NCLB and Teachers& NCLB and Social Studies and Sciences

Remarks

NATIONAL CENTER ON
Performance Incentives

CALDER

Overview

- A common theme of the three presentations is “the unintended consequences of NCLB”.
- Underlying policy issue: NCLB increases oversight, accountability, toward a narrowly defined set of tasks.
- Trade-offs are being made *all the time* so some shifting of effort away from “un-incented” tasks is inevitable.
- What we need to ascertain is: (1) what are the distortions? , (2) are students better off without the policy or with the policy?, and (3) are there cost effective way through which we can reduce these unintended consequences?

David P. Sims “Going down with the ship? The effect of school accountability on the distribution of teacher experience in California.”

- Looks at elementary and secondary schools in California from 2002-2006 to determine the effect of failing to meet academic performance thresholds on teacher experience under the NCLB accountability system.

- While potentially important, this may be a mitigating force rather than a driving effect since several studies find that accountability increases student achievement.

- Carnoy, M., Loeb, S. (2002). (using NAEP)
- Hanushek, Eric and Margaret Raymond (2005). (using NAEP)
- Cecilia Elena Rouse & Jane Hannaway & Dan Goldhaber & David Figlio, (2007). (RD design, effects on practices, Three years later the impacts persist.)
- Chiang, Hanley " unpublished (RD design, positive effect on same test)
- Jonah Rockoff and Lesley Turner (2009) (RD design, positive effect on same test)

Mechanisms

- North Carolina's accountability system had an adverse effect on schools serving low-performing students by making them even less able to retain teachers.
  - Charles T. Clotfelter , Helen F. Ladd , Jacob L. Vigdor , Roger Aliaga Diaz, (2004) "Do school accountability systems make it more difficult for low-performing schools to attract and retain high-quality teachers?" **Journal of Policy Analysis and Management 23(2), 251 - 271.**
- Are there "goodies" given to high performing schools that would serve as a compensating differential?
  - In North Carolina teachers in "exemplary" schools receive financial bonuses of $1500.
- Are these effects driven by increased teacher turnover, or changes in the characteristics of newly hired teachers? (one could pull this from the data).
- Are teachers responding to changes in the student composition (due to school choice)?
  - C. Kirabo Jackson. "Student Demographics, Teacher Sorting, and Teacher Quality: Evidence From the End of School Desegregation" *The Journal of Labor Economics* 27.2 (2009): 213-256.

Effect on Student Performance

- Emphasis on novice teachers is nice.
- How are other characteristics affected (selectivity of college, scores on license exams, fully certified, teaching in field)?
  - Clotfelter, Charles T., Helen F. Ladd, and Jacob L. Vigdor. 2007. How and why do teacher credentials matter for student achievement? Working Paper no. 12828.
- One can put all teacher quality into one index of predicted value-added based on some student achievement model (or use student data to look at estimated value-added).
  - Boyd, Lankford, Loeb, Rockoff, Wyckoff (2007)
  - Jackson and Bruegmann (2009)
- One can also look at turnover.
  - Turnover might have an independent effect on outcomes. Can this be quantified?

One could use the rule more

- "The fifteen percent rule only binds when there are at least 50 students in the subgroup, and the percent rule does not matter if there are 100 students in the subgroup."
- Your use the rule to identify failure is very nice. However, there are really two distinct sources of variation attributable to the rule.
- This gives you three distinct RD designs that can be illustrated visually–
  - (1) conditional on having 50 - 100 there is a sharp RD at 15%.
  - (2) conditional on having 15% there is a sharp RD at 50.
  - (3) conditional on <15% there is a sharp RD at 100.

Is there gaming of the subgroups?

- It would be helpful to show that there was no strategic classification of students to reduce the chances of failing.
  - Jacob, Brian A. "Accountability, Incentives And Behavior: The Impact Of High-Stakes Testing In The Chicago Public Schools," Journal of Public Economics, 2005, v89(5-6,Jun), 761-796.

Implications

- Failing schools may lose high quality teachers as a direct result of accountability.
- We need policy that mitigates this ill-effect.
  - Provide stronger support for teachers in failing schools.
    - Jonah Rockoff "Does Mentoring Reduce Turnover and Improve Skills of New Employees? Evidence from Teachers in New York City,"
  - Limit the extent of student flight.

Boyd, Lankford, Loeb, and Wyckoff (2006)

- **Asks whether the introduction of testing in the fourth grade increased the turnover and attributes of fourth-grade teachers relative to other grades.**
  - turnover *lower* for 4th grade post reform
  - especially true for "high-ability" teachers
  - less true for more experienced teachers
- **It would be interesting to see how the results relate to teacher value-added (since leavers are often weaker that those who remain).**
- **Do we think that this is a (positive) feedback effect? (I have heard of this anecdotally).**

Effect on one grade vs. on school

- There could be spillovers across grades.
  - principals move the good teachers to tested grades and encourage weaker teacher to move?
- The partial effect on a particular grade (when other are not tested) could be difference from the effect where all grades are tested.
- Possible Extension: To get at overall effect on teaching one could do a before/after cross-state analysis using the SASS or CCD data. [perhaps this was done in *Cuhna and Loeb (2007)*]

Policy implications

- **Effects on entire school are not obvious but findings are highly suggestive.**

- **Given that teaching is a field with little feedback, testing could actually make teaching more attractive (particularly for those who will do well).**
  - However, the effect when there are sanctions is less clear.
- **What test are used for will mediate their impact.**
  - To distribute awards or to shame teachers.
  - For professional development.
  - For school accountability.
  - Emphasis on raw scores vs. gains.
  - Compare to others (in school) or baseline scores.
  - Made public.

Boyd, Lankford, Loeb, Rockoff, Wyckoff "The Narrowing Gap in New York City Teacher Qualifications and Its Implications for Student Achievement in High-Poverty Schools"

- Has the Highly Qualified Teacher provision affected teacher quality?

- They find that measurable characteristics of teachers are more equal across schools in 2005 than they were in 2000. Schools with large proportions of poor students and students of color, on average, have teachers whose observable qualifications are much stronger than they were five years ago.

Numbers vs. Shares

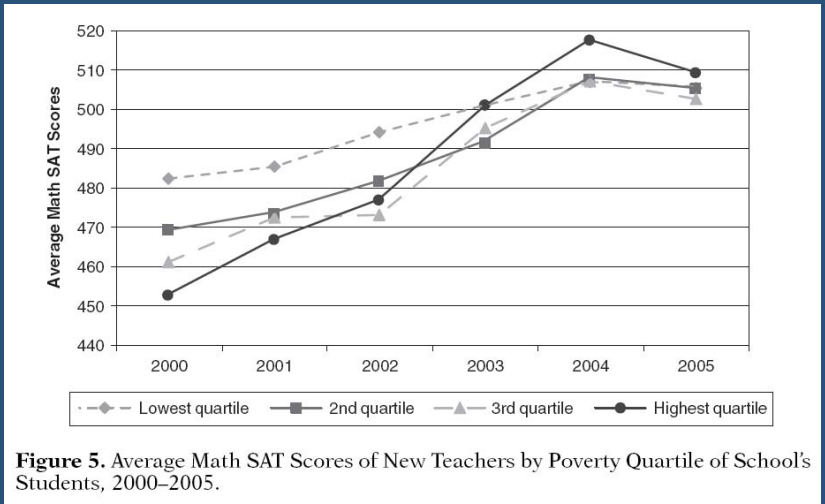
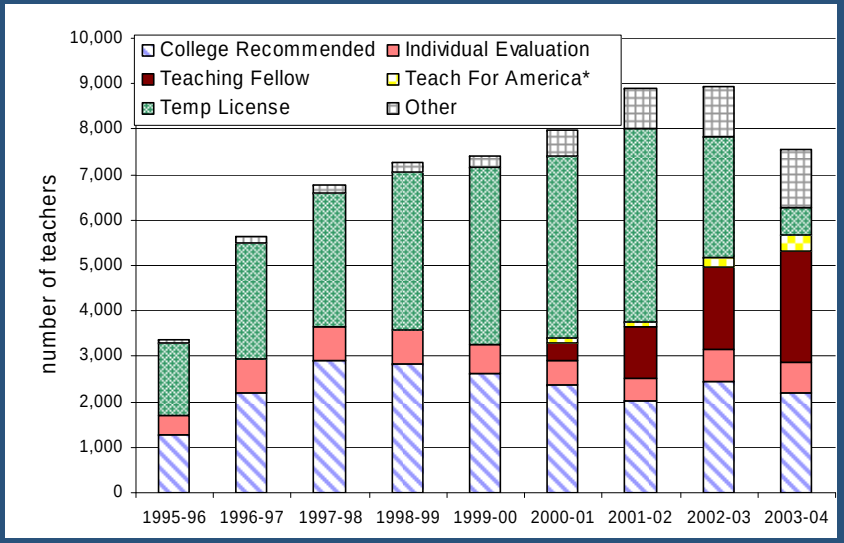


Figure 5. Average Math SAT Scores of New Teachers by Poverty Quartile of School's Students, 2000–2005.



- If the supply of strong teachers is low, the emphasis on stronger teachers could lead to a shortage.
- This is likely to be a shorter run concern (but a worry none the less).
- An analysis by numbers rather than shares might be helpful I (also look at turnover).

Possible distributional effects across schools.

- An unforeseen effect could be the cascading of teachers to affluent schools since all schools have increased demand for quality teachers (i.e. higher turnover and unfilled vacancies at inner city schools).
  - Similar to the effect of class size reductions in CA.
    - Bohrnstedt, George W., and Brian M. Stecher. (1999)
- How would the impact analysis change if one were to take unfilled positions and turnover into account?

Policy implications

- “average improved achievement for students in the poorest decile of schools of 0.03 standard deviations”
  - Comparing similar measures to value-added measures imply the real effect on teacher quality could be as much as three times that size.
- Requirements on teacher quality can be effective at screening poor teachers – particularly at difficult to staff schools.
- However, one should consider the supply of teachers.
- Success will depend on ability to observe teacher quality.
  - Much research needed on professional development.
  - More research is needed to determine how to *create* a high quality teacher.

Dale Ballou and Jeffrey Springer “HAS NCLB SET BACK ACHIEVEMENT IN SCIENCE AND SOCIAL STUDIES?”

- This is a question asked by many with little empirical evidence guidance.

- This is test of the classic multitask principal agent model.
  - Holmstrom, Bengt, and Paul Milgrom. 1991. “Multitask Principal-Agent Analysis: Incentive Contracts, Asset Ownership and Job Design” *Journal of Law, Economics and Organization* 7 (Special Issue): 24–52.

Summary

- **Compare schools that have a strong incentive to change instructional emphases in response to NCLB, to schools that have much weaker incentive to do so.**
  - High incentive schools include those at high risk of failing to make AYP (based on pre-NCLB achievement in math, reading) and Title I schools (since non-Title I schools are not sanctioned under NCLB).

- **In both states they see an increase in science and social studies achievement between 2003 and 2007 in both "high-incentive" and "low-incentive" schools.**

- **In general, increases are even greater in the high-incentive schools.**

Possible Mechanisms

- Schools as risk of failure undergo broad across the board changes that affect all subjects.
  - Cecilia Elena Rouse & Jane Hannaway & Dan Goldhaber & David Figlio, 2007. "Feeling the Florida Heat? How Low-Performing Schools Respond to Voucher and Accountability Pressure," NBER Working Papers 13681.
- There are complementarities between math and reading and science and social studies.
- The scope for substitutions across subjects is limited.
- There are some ceiling effects.

Possible Specification

- With no pre-treatment data, the exogenous variation comes from years of exposure to NCLB.
- For each subject, one could look at “years of exposure” interacted with “high incentive” with school, year, and grade fixed effects.

Do we know what the first order effect is?

- What is the effect of being “high incentive” pre and post NCLB on Math and ELA overall?

- What is the effect of being “high incentive” pre and post NCLB on Science and Social Studies overall?

- From a policy perspective we might also care about the overall effect (based on a model that does not condition on math and ELA).

Possible Extensions:

- Is there any RD type design that can be implemented?
  - Use RD in Title 1 eligibility (at 40% low income)
  - Use RD in earning an F vs. a D

- This would assuage concerns about endogeneity, difference across dissimilar schools, and ceiling/plateau effects in comparing high-achieving schools to low-achieving schools. (different plateau reflect different proportions of tough to reach students)

Long Run vs. Short Run

- Most ways in which schools could substitute from Sci/SS towards Math/ELA would not operate in the short run.
  - i.e. purchasing better textbooks in Math and leaving older textbooks for Science. Focusing recruiting on hiring strong teachers in math and science, and neglecting soc/sci.
- Over the short run, if there is no change in personnel and little change in the distribution of school resources there may be little or no effect, even if there would be one in the long-run.

Big Picture

- **The focus on Math and English may not have detrimental effects on Science and Social studies.**
  - This may be a short run effect.
  - With the caveat that social studies and science are used in state accountability.
- **For an overall effect we should look at broader outcomes like school dropout and college enrolment?**
  - If there are improvements in math and reading, with ill effects in SS/Sci the overall effect could be positive (the end focus should be on achievement overall).