that it is possible to improve the achievement of disadvantaged students without singling out their performance.
- Florida's new grading system provides little incentive to boost student performance in highest-performing schools.
  - NCLB sub-grouping requirements may put pressure on these schools
  - May suggest that the ideal school accountability system would still put accountability pressure on high-performing schools to improve the performance of minorities and economically disadvantaged students.