

Senate Committee on Education**February 25, 2013****Hearing on SB 169**

**Testimony by Dr. Marcus Winters, Senior Fellow at the Manhattan
Institute, Assistant Professor in the College of Education at the
University of Colorado Colorado Springs**

Chairman Abrams and distinguished members of the committee:

Thank you for offering me the opportunity to speak to you today as you consider legislation to enact a policy requiring students to demonstrate a minimal level of proficiency prior to entering the fourth grade. I truly appreciate the opportunity to present to you the results of my recent empirical evaluation of the sustained impact of a similar program that has been operating in Florida since 2002. Our research has found strong evidence that the remediation treatment offered under Florida's policy has led to substantial improvements in student proficiency that have been sustained for a meaningful period of time.

Like the policy under consideration today, Florida's test-based promotion policy is meant to end the longstanding practice of "social promotion", whereby children are promoted to the next grade level for

Senate Education Committee

Date 2-25-13

Attachment 3

socialization reasons even if they have not developed the skills necessary to succeed there. Florida's policy provides remediated students with several treatments -- they attend summer school, are required to be assigned to a "high quality" teacher during the retained year, are provided with targeted reading instruction, and are provided with an educational plan. However, what makes the policy particularly controversial is that it requires students to demonstrate possession of some minimal skill in order to avoid grade retention.

Opponents of test-based promotion policies point to a wide body of research that seemingly shows that retention harms later student outcomes. But while 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. The research techniques utilized by most earlier studies of grade retention fail to meet modern research standards. Consequently, with only a few recent exceptions, most previous studies on the effect of retention on student outcomes provide misleading conclusions.

Many prior studies on grade retention have simply compared the later outcomes of retained students to 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. A researcher can't account for a characteristic like the student's maturity level because it does not appear in her dataset. Nonetheless, the student's maturity level is very likely to be related to their 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.

Fortunately, researchers have developed several techniques capable of making true apples-to-apples comparisons. However, most of the studies cited by opponents of test-based promotion policies do not use the more rigorous methods. Of the twenty-two 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”.

When considering test-based promotion policies, I'd urge policymakers to only consider the few existing papers that use what is known as a "regression discontinuity" design to evaluate modern programs. These papers deserve particular attention because, unlike even very sophisticated matching strategies, under minimal assumptions regression discontinuity accounts for both observed *and unobserved* differences between remediated and socially promoted students. Regression discontinuity is one of the few research strategies strong enough that the U.S. Department of Education's What Works Clearinghouse considers it capable of making causal estimates.

Research using this methodology has found that Chicago's test-based promotion policy for third grade students had a short run positive effect that faded away relatively quickly, and a recent evaluation of the short-run effect of New York City's third grade promotion policy found positive results.

In our recent paper, we utilize a regression discontinuity design to study the sustained impact of remediation under Florida's test-based promotion policy on student achievement. Along with our use of a high-quality research design, a distinguishing feature of our analysis is that we

follow students for up to five years after the remediation decision -- as late as the seventh grade. Our research has been peer reviewed and published in the well-regarded economics journal *Education Finance and Policy*.

Though the statistical details powering our study are complicated, the intuition behind the regression discontinuity approach is relatively easy to understand. The procedure takes advantage of the fact that a student's likely exposure to Florida's remediation policy depends on where their third grade reading score falls relative to a known benchmark. Under Florida's policy students needed to score at or above Level 2 on the state's third grade reading exam in order to be default promoted to the fourth grade; students who scored in Level 1 were retained unless they received an exemption.

An important implication of this policy design is that students with scores very near, but on either side of 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 guessed right on the exam. Students with test scores within a narrow neighborhood of the cutoff for remediation eligibility thus were very similar to each other 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.

Our analysis 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-- to those of peers with scores just above the threshold -- the vast majority of whom were promoted. Because among this group of students randomness played a significant role in determining whether or not 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 allows us to measure the effect of remediation under Florida's policy independent of other factors, such as the student's maturity level.

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 allows 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 was available.

We found that remediation and retention under Florida's policy was related to substantial improvement in both math and reading. 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 was substantially larger than the seventh grade performance of their socially promoted peers. The results were similar on both the state's high-stakes tests as well as low-stakes standardized math and reading exams. We also found that remediation had a positive effect on performance on an elementary science exam. Finally, that the effects of the policy were similar for multiple cohorts of students lends credibility to our estimates.

The magnitude of the sustained effect of third grade remediation under Florida's test-based promotion policy is noteworthy relative to the effects of other interventions considered to have large academic benefits. 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.

Further, that the effect of treatment under Florida's remediation policy remains statistically significant five years after the intervention distinguishes it from other educational interventions. For instance, research has found that

the positive effects of the Head Start program fades to the point of statistical insignificance by the end of the first grade.

The results of our analysis 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.

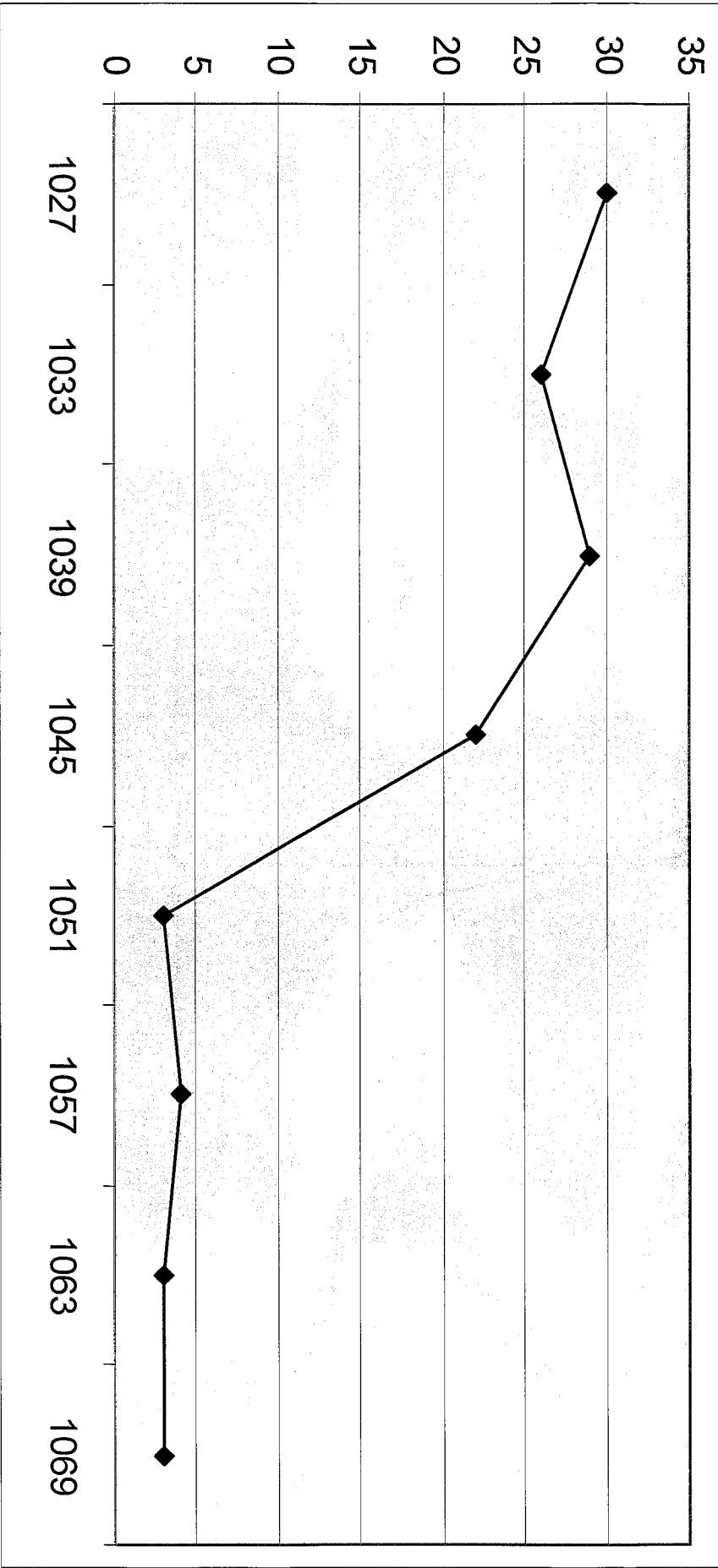
It is important to note that our results only strictly apply to test-based promotion policies identical in structure to Florida's program. We are not able to completely disaggregate the effect of retention from that of summer school attendance and other coinciding interventions. However, we do provide 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.

What we can say is that Florida-style test-based promotion has a large and sustained positive effect on student achievement. Our results suggest that we can look to Florida's experience as a guide for designing a remediation policy for Kansas's students as well.

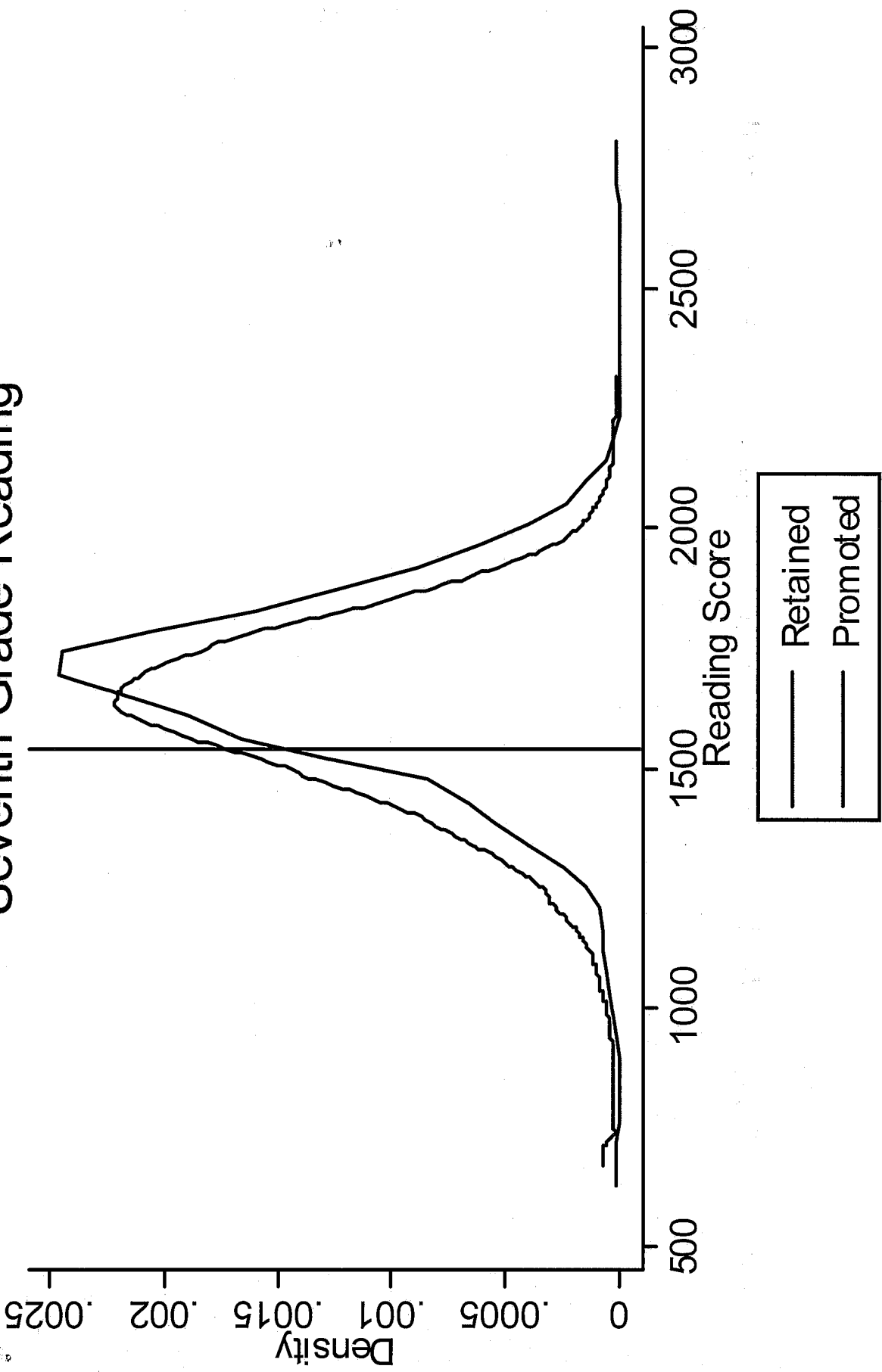
Thank you. I look forward to your questions.

Percentage of Students Retained



3-11

Seventh Grade Reading



kernel = epanechnikov, bandwidth = 39.8757