

# THE BENEFITS OF FLORIDA'S TEST-BASED PROMOTION SYSTEM

Marcus A. Winters  
*Senior Fellow, Manhattan Institute*



State and municipal policymakers are increasingly addressing the practice of social promotion in schools—moving children along to the next grade whether or not they have mastered the curriculum—by mandating test-based grade promotion.

This paper draws conclusions about the effects of a policy limiting social promotion. To do so, it employs a methodology known as regression discontinuity, which is capable of producing causal estimates of policy effects to study the impact of Florida’s test-based promotion policy on later student achievement. Under this program, students must take an exam to automatically pass from third to fourth grade (some students scoring below the automatic promotion threshold may still advance at teacher discretion). Students who are retained in third grade also receive a rigorous remediation regime aimed at improving their long-term performance.

By studying the long-term performance of children who just barely passed the test, as well as those who were just barely left behind, it was possible to compare two essentially identical populations: one set of students who moved forward despite only borderline understanding of the material; and another set who stayed behind a year and received tutoring, mentoring, and other remedial interventions.

On average, the students who were remediated did better academically, in both the short and long term, than those who were promoted. Tellingly, the benefits of the remediation were still apparent and substantial through the seventh grade (which is as far as the data can be tracked at this point).

These results contrast with previous work cited by supporters of social promotion finding that grade retention has strong negative consequences for the student’s later academic outcomes. This paper takes the view that there is considerable reason to question the validity of much of that research because most prior studies on grade retention use methods that are flawed or inadequate. Notably, these studies do not take into account “unobserved differences” between students studied. Unobserved differences are characteristics, such as maturity level or home environment, that aren’t accounted for in the researchers’ datasets, but which may have an enormous bearing on student performance.

The results of this study demonstrate that a test-based promotion policy structured similar to Florida’s policy should be expected to improve student performance relative to a policy of social promotion. Florida’s system is an example for policy makers across the country to emulate.

## ABOUT THE AUTHOR

---

Marcus A. Winters is a senior fellow at the Manhattan Institute and an assistant professor at the University of Colorado Colorado Springs. He conducts research and writes extensively on education policy, including topics such as school choice, high school graduation rates, accountability, and special education. Winters has performed several studies on a variety of education policy issues including high-stakes testing, performance-pay for teachers, and the effects of vouchers on the public school system. His research has been published in the journals *Educational Evaluation and Policy Analysis*, *Education Finance and Policy*, *Economics of Education Review*, *Teachers College Record*, and *Education Next*. His op-ed articles have appeared in numerous newspapers and magazines, including *The Wall Street Journal*, *The Washington Post*, *USA Today*, the *New York Post*, the *New York Daily News*, the *Weekly Standard*, and *National Affairs*. He is often quoted in the media on education issues. Winters received his B.A. in political science from Ohio University in 2002, a Ph.D. in economics from the University of Arkansas in 2008.

# CONTENTS

---

1	<b>Introduction</b>
2	<b>Background</b>
2	<b>The Limitations of Previous Research on Grade Retention</b>
4	<b>Florida’s Test-Based Promotion Policy</b>
4	<b>The Regression Discontinuity Approach</b>
	Figure 1. Illustration of Regression Discontinuity
6	<b>Results</b>
	Table 1. Results—Reading
	Table 2. Results—Math
	Figure 2. Fourth Grade Reading
	Figure 3. Fifth Grade Reading
	Figure 4. Sixth Grade Reading
	Figure 5. Seventh Grade Reading
7	<b>Results in Context</b>
	Figure 6. Treatment Effects of Educational Interventions
9	<b>Conclusion</b>
10	<b>Endnotes</b>
11	<b>References</b>



# THE BENEFITS OF FLORIDA'S TEST-BASED PROMOTION SYSTEM

Marcus A. Winters

## INTRODUCTION

The practice of social promotion in schools—promoting children to the next grade level even if they have not mastered the curriculum—has become a flashpoint in recent years. Increasingly, legislators and policymakers are scrutinizing social promotion policies and are indeed passing laws and regulations aimed at limiting or prohibiting the practice.

Though most modern policies aimed at addressing social promotion incorporate several treatments meant to remediate low-performing students, such policies are particularly controversial because they substantially increase grade retention. Opponents of such policies point to a wide body of research that seemingly shows that retention harms later student outcomes. Although there is a great deal of research on the topic, very little of it is of high enough quality to be a useful guide for policymakers.

This paper discusses the results of recent research that measured the effect of remediation under Florida's test-based promotion policy on student achievement. Using test-score data over several years, the

author examined the progress of third-grade students by comparing groups of children who had been barely promoted with those who had just missed the cutoff and were remediated. The study provides strong evidence that remediation under Florida's policy has a substantial positive effect on student performance. The data show a very large short-run effect on student achievement that fades somewhat over time. However, several years after the intervention—the data follow students up to the seventh grade—the positive effect of remediation remains distinguishable and quite large.

The research described in this paper, conducted with University of Arkansas professor Jay P. Greene, has been peer-reviewed and is scheduled for publication in the summer issue of the journal *Education Finance and Policy*. Readers interested in technical details and specification checks should look to that academic work. The purpose of this paper is to interpret the results of the academic article to better inform the ongoing policy debate.

## BACKGROUND

Seeking to end the practice of social promotion, several school systems have recently enacted, or are considering adopting, test-based promotion policies that attempt to augment teacher discretion with objective measurement. These remediation policies often include several treatments for low-performing students, including assignment to summer school. What makes these policies particularly controversial is that they require students to demonstrate possession of some minimal skill in order to avoid grade retention. Such policies have been in effect for several years in Florida, New York City, and Chicago; they were recently adopted by state legislatures in Oklahoma, Arizona, and Indiana; other states, including Colorado, Iowa, New Mexico, and Tennessee, are reported to be considering similar legislation.

Opponents of social promotion have argued that, while retention might hurt a student's feelings (at least in the short run), the school is doing the child no favors by promoting him to a higher grade level for which he lacks the proficiency to succeed. Proponents of

ending social promotion point to the third grade as a particularly important gateway because it is commonly said that after this point, students stop "learning to read and begin reading to learn."

But proponents of social promotion argue that keeping children back a year harms students more than it helps them. The intuition behind their argument largely centers on student feelings: social promotion is the default for many public school systems largely because educators are concerned that holding a student back a grade will harm his self-esteem, leading to negative effects on later learning and life outcomes.

At first blush, the evidence appears to be against the use of grade retention. Those opposed to test-based promotion policies point to a wide body of research purporting to find that retention harms later student achievement. However, the reliability of much of that earlier research has been called into serious question. In particular, the research techniques utilized are, in most cases, inadequate: most previous research has failed to adequately account for the many differences between retained and promoted students that are not listed in the data. Research that does not account for so-called unobserved differences between retained and promoted students cannot credibly estimate the effect of retention on later student outcomes.

## THE LIMITATIONS OF PREVIOUS RESEARCH ON GRADE RETENTION

School systems adopting policies that dramatically increase grade retention do so despite a large body of research finding that it is harmful to student achievement.<sup>1</sup> Opponents of grade retention frequently point to this research as proof that retaining students will inevitably harm later student outcomes. The ferocity with which this research is wielded is often eyebrow-raising. Commenting on the recent expansion of such policies in several school systems across the nation, Arizona State University professor David Berliner recently said, "It seems like legislators are absolutely ignorant of the research, and the research is amazingly consistent that holding kids back is detrimental."<sup>2</sup>

But is the research really that clear? There is ample reason to question the validity of findings for much of the existing body of research on retention. Most, though not all, previous studies on the effect of retention on student outcomes provide misleading conclusions because they fail to adequately account for all the differences between retained and promoted students.

A valid study must measure the effect of an intervention by comparing the outcomes of a group of individuals who were exposed to it (the treatment group) with that of another group of individuals who were not (control group). The purpose of the control group is to represent what would have happened to the treated group had the intervention not occurred.

The most important problem facing any empirical researcher is ensuring that the differences between the outcomes of the treatment and control groups can be entirely attributed to the intervention in question. One must reach a very high bar to make such “causal” interpretations of research findings. There must be substantial reason to believe that members of the treatment and control groups in a study are identical in every way—both observed in the researcher’s data set (for instance, race/ethnicity) and unobserved in that data set (did his parents read to him as a child?)—except for their access to the treatment.

Many previous studies on grade retention have simply compared the later outcomes of retained students with those of promoted students from their class, holding constant some observed characteristics about them, such as their race/ethnicity and socioeconomic status. The problem with this approach is that when retention is determined by the teacher rather than by some administrative rule, there is ample reason to believe that there are differences between the promoted and retained students that are observed by the teacher but invisible to the researcher.

For instance, a teacher might look at two students with identical test scores at the end of the year but determine that one of the students has the maturity level to be promoted while the other is immature and thus should be retained. Researchers can’t account for

a characteristic like the student’s maturity level because it does not appear in their data set. Nonetheless, the maturity level of students is very likely to be related to students’ academic achievement in later years. Thus, when the researcher observes that the promoted student outperforms the retained student on later standardized tests, it is unclear whether the difference can be attributed to grade retention or if it is just an artifact of the personality differences between the two students.

Random assignment to treatment is the strongest available research technique because it ensures that the treatment and control groups are identical at baseline. These types of “gold standard” studies are often used in medical trials and have recently become more common in the social sciences as well. However, random assignment is not always feasible. It would be beneficial for researchers; but schools, for obvious reasons, do not randomly assign students to be promoted or retained.

Fortunately, researchers have developed several techniques capable of closely replicating random assignment studies. However, most of the studies cited by opponents of test-based promotion policies do not use the updated methods. Of the 22 papers evaluating the effect of grade retention on achievement published between 1990 and 2006 that were identified in a recent meta-analysis, only six could be defined as “high quality,” meaning that they included comparison groups with similar observed characteristics at baseline and adequate statistical controls.<sup>3</sup> The meta-analysis discovered that studies of higher quality report more positive effects from grade retention than do lower-quality evaluations. Nonetheless, even these higher-quality papers do not tend to find that retention leads to substantial academic improvements.

There are limitations even among the very few papers deemed “high quality.” Most important, several of these papers utilize statistical matching techniques based solely on observed characteristics of retained and promoted students. Thus even some of the most sophisticated papers do not account for unobserved differences between retained and promoted students.

One of the few notable exceptions in the previous research is a series of studies that utilizes a regression discontinuity strategy to study the effect of remediation under test-based promotion policies on short-run outcomes in Florida and Chicago. These papers deserve particular attention because—in contrast even to very sophisticated matching strategies—under minimal assumptions, regression discontinuity accounts for both observed *and unobserved* differences between retained and promoted students. That is, this research strategy closely replicates the results of a randomized experiment. Regression discontinuity is one of the few research strategies strong enough for the U.S. Department of Education’s What Works Clearinghouse to consider capable of making causal estimates.

Results are mixed from studies using a regression discontinuity design to study the effect of Chicago’s policy on the achievement of third- and sixth-grade students.<sup>4</sup> My own previous research, also coauthored with Jay Greene, used a regression discontinuity design and found that third-grade students remediated under Florida’s test-based promotion policy benefited one and two years after the retention decision.<sup>5</sup>

This paper is an extension of my earlier work evaluating the effect of Florida’s policy. The primary contribution of this paper is to follow student achievement over a much longer period of time to discover whether the effect of remediation under Florida’s policy fades as students progress through school. I am also able to measure the effect of the remediation treatment on multiple cohorts of entering third-grade students.

## FLORIDA’S TEST-BASED PROMOTION POLICY

Florida’s test-based promotion policy was among a series of education reforms adopted under the governorship of Jeb Bush, who was first elected in 1998. Students who entered the third grade in the fall of 2002 were the first subjected to the mandate. The law has applied to all subsequent cohorts of third-grade students in the state.

Florida’s policy requires that third-grade students score at or above Level 2 (the second-lowest of five levels) on the Florida Comprehensive Assessment Test (FCAT) in order to be promoted by default. However, scoring below the benchmark does not guarantee that the student will be held back: students can receive one of several exemptions and be promoted despite their low performance. In fact, nearly half of the students with test scores below the threshold in the policy’s first year were promoted.<sup>6</sup>

Under Florida law, students retained according to the policy are also subjected to additional interventions during the retained year; they are required to attend a summer reading camp and are assured of being assigned to a “high-quality teacher” during the following school year. Schools must also develop academic improvement plans for each remediated student that addresses the student’s specific needs during the retention year.

Because the other interventions are also triggered by the policy, we are not able to determine the effect of retention alone on student achievement separate from these other interventions. That is, the results should be thought of as an average effect of the entire remediation treatment on student achievement and not just of retention itself. However, in the technical version of this paper, we show that assignment to a “high-quality” teacher during the retained year does not appear to be driving any of the reported results.<sup>7</sup>

## THE REGRESSION DISCONTINUITY APPROACH

The analyses underlying the results discussed in this paper utilize a regression discontinuity design. Readers interested in the technical details of our approach should look to the academic work on which the discussion in this paper is based. However, the intuition behind the regression discontinuity procedure is relatively simple to understand.

The procedure takes advantage of the fact that students’ likely exposure to Florida’s remediation policy

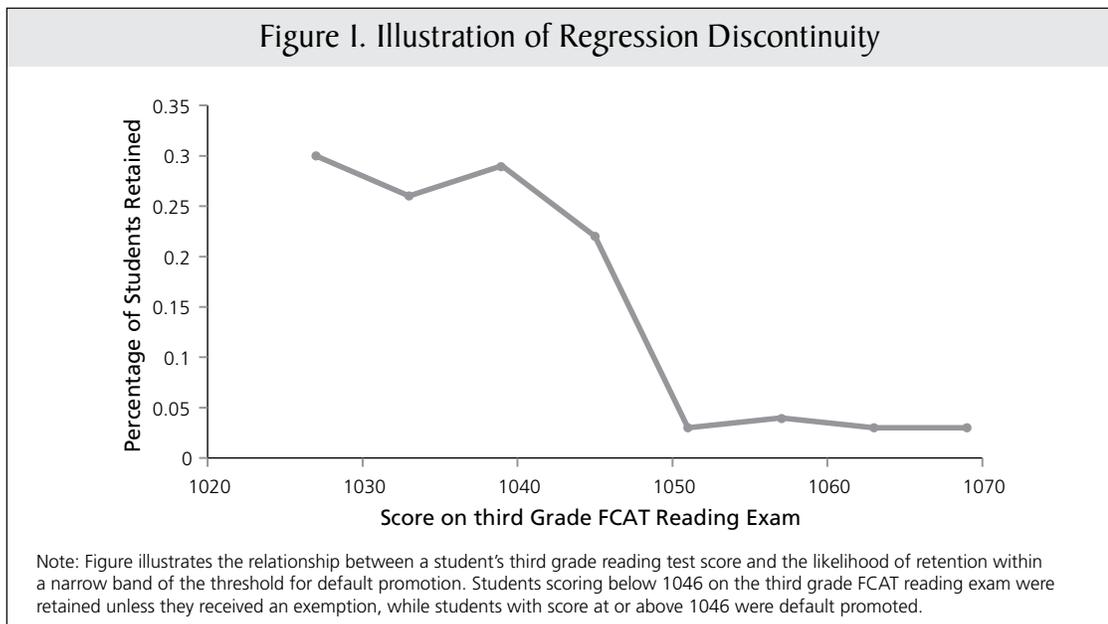
depends on where their third-grade reading score falls relative to a known benchmark. As mentioned, under Florida’s policy, students needed to score at or above Level 2 on the state’s third-grade reading exam to be default-promoted to the fourth grade; students who scored in Level 1 were retained unless they received an exemption. The cutoff for the intervention was a score of 1046 on the test’s developmental scale.

An important implication of this policy design is that students with scores very near the Level 2 benchmark have academic proficiencies that are very similar to one another. The difference between a student scoring just above or below the threshold was often one or two questions on the exam—differences that can be chalked up to luck, rather than meaningful differences in knowledge or ability. Students with test scores within a narrow neighborhood of the cutoff for remediation eligibility thus were very similar to one another except that one group faced the possibility of remediation under the policy while the other group did not and was instead default-promoted to the next grade level.

Figure 1 illustrates the relationship between student third-grade reading scores and the probability of retention during the 2003–04 school year. The dots on the figure represent possible scores on the exam

according to the scoring scale—thus, the figure shows the four possible scores above and below the threshold. The figure shows a clear discontinuity in the probability that a student is retained under the policy at the threshold score. For instance, though the difference in their scores was likely only a single additional correct answer on the exam, about 22 percent of the students with the next-lowest possible score below the test-score threshold were retained in 2003–04, while only about 3 percent of the students with the next-highest possible score on the test were retained.

Our analyses essentially compares the later academic outcomes of students with third-grade reading scores just below the threshold for default promotion—many of whom were retained and received the remediation treatments—with those of peers with scores just above the threshold (the vast majority of whom were promoted), controlling for other observed characteristics about them.<sup>8</sup> Because among this group of students, randomness played a significant role in determining whether they were subjected to the intervention, we can say with high confidence that the treatment and control groups are identical in every way, both observed and unobserved, except for their exposure to the remediation treatment. Thus, unlike many other papers on this topic, the regression discontinuity procedure enables us to measure the effect of remediation under



**Table I. Results—Reading**

	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Entered 3rd Grade 2003-04					
Retention (predicted)	0.311***	0.224***	0.231***	0.183***	
	[0.0404]	[0.0447]	[0.0459]	[0.0491]	
Entered 3rd Grade 2004-05					
Retention (predicted)	0.292***	0.238***	0.210***		
	[0.0399]	[0.0499]	[0.0470]		
Entered 3rd Grade 2005-06					
Retention (predicted)	0.385***	0.327***			
	[0.0447]	[0.0490]			
Entered 3rd Grade 2006-07					
Retention (predicted)	0.409***				
	[0.0490]				

Florida’s policy independent of other factors, such as the student’s maturity level.

We compare the academic outcomes of remediated and socially promoted students when they reach the same grade levels during their academic careers. Like some other earlier researchers, we argue that the within-grade comparison is the most policy-relevant because it best aligns with what schools are interested in: the student’s performance relative to his same-grade peers.<sup>9</sup> In addition, we point out that, long-term, an additional year of schooling is potentially one of the most important interventions.

In this analysis, we follow four cohorts of students from their initial third-grade year. The first cohort we consider is the entering third-grade class of 2003–04. Our data allow us to follow this group of students through the seventh grade. To test the robustness of our results, we also follow each subsequent cohort of third-grade students for which data are available, ending with the fourth-grade performance of students who entered the third grade in 2006–07.

## RESULTS

Table 1 reports the results from our analyses of student reading scores. The table shows the coefficient estimate of the effect of remediation

on student achievement within the grade level listed in the column head. Results are in standard deviation units, which we will put into greater context below.

Under Florida’s policy, the results from reading show a very large and sustained effect of remediation on student achievement. Remediation has a very large effect in the grades immediately following it. That effect appears to fade as the student progresses through middle school. However, by the seventh grade, the performance of remediated students from the second cohort subjected to the policy was about 0.183 standard deviations greater than it was for their socially promoted peers. Though I will put that result into greater context below, it is important to note at this point that this effect is substantial.

Also notable in the table is that the effect of remediation appears to be very similar for each of the cohorts evaluated. Remediated students who entered the third grade in 2005–06, for instance, had very similar achievement relative to their socially promoted peers, as did students who entered the third grade in 2003–04. The fact that multiple cohorts appear to have experienced similar outcomes because of the remediation suggests that our findings are quite robust.

Table 2 reports very similar results for the effect of remediation on student math achievement. Again, we see a very large immediate effect from remediation that appears to fade somewhat over time. However,

Table 2. Results—Math					
	Grade 4	Grade 5	Grade 6	Grade 7	Grade 8
Entered 3rd Grade 2003-04					
Retention (predicted)	0.360***	0.268***	0.199***	0.174***	
	[0.0431]	[0.0470]	[0.0526]	[0.0525]	
Entered 3rd Grade 2004-05					
Retention (predicted)	0.368***	0.276***	0.217***		
	[0.0455]	[0.0461]	[0.0483]		
Entered 3rd Grade 2005-06					
Retention (predicted)	0.416***	0.328***			
	[0.0438]	[0.0506]			
		Cohort 5			
Entered 3rd Grade 2006-07					
Retention (predicted)	0.523***				
	[0.0479]				

by the time they are in seventh grade—five years after the remediation decision—treated students are substantially outperforming their socially promoted peers in math.

Figures 2 through 5 illustrate the effect of remediation on student reading achievement for the entering third-grade class of 2003–04 in the fourth through seventh grades. The figures—known as kernel densities—are visual representations of the student reading scores. The dark line represents the scores of students who were remediated after their initial third-grade year, and the light line represents the scores of students who were promoted after the third grade. Only students with initial third-grade test scores within a very narrow band of the eligibility threshold are included in the figures. Thus, when considering the figures, it is important to keep in mind that these students had essentially identical test scores when they were in the third grade together.

Consistent with the results reported in Table 1, the figures show that in each grade subsequent to the third, remediated students are, on average, outperforming their socially promoted peers. Though relatively high and low performers exist in both the remediated and promoted groups, the pattern is clear: on average, remediated students substantially outperform their socially promoted peers beginning in the fourth grade,

lose some of this ground through middle-school grades, but are still achieving at higher levels as late as the seventh grade.

## RESULTS IN CONTEXT

Our results indicate that by the time they are in the seventh grade—five years after the remediation decision—remediated students, on average, outperform their socially promoted peers by about 0.18 standard deviations in reading and 0.17 standard deviations in math. Let’s put those results into context by comparing them with those found for other education interventions.

That the effect of treatment under Florida’s remediation policy remains statistically significant five years after the intervention distinguishes it from other educational interventions; commonly, most benefits fade over time. For instance, research (conducted using a “gold-standard” randomized design) has found that the positive effects of the Head Start program fade to the point of statistical insignificance by the end of the first grade.<sup>10</sup>

The magnitude of the sustained effect of third-grade remediation under Florida’s test-based promotion policy is also noteworthy. Figure 6 puts the size of this

Figure 2. Fourth Grade Reading

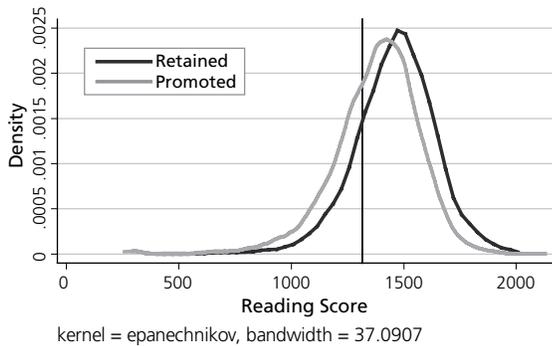


Figure 3. Fifth Grade Reading

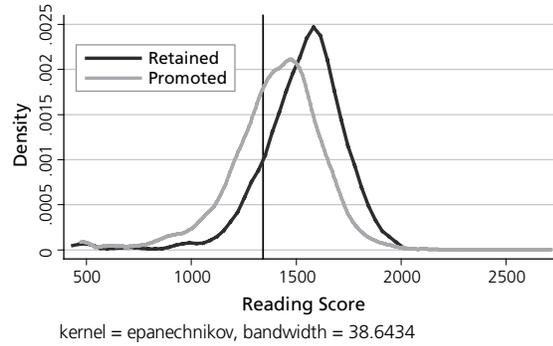


Figure 4. Sixth Grade Reading

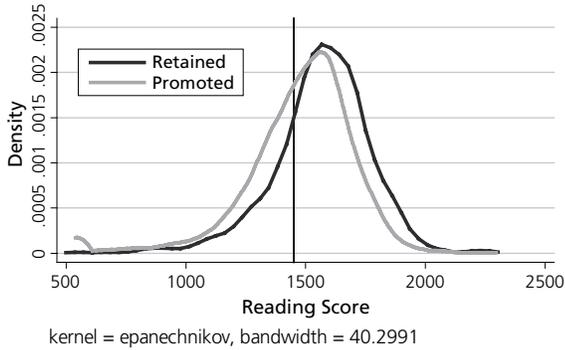


Figure 5. Seventh Grade Reading

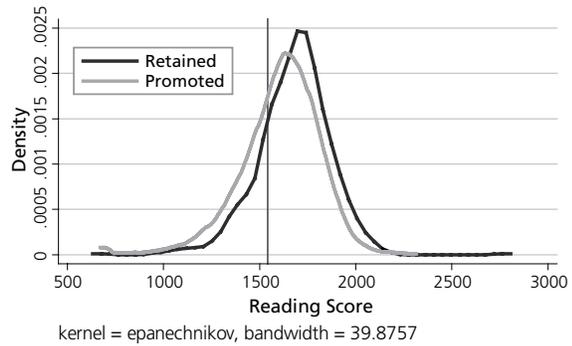
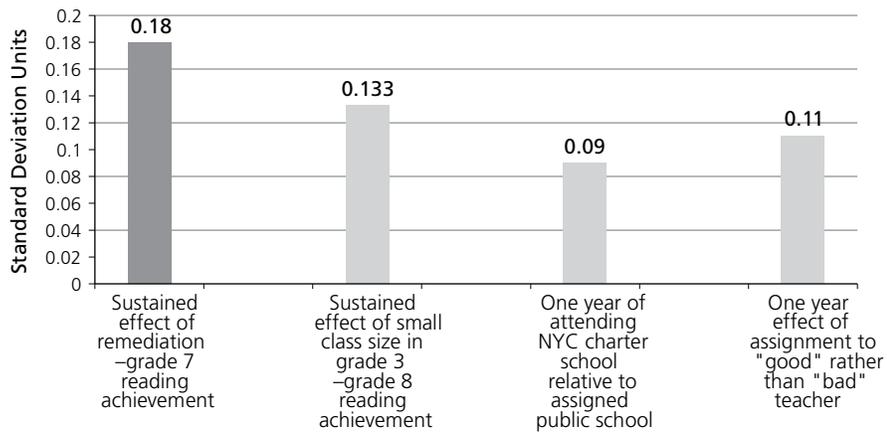


Figure 6. Treatment Effects of Educational Interventions



Source: Class Size - Nye et al. (1999); NYC Charter School - Hoxby, Murarka and Kang (2009); Teacher Quality - Rivkin, Hanushek, and Kain (2005)

result into context by comparing it with what some high-quality research finds to be the effect of other important policy interventions often characterized as having a large effect on student outcomes. The sustained benefit of Florida’s remediation policy is substantially larger than the one-year effect of a student being assigned to a “good” instead of a “bad” teacher or the one-year effect of attending one of New York City’s charter schools. The sustained effect of remediation after five years is also larger than what research has found to be the five-year effect of assignment to a small class size in the third grade.

## CONCLUSION

The results of our analyses are very encouraging for the use of Florida’s test-based promotion policy. We find evidence that students remediated under the policy make large academic gains relative to their socially promoted peers—gains that are meaningful and sustained at least through middle school.

There remains much to learn about the overall effects of Florida’s policy. In future years, it will be

important to evaluate the effect of early remediation on the probability that a student graduates from high school. Research analyzing whether the academic gains resulting from the treatment are worth the cost of the program to the taxpayer is also needed. Finally, the effect that the policy has on students when they first enter the third grade has not yet been examined adequately.

Given that policies can differ across school systems, it is important to note that our results strictly apply only to test-based promotion policies identical in structure to Florida’s program. We are not able to completely disaggregate the effect of retention from that of summer school attendance and other coinciding interventions. However, in the academic version of this paper, we do provide evidence that the policy’s requirement that a student be assigned to a “high-quality” teacher the following year does not appear to drive the effects from treatment. Nonetheless, policymakers should be aware that we can say that Florida-style test-based promotion has a large and sustained positive effect on student achievement and that they should use Florida’s experience as a guide for designing other remediation policies.

## ENDNOTES

---

1. For literature reviews expressing this view of the research, see Holmes 1989; and Jimerson 2001.
2. Quoted in “Third Grade Again: The Trouble with Holding Students Back,” *TheAtlantic.com*, February 14, 2012.
3. Allen et al. 2009.
4. See Jacobs and Lefgren 2004; Jacob and Lefgren 2007; and Roderick and Nagaoka 2005.
5. Greene and Winters 2007.
6. Greene and Winters 2009.
7. In this paper, I show results from estimating the effect of remediation on students from the second through fifth cohorts of students subjected to the policy. I omit discussion of the first cohort subjected to the policy for technical reasons related to the interpretation of the results interpretation. In short, the first cohort subjected to the policy—the entering third-grade class of 2002–03—represents a special case. Because they were the first cohort subjected to the policy, students who were promoted to the next grade out of this group entered the fourth grade with peers who were of far greater quality than did later cohorts of students subjected to the policy. Essentially, students who were promoted to the fourth grade at the end of 2002–03 no longer shared classrooms with a large number of very low-performing students, who were instead retained in the third grade. Because peer quality has been found to influence student achievement, the situation faced by this group of students cannot be generalized to the effect of the policy on later cohorts. The results from these students and the discussion of this measurement issue are provided in the more technical version of this paper.
8. Our primary approach utilizes a two-stage least-squares procedure that uses an indicator for whether the student’s test score was below the threshold as an instrument for remediation. Analyses are restricted to those individuals whose initial third-grade test score was within a small neighborhood around the eligibility threshold. The models, described in the academic version of this paper, control for student gender, race/ethnicity, free or reduced-price lunch eligibility, disability status, and a fixed effect for the student’s school. As also reported in the academic version of this analysis, we provide a series of visual and empirical tests showing the validity of the regression discontinuity design.
9. See, e.g., Alexander, Entwisle, and Dauber 1994.
10. Puma, Bell, Cook, and Heid 2010.

## REFERENCES

---

- Allen, C. S., Q. Chen, V. L. Wilson, & J. N. Hughes (2009). Quality of research design moderates effects of grade retention on achievement: A meta-analytic, multilevel analysis. *Educational Evaluation and Policy Analysis*, 31: 480-499.
- Greene, J. P., & Winters, M. A. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, 2(4), 319-340.
- Greene, J. P., & Winters, M. A. (2009). The effects of exemptions to Florida's test-based promotion policy: Who is retained? Who benefits academically? *Economics of Education Review*, 28, 135-142.
- Holmes, C. T. (1989) Grade-level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.) *Flunking grades: Research and policies on retention* (pp. 16-33). London: Falmer.
- Hoxby, C., Murarka, M. S., & J. Kang (2009). "How New York City's Charter Schools Affect Achievement, August 2009 Report." Second report in series. Cambridge, MA: New York city Charter Schools Evaluation Project, September 2009.
- Jacob, B. A., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *Review of Economics and Statistics*, 86, 26-244.
- Jacob, B. A., & Lefgren, L. (2007). The effect of grade retention on high school completion. NBER Working Paper W13514.
- Jimerson, S. R. (2001). A synthesis of grade retention research: Looking backward and moving forward. *California School Psychologist*, 6, 47-59.
- Nye, Barbara, Larry V. Hedges & Spyros Konstantopoulos. (1999) "The Long-Term Effects of Small Classes: A Five-Year Follow-Up of the Tennessee Class Size Experiment." *Educational Evaluation and Policy Analysis*. 21(2): 127-142.
- Puma, Michael, Stephen Bell, Ronna Cook, Camilla Heid, et al. (2010). "Head Start Impact Study: Final Report." *U.S. Department of Health and Human Services*. Washington, D.C.
- Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). "Teachers, schools, and academic achievement." *Econometrica*, 73(2), 417-458.
- Roderick, M., & Nagaoka, J. (2005). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *Educational Evaluation and Policy Analysis*, 27, 309-340.





CENTER FOR STATE AND  
LOCAL LEADERSHIP

Michael Allegretti  
*Director*

FELLOWS

Rick Baker  
Josh Barro  
Daniel DiSalvo  
Edward Glaeser  
Nicole Gelinas  
Steven Malanga  
Edmund J. McMahon  
Fred Siegel  
Jacob Vigdor  
Marcus A. Winters

The Manhattan Institute's **Center for State and Local Leadership (CSLL)** promotes promising new approaches to reform of state and local government. CSLL works on a broad range of issues, including public sector reform (specifically of pensions and health benefits), education reform, prisoner reentry, policing, public housing, infrastructure, immigration, and public service delivery. By spotlighting new ideas, and providing the research and proposals to inform creative new policies, the CSLL hopes to lay the groundwork for an environment in which commerce, employment, and a rich civic life can flourish.

The CSLL operates across the country, working in states such as California, Illinois, and Rhode Island, and cities such as Newark, New Jersey and Detroit, Michigan. The CSLL's tools include regular writing and research reports by affiliated Manhattan Institute scholars and senior fellows, along with public events and media appearances. The CSLL operates [www.PublicSectorInc.org](http://www.PublicSectorInc.org), a website devoted to analysis of the implications—financial and political—of the power wielded by public sector unions and allied elected officials. The CSLL also annually selects and showcases the Manhattan Institute's Urban Innovator Award, which recognizes a state or local leader whose combination of innovativeness and skill at implementation has led to groundbreaking improvements in public service, and the Manhattan Institute's Social Entrepreneurship Awards, which recognize those who identify social needs and take it upon themselves to address them privately.

[www.manhattan-institute.org/csll](http://www.manhattan-institute.org/csll)

The Manhattan Institute is a 501(C)(3) nonprofit organization. Contributions are tax-deductible to the fullest extent of the law. EIN #13-2912529