ARIZONA STATE UNIVERSITY

EPRU | EDUCATION POLICY RESEARCH UNIT

Colorado
University of Colorado at Boulder

EPIC Education and the Public Interest Center

DOCUMENT REVIEWED:	"High-Achieving Students in the Era of NCLB"
AUTHORS:	Tom Loveless, Steve Farkas, & Ann Duffett
PUBLISHER/THINK TANK:	Thomas B. Fordham Institute
DOCUMENT RELEASE DATE:	June 18, 2008
REVIEW DATE:	July 29, 2008
REVIEWER:	Gregory Camilli
E-MAIL ADDRESS:	camilli@rutgers.edu
PHONE NUMBER:	(732) 932-7496

SUGGESTED CITATION:
Camilli, G. (2008). *Review of "High-Achieving Students in the Era of NCLB"* Boulder and Tempe: Education and the Public Interest Center & Education Policy Research Unit. Retrieved [date] from http://epicpolicy.org/thinktank/review-high-achieving-students.

Summary of Review

A recent report from the Fordham Institute considers potential instructional policies for high-achieving students that should be considered in the forthcoming reauthorization of the No Child Left Behind Act. The report finds: 1) achievement growth among high-achieving students has been slower than that of low-achieving students; 2) this trend can be traced to state accountability practices; and 3) teachers would support new policies targeted to high achievers. This review examines several premises of the report's conclusions, both implicit and explicit. It concludes that evidence regarding the effects of accountability is inconsistent. It also concludes that teachers have a more nuanced view of allocating resources to high- and low-achievers than is recognized in the report.

I. INTRODUCTION

The Thomas B. Fordham Institute's recent report *High-Achieving Students in the Era of NCLB*[1] concerns the development of high-achieving students in the policy context of the No Child Left Behind Act (NCLB), which is the current incarnation of the Elementary and Secondary Education Act (ESEA). The report is composed of two strands of analysis. In Part 1 of the report, Tom Loveless explores the link between the effects of the NCLB law and growth of high- and low-achieving students in mathematics and reading achievement as measured by the National Assessment of Educational Progress (NAEP). In Part 2, authored by Steve Farkas and Ann Duffett, the results of a survey are presented in which teachers were asked to reflect on how schools are serving high-achieving students while meeting the requirements of NCLB.

Loveless shows that high-achieving students have made relatively less (or "languid, " p. 35) progress since 1990 than low-achieving students, and that the narrowing of the gap between high and low achievers is associated with increases in state accountability practices dating from about 2000. Because this suggests that incentives provided by federal legislation work, an experiment is proposed for discovering how schools can be motivated to better serve high achievers. High-achieving students—including Black and Hispanic high achievers—are also described in some detail in this section.

Farkas and Duffett find that teachers indicate at-risk students are their top priority, but that teachers also believe that all students deserve their teacher's attention and that sufficient resources are not available to high achievers. They note that teachers think differentiated instruction is difficult and that teachers support homogenous grouping to enable advanced students to learn faster and in greater depth.

These two independent sets of findings are stitched together in an executive summary and in a foreword by Chester Finn and Michael Petrilli. These two sections provide the majority of the policy recommendations as well as valuable insights into the policy goals of the Fordham Institute, which sponsored the report. Indeed, it is in the foreword where readers find the key link from the report's findings to the potential redesign of ESEA. But these sections also reveal a glaring methodological inconsistency. For instance, the executive summary maintains that

> Neither of these studies sought a causal link between [NCLB] and the performance of high-achieving students. We cannot say that NCLB "caused" the performance of the nation's top students to stagnate any more than it "caused" the achievement of our lowest-performing students to rise dramatically. All we know is that the acceleration in achievement gains by low-performing students is associated with the introduction of NCLB (and, earlier, with state accountability systems) (p.6).

Yet on the very next page of text, the foreword maintains that

> The major finding of this dual study is that, in one respect at least, [NCLB] is working precisely as designed … Congress was quite clear about NCLB's objectives. Right on the cover, it's termed "An Act to close the achievement gap (p. 8)."

Though correlation is not causation, a com-

parison of these two passages indicates that the line is sometimes blurry. In fact, cause-effect implications occur throughout the report, and are sometimes used with little if any qualification to tie empirical results to policy recommendations.

After the findings and conclusions of the report are summarized, results and interpretations in these two sections will be examined. As argued below, there is substantial doubt about the generalizability of gap results presented in Part 1, and a more balanced model of social justice is evidenced by teachers than recognized in Part 2. Note that Parts 1 and 2 are independent studies, and for this reason they are discussed separately below, except in the final section of this review.

II. Findings And Conclusions Of The Report

For the purpose of this review, the major topics of the report are organized into four areas: purpose, student achievement, teacher perceptions, and policy implications.

Purpose

Two purposes for conducting the study are described in the foreword. One is the concern for equity: all students should have the opportunity to excel. A second motivation is economic:

> If we want to compete in the global economy, don't we need all our young people—including our highest achievers—to make steady progress too? And if so, isn't our current approach to standards-based reform in need of a make-over? (p. 8)

This economic concern is echoed in Part 1 with a quote (p. 14) by Susan Goodkin:[2]

> By forcing schools to focus their time and funding almost entirely on bringing low proficiency students up to proficiency, NCLB sacrifices the education of the gifted students who will become our future biomedical researchers, computer engineers, and other scientific leaders.

These rationales are consistent with the theory that investing in education, especially science and mathematics education, will help to keep America economically competitive in the world (see also the reference to Sputnik on p. 14). In our era of scientific-based reasoning, this contention is often presented as self-evident. But it is at best simplistic and, at worst, an umbrella for a number of inexplicit values. It is, therefore, a topic worth brief exploration.[3]

Economic development has many influences. For example, the progress of stem cell research is influenced by presidential opposition; software coding is outsourced to countries with lower cost structures; and scientific leadership in alternative energy sources is often diverted by special interests—not to mention international trade, monetary policies, stock market and housing bubbles, or healthcare and Social Security deficits. Though international competitiveness is a frequent rationale for ratcheting up educational achievement, there is a lack of consensus regarding the macro effects of education as well as whether the United States is losing ground in science and technology. For example, according to a report from the Rand National Defense Research Institute, the U.S. has recently grown "faster in many measures of [science and technology] capability than did Japan and Europe" (p. xv), and the U.S. "produces 35 percent, 49 percent, and 63 percent, respectively, of total world publications, citations, and

highly cited publications, employs 70 percent of the world's Nobel Prize winners and 66 percent of its most-cited individuals" (p xvi).[4] There may be an urgent problem, but the problem needs to be more thoroughly described in order to be solved. In parallel, the purposes—and especially the economic objectives—of policies oriented to high achievers need to be more explicit.

Beyond concerns for equity and economic strength, a key thread in the report is the reauthorization of ESEA. The executive summary contends that a just education system is one in which all students are challenged, and a redesign of the current standards-based approach according to this principle would lead to an education law "dramatically different" (p. 6) from NCLB. Accordingly, Finn and Petrilli argue that we need to bring some "honesty" (p. 12) into the debate about how to value high achievers relative to other students.

Achievement

There were two stated reasons in Part 1 for examining the growth of high and low achievers. The first was to check how high achievers have fared since the advent of NCLB in reading and mathematics, using data from the National Assessment of Student Progress (NAEP) at grades 4 and 8. Also using NAEP data, the second was to examine the size of the gap between students at the 10th and 90th percentiles; according to the goals of NCLB, the gap should narrow.

At the 10th percentile, gains were observed for all combinations of subject area, grade level, and achievement level, except 8th grade reading. In contrast, minimal achievement growth was observed at the 90th percentile. The 90th-10th percentile (or high-low, for short) gap also narrowed for all combinations, except 8th grade reading. This was due to a relatively larger (termed "impressive" on p. 18) gain for low achievers; in contrast, the long-term trend for higher achievers appears unaffected by NCLB. No support was found for the "Robin Hood" hypothesis, namely, that NCLB's purported benefits for lower-achieving students are attained at a cost to higher-achieving students.

Loveless cautions that it is debatable as to when the sanctions of NCLB became effective. Thus, changes in test scores or gaps that date from 2000-2002 are treated as inconclusive evidence of the effects of NCLB. However, he also notes that "general" accountability (p. 23) might account for the 2000-2003 changes in score gaps. Using the state accountability scoring system of Carnoy and Loeb,[5] and examining the years from 1996-2002, the high-low gap is shown to have narrowed more for high-accountability states, except in 8th grade reading.

A detailed statistical description of high achievers is also given in Part 1, and this information should be of substantial interest to policy makers. In comparison to lower achievers, for instance, higher achievers tend to have myriad socioeconomic advantages, to attend suburban schools, to take more advanced math classes, and to have teachers who majored or minored in math. High achievers who are Black, Hispanic, or from low-income families attend schools that on average have more constraints than their White counterparts, resulting in fewer opportunities to learn. According to Loveless, NCLB may place such students "at risk of suffering any lost opportunities stemming from NCLB's incentives" (p. 36).

Teacher perceptions

As noted above, a key item (highlighted in

the foreword) on the teacher survey concerns equity:

> For the public schools to help the U.S. live up to its ideals of justice and equality, do you think it's more important that they (A) Focus on raising the achievement of disadvantaged students who are struggling academically OR (B) Focus equally on all students, regardless of their backgrounds or achievement levels?

About 86% of teachers agreed with option B. Yet only 23% reported that the needs of advanced students were a top priority at their schools. In response to another key question, 76% of teachers overall would like to see the nation "relying more on homogeneous classes for advanced students so that they learn faster and in greater depth" (p. 68). It is significant to note that Finn and Petrilli (p. 12) consider the answer (86% agreement) to the statement that all students should have the opportunity to excel to be "an overwhelming repudiation of one of NCLB's core tenets."

Examples of other teacher perspectives included the following:

1. Teachers are not overly concerned that focusing on advanced students will harm the self-esteem of other students.
2. Teachers are "less worried" about skewing advanced classes by race.

The questionnaire was also designed to explore instructional strategies for high achievers, including:

3. Most teachers do not report receiving training on teaching advanced students, and they favor more professional development for this purpose.
4. Teachers report that differentiated instruction is difficult.
5. Grade acceleration is rare, and should be reconsidered.
6. Teachers favor magnet programs that bring advanced students together.

In sum, one important empirical finding is that teachers appear to agree, on average, with strategies that require social and even racial stratification.

IV. The Report's Use of Research Literature

A cursory literature review is provided for the achievement analysis in Part 1. The three studies that are discussed are provided to show mixed evidence of achievement growth rather than to provide a balanced survey of the research. No literature is provided for the survey component given in Part 2, nor are results from other surveys discussed.

It would have been helpful if the report had acknowledged the national accountability research that preceded it[6]—though the frequently cited study by Carnoy and Loeb[7] was footnoted. There is indeed some common support for the report's claim that accountability policies have positively affected some measures of student achievement, but three points should be kept in mind.

First, in a more recent study, researchers Carnoy and Loeb tracked students in all 50 states in a more comprehensive manner. They concluded that mathematics improvements from the late 1990s "tapered off from 2000 to 2003."[8]

Second, most results about policy effects on mathematics achievement have been based on NAEP total scores, rather than content strands (number properties and operations;

measurement; geometry; data analysis and probability; algebra).[9] If accountability has effects on the skills of low-achieving students, those skills are more likely to be basic. The same is true with reading. There is no reason to believe that because accountability systems may have had an effect in the past with respect to basic skills, they will be effective in the future with higher-order skills.

Third, the estimated effects of accountability models on score gaps for the most part have been small,[10] and few researchers have sought to explore simultaneous influences on the size of achievement gaps, e.g., the effects of pre-school education.

V. Review of the Report's Methods

The methods used were simple but direct and useful, as discussed below. Additional issues related to methodology are discussed in the next section on validity of conclusions.

Achievement

In Part 1, the NAEP Data Explorer tool was used to track mathematics and reading scores at the 90th and 10th percentiles. No attempt was made to control for moderator variables (e.g., teacher qualifications, state education spending, and median family income), which could have provided alternative explanations to policy effect. But, to be fair, few studies have taken this multidimensional approach to estimating policy effects.

Teacher Perceptions

Part 2 of this report is based on questionnaires sent to a nationally representative sample of teachers from grades 3-12. From a database of current teachers in the U.S. (containing their school addresses), a sample of 6,000 was generated. A total of 900 questionnaires were returned, representing a 15% response rate. In addition, quotations from qualitative data generated by five focus groups were used to illustrate key points.

VI. Review of the Validity of the Findings And Conclusions

Findings in the report regarding student achievement are largely based on the analysis of NAEP data sets. The primary validity issues examined are whether causal inferences were indeed drawn, despite the caveats, and whether these inferences were adequately supported by the data. For the teacher survey, validity issues concern the sampling procedures as well as the wording of questions.

Achievement

Recall that the report used the Carnoy-Loeb scores to establish an association between accountability and the achievement gap. But the reader is cautioned against drawing causal inferences from these associations:

> This analysis cannot test causal theories relating to NCLB (or anything else) since the NAEP data are cross-sectional, offering a snapshot of how students are performing at a single point in time. (p. 16)

Despite this warning, a causal trap is nonetheless set:

> If the larger gains at the bottom of the achievement distribution are associated with the incentives of accountability systems, this trend suggests a missed opportunity to promote achievement among high achievers. (p.25)

The policy recommendation then follows

> Accountability systems work about as intended. The key is to get the incentives right. To promote the continued progress of high-achieving students, policymakers should consider creating incentives for schools to boost more students into the upper echelons of achievement. (p. 36)

This, apparently, is the source of the statement in the foreword to the effect that NCLB is working "precisely" as intended.

Correlation does not imply causation, but, as the logic appears to flow, the correlation in question suggests a causal effect that is reliable enough to be used for policy recommendations. The problem here is that effects can have multiple causes, and no effort is made in Part 1 to isolate the unique contribution of accountability (of any kind). To formulate evidence-based policies, however, the same degree of rigor is required as one would use in estimating a causal effect in a randomized experiment.

If achievement growth at the 10th percentile received a post-NCLB boost relative to the 90th percentile (i.e., the high-low gap narrowed), then NCLB-like strategies might possibly be used to leverage outcomes for high achievers. However, a counter argument can be made that the gap did not shrink in a way consistent with general accountability influences. There are in fact two NAEP data sets from which gap statistics can be calculated. While the NAEP state data were used in Part 1, an examination of the high-low gap is also possible using the NAEP long-term trend data set. Unfortunately, gap estimates from these two sources are not equivalent, and this limits the generalizability of the gap trends presented in Part 1.

For example, consider 4th NAEP mathematics. Using the NAEP state data, the gap shrinks from 2000 to 2005 by 8 points, an effect size of about .25. For the long-term NAEP data, however, the 1999-2004 gap at age 9 (about 4th grade) decreases by only 3 points, an effect size of less than .10. Moreover, consider the high-low gap trend from the long-term data in more detail:

Table 1: Gap between low- and high-achievers in NAEP mathematics data (age 9)

Year	10th %	90th %	Gap
1978	171	264	93
1982	173	263	90
1986	177	264	87
1990	186	271	85
1992	185	271	86
1994	187	272	85
1996	187	274	87
1999	187	275	88
2004	197	282	85

Here, there is no evidence of a shrinking gap at age 9 in mathematics, except, perhaps, from 1978-1986, well before NCLB. There has been more growth at the 10th than at the 90th percentile since 1978, but most of this differential growth occurred prior to 1990. For other age-subject combinations in the long-term data, evidence of a shrinking gap consistent with accountability influences is similarly weak. In sum, the state and long-term data sets give different results, and this problem needs to be addressed before gap statistics can be used confidently to describe the effects of accountability policies.

This statistical ambiguity in estimating gaps may arise in the ethereal realms of psychometrics. Determining the scale score that corresponds to a particular percentile rank depends fundamentally on technical measurement procedures. Yet to my knowledge,

an NAEP technical manual has not been issued since 1998. Thus, there is no way to determine whether the disparity in gap trends is due to test specifications, sampling procedures, scaling, equating, or the like. This is not to say that the achievement results given in the report are inaccurate, but rather that a discrepancy exists that presents a significant obstacle to a confident interpretation of NAEP gaps and trends in gaps.

Teacher Perceptions

The 15% response rate for the teacher survey is a limitation, and more descriptive information would have helped readers to evaluate potential response biases. Breaking down responses by urbanicity or level (e.g., elementary, middle, high school) would have also been useful, especially the latter, because grouping practices tend to be different among these levels.[11]

In a problematic questionnaire item, 57% of teachers agreed that "Sometimes, when teachers use advanced students to tutor other students, it's because they have run out of ways to challenge the high achievers." I wonder how a question like this comes into being. One could agree with the statement that "Sometimes, I eat lunch when I'm not hungry," but might tend to disagree with the statement "I eat lunch when I'm not hungry." This type of wording can exacerbate response bias.

I would add that while the report generally does not characterize NCLB as having a "Robin Hood" effect on gifted students, the analysis of survey results does make that leap: "[Teachers] have seen more attention paid to struggling students because of the schools' drive to move more 'bubble' students to proficiency" (p. 69).

It is significant to note that Finn and Petrilli (p. 12) consider the answer to the equity item above (86% agreement with equal focus on all children) to be "an overwhelming repudiation of one of NCLB's core tenets." However, immediately following the equity item teachers were asked the following:

> If you had to pick, what should be a greater priority for the nation's schools: (A) Maximizing the achievement of academically advanced students OR (B) Closing the achievement gap?

Only 26% of teachers agreed with option A, while 57% chose option B. This result suggests that if we listened to teachers, their opinions would reveal a far more nuanced understanding of the problems and issues of high achievers than is evident in the foreword.

Chain of Reasoning

The chain of reasoning in the report seems to go something like: (a) the intent of NCLB was to focus education dollars on low-achieving kids, (b) there is a pattern of NAEP results that is not inconsistent with the hypothesis that accountability has had small positive effects on low achievers, (c) teachers would support more attention to the needs of high achievers, (d) high-achieving Black, Hispanic, and poor students are especially placed at risk by NCLB, and (e) therefore to remain economically competitive, we need to bring more focus to high achievers by taking the following two steps: (1) having web-based courses in advanced subjects like algebra for students who do not otherwise have access to advanced courses, and (2) tracking students into more homogenous ability groups for instruction. To be sure, tracking isn't explicitly recommended. However, Loveless has recently authored a work favorable to tracking practices,[12] and

the Fordham Institute, as acknowledged in the foreword, will carry out additional research in this area.

To help ensure that public schools that accept federal funding attend to the needs of both low and high achievers, Finn and Petrilli suggest that incentives might be useful. These could be implemented as "growth models," "multiple indicators," or "multiple measures" (p. 12). The rationale for this recommendation is that NCLB, state accountability, or both shrank the high-low achiever gap. Thus, some type of accountability mechanism could be used to equalize instructional opportunities for all children (i.e., low-achieving students, "the best and brightest," and everyone in between). However, as noted above, the evidence underlying this proposal is weak. In Part 1, a more modest approach is taken in proposing an experiment to investigate the efficacy of potential incentives to schools for addressing the needs of high achievers.

More generally, the push in this report for more stratification of educational opportunities runs contrary to a large body of empirical research. There is much research available, ranging from kindergarten to college students, on the effects of selective education. Studies on promotion to first grade show conclusively that children, even at-risk children, benefit from promotion—a kind of detracking.[13] The research on tracking similarly supports heterogeneous classrooms, though some consider this evidence to be inconclusive.[14] Recent research in higher education suggests that students attending more selective colleges do not have an advantage in graduation rates or long-term outcomes over similar students at less selective schools.[15] The same is true for law schools.[16] In general, the evidence is stacked against instructional approaches that physically separate students into enduring groups. This is a topic worth revisiting in the future designs of ESEA, provided the evidence is weighed honestly.

VII. USEFULNESS OF THE REPORT FOR GUIDANCE OF POLICY AND PRACTICE

There are two components of assessing how academically talented students fare in American schools. The first is informational and concerns identifying inequalities in educational opportunity. For example, survey results could be used to determine the prevalence of schools for which AP courses are not yet available for which there are sufficient numbers of students who would take those classes. A similar strategy could be taken to determine the prevalence of and demand for teachers qualified to teach advanced courses. Run the numbers, then tailor solutions to mismatches when they are found. Moreover, in designing effectiveness studies of instructional practices for high achievers, it would be most useful to examine the factors that drive between-school outcomes. Federal interventions in education are notorious for targeting the same accountability model to schools with very different challenges. Save for the experiment proposed in Part 1 of the report, it is all too easy to infer that another one-size-fits-all solution is in the making.

The second component of assessing the circumstances of high achievers would involve using the research literature on the effects of instructional policies to its fullest potential. The two studies collected in this report attempt to inform broad policies on the strength of two fairly narrow analyses. A broader synthesis of studies on instruction would be useful prior to formulating and recommending instructional mandates. Likewise, many serious research studies have tackled the issue of effectiveness of NCLB,

and an empirical synthesis of these studies would also be useful. The current report would make a contribution to this synthesis, but does not stand on its own as a document for informing instructional policies for high achievers. I applaud the spotlight placed on Black, Hispanic, and poor high achievers in Part 1; it was long overdue, and it provides a meaningful statistical description of these student populations. In fact, NCLB may have contributed little to the achievement growth of any high-achieving student.

As for the redesign of ESEA, there is currently substantial appreciation of the distinction between policy enactment and the funding of federal education mandates. Yet even with a firm grasp of incentives for developing high achievers, there would be a concern that adding more criteria to the current set of 41 NCLB indicators might further stretch instructional resources. Policy makers who implement evidence-based practices for high-achieving students will thus need to grapple directly with the issue of resources and the drawbacks of increased bureaucracy. In schools with underserved high achievers, increasing opportunity requires access to qualified teachers—with or without web-based instruction. In schools with high achievers who are thriving, mandates may interfere with successful instructional practices. It is important to get the incentives right for developing high achievers, but it is more important to get the vision of equitable instruction right in highly diverse student populations.

Notes and References

[1] Loveless, T., Farkas, S., and Duffett, A. (2008). *High Achieving Students in the Era of NCLB.* Washington, DC: The Thomas B. Fordham Institute.

[2] Goodkin, S. (2005, December 27). Leave no gifted child behind. Washington Post, A25. Goodkin is Executive Director of the California Learning Strategies Center.

[3] The report's contention about economic competition is different from the claim that a college education provides a rate of return to individuals, which has substantial empirical support. For example, see

Rouse, C. (1999). Further estimates of the economic return to schooling from a new sample of twins. *Economics of Education Review*, *18*, 149-157.

[4] Golama, T. & Hosek, J. (2006). U.S. Competitiveness in Science and Technology. Santa Monica, CA: Rand National Defense Research Institute.

[5] Carnoy, M. & Loeb, S. (2003). Does external accountability affect student outcomes? A cross-state analysis. *Education Evaluation and Policy Analysis*, 24, 305-331.

[6] Amrein, A. L., & Berliner, D. C. (2002, March). High-stakes testing, uncertainty, and student learning. *Education Policy Analysis Archives*, 10. Retrieved April 20, 2004, from http://epaa.asu.edu/epaa/v10n18/.

Nichols, S. N., Glass, G. V., & Berliner, D. C. (2005). *High-Stakes Testing and Student Achievement: Problems for the No Child Left Behind Act*. Education Policy Studies Laboratory: University of Texas, San Antonio and Arizona State University.

Braun, H. (2004, January). Reconsidering the impact of high-stakes testing. *Education Policy Analysis Archives,* 12. Retrieved April 20, 2004, from http://epaa.asu.edu/epaa/v12n1/.

[7] Carnoy, M. & Loeb, S. (2003). Does external accountability affect student outcomes? A cross-state analysis. *Education Evaluation and Policy Analysis*, 24, 305-331.

[8] A brief description is given in Viadero, D. (2005, May 11). Second look at tougher accountability yields new results. *Education Week*, Bethesda, MD: Editorial Projects in Education Inc.

[9] The 2000 NAEP mathematics framework is given at http://www.nagb.org/pubs/96-2000math/toc.html, and 2005 framework at http://www.nagb.org/pubs/m_framework_05/toc.html.

[10] E.g., Nichols, S. N., Glass, G. V., & Berliner, D. C. (2005). *High-Stakes Testing and Student Achievement: Problems for the No Child Left Behind Act*. Education Policy Studies Laboratory: University of Texas, San Antonio and Arizona State University.

Braun, H. (2004, January). Reconsidering the impact of high-stakes testing. *Education Policy Analysis Archives*, 12. Retrieved April 20, 2004, from http://epaa.asu.edu/epaa/v12n1/

[11] Loveless, T. (1999). *The tracking wars: State reform meets school policy*. Washington, DC: Brookings Institute Press.

[12]. Loveless, T. (1999). *The tracking wars: State reform meets school policy*. Washington, DC: Brookings Institute Press. (Tom Loveless is the author of Part 1 of the report under review.)

[13] Hong, G. & Raudenbush, S.W. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis*, 27, 205-224.

[14] For example, see:

Gamoran, A., & Berends, M. (1987). The effects of stratification in secondary schools. *Review of Educational Research*, 57, 415-435.

Burris, C. C., Wiley, E. W., Welner, K. G. & Murphy, J. (2008). Accountability, rigor, and detracking: Achievement effects of embracing a challenging curriculum as a universal good for all students. *Teachers College Record*. *110*(3), pp. 571-608.

Oakes, J. (2005) *Keeping Track: How Schools Structure Inequality (Second Edition).* New Haven: Yale University Press.

[15] Alon, S. & Tienda, M. (2005). Assessing the "mismatch" hypothesis: Differences in college graduation rates by institutional selectivity. *Sociology of Education*, 78, 294-315.

Dale, S.B. & Krueger, A.B. (2002). Estimating the payoff to attending a more selective college: an application of selection on observables and unobservables. *Quarterly Journal of Economics*, 117, 1491-1527.

[16] Rothstein, J. & Yoon, A. (2007, October). Mismatch in Law School (updated). Princeton University Education Research Section Working Paper #16.