Test-Based Grade Retention

AEFP Live Handbook
Umut Özek
K-12 Education
Standards and Accountability

Introduction

Holding back students deemed not ready for more challenging course content is a practice nearly as old as formal education in the United States.[1] Yet this practice has evolved considerably over the past three decades.

In the twentieth century, grade retention decisions were mainly discretionary, with teachers, parents, and/or principals exercising their individual or collective judgment about whether a student should be required to repeat a grade. In the 1990s, concerns about "social promotion" along with the increasing popularity of educational accountability and standardized testing led to the implementation of universal policies that link grade retention decisions to student performance on standardized tests in many states and school districts.

For example, in 1997, President Bill Clinton cited a pioneering test-based retention program in Chicago Public Schools and called on all school districts to "adopt high standards, to abolish social promotion, [and] to move aggressively to help students make the grade through tutoring and summer schools."[2] A decade later, six states and twelve large school districts had adopted test-based promotion policies. In 2023, 26 states and the District of Columbia had policies that either allow or require retention of third-grade students who lag on their third-grade reading scores.[3]

Test-based grade retention policies vary by grade, subject, criteria for retention, and the instructional support provided to students flagged for retention before and after the ultimate retention decision. For example, Florida's third-grade retention policy, which was enacted in 2002 and provided a blueprint for many other states, uses student test scores on the spring reading test to flag students for retention. These students are then eligible to participate in a summer reading program to improve their reading skills. Not all students flagged for retention on the basis of their test scores are ultimately retained because of good-cause exemptions.[4] Further, for students who are ultimately retained, schools are required to develop academic improvement plans that specifically address their needs, to assign these students to high-performing teachers (categorized as such according to their student performance and performance appraisals), and to provide 90 minutes of daily reading instruction in the following school year. Grade retention policies in school districts in other states (e.g., Ohio, New York City, Texas), on the other hand, use multiple tests (the first to flag students for retention and instructional support and the final one to make the ultimate retention decision) or do not include explicit requirements for the instructional support to be provided for retained students (as in, e.g., Texas).

Despite their increasing popularity, test-based grade retention policies remain controversial.[5] Proponents argue that an additional year of instruction in the same grade provides struggling students with a realistic chance to catch up academically. Opponents of these policies argue that grade retention imposes significant emotional burdens on students (e.g., stigma associated with being labeled as failing, challenges of adjusting to new peers), which might in turn lead to student disengagement from schooling. Critics also point to the monetary expenses of grade retention policies such as the cost of the additional year of instruction and the forgone wages for retained students.

What does the research say? The overarching conclusion of early literature on discretionary retention policies is that retained students perform significantly worse than their promoted peers in the years that follow. However, more recent studies, which better address identification challenges by leveraging natural experiments with retention policies (e.g., in regression discontinuity designs), portray a more nuanced picture. In this chapter, I discuss some of the overarching findings from this recent body of research and highlight the potential avenues for future research.

Key findings

Key finding #1: *Grade retention in elementary school (typically grades 3–5) usually improves student outcomes in middle and high school when implemented as part of a broader remediation effort.*

Several recent studies in Florida, Texas, Indiana, Mississippi, Chicago, and New York City suggest that early grade retention increases test scores in elementary and middle school, reduces the need for future remediation, and increases the likelihood that the students take advanced courses in middle and high school. These benefits are more pronounced in settings where grade retention is coupled with instructional support for both retained and other at-risk students who were eventually promoted.

Key finding #2: *Emerging evidence suggests that the short-term benefits of early grade retention policies could extend beyond the students held back.*

Spillover effects of early retention include positive effects on the students flagged for retention but not ultimately retained, possibly because the higher expectations lead students and teachers to act to improve learning and avoid retention. One study also suggests positive effects on the younger siblings of retained students.

Key finding #3: *Recent evidence suggests that the long-run cost of early grade retention to states could be considerably lower than an additional year of schooling.*

This is primarily driven by the reduced need for future remediation for students who were retained in elementary school, while students who are at risk of retention but ultimately promoted often take longer than four years to graduate high school.

Key finding #4: *Early evidence about the effects of grade retention in elementary school on postsecondary and labor market outcomes suggests zero or negative effects, yet more research is needed for these long-term effects to be better understood.*

In contrast to the vast evidence on the short-term test score effects of early grade retention, only two studies to date have looked at its associated postsecondary and labor market outcomes, revealing precisely estimated zero effects on postsecondary enrollment in Florida and significant negative effects on earnings in Texas.

Key finding #5: *Grade retention in middle and high school typically leads to worse educational outcomes, with little or no effect on academic achievement and higher levels of student disengagement.*

Even in settings where retained middle and high school students are provided with instructional supports similar to those of their counterparts in elementary grades, evidence suggests that retention reduces persistence in high school, the high school graduation rate, and the likelihood of college enrollment and degree attainment and increases the likelihood of involvement in criminal activity and arrest as an adult.

Evidence

Early evidence on the effects of test-based grade retention in elementary grades has come from Chicago, New York City, and Florida, three of the earlier adopters of these policies that provided the blueprints for many other school districts and states. These have been followed by studies set in Indiana, Mississippi, and Texas. To identify the causal effects of grade retention on student outcomes, most of these studies use regression discontinuity designs to compare students whose test scores fell right below a retention cutoff to those with scores right above it, under the assumption that these students are similar along all baseline attributes other than the policy-driven retention decision.

An important challenge researchers face when evaluating the effects of grade retention is when to measure student outcomes. Policy evaluations in other contexts typically compare the outcomes of treatment and comparison groups at the same point in time after the treatment (e.g., 2 years after the treatment). In the context of grade retention, this same-age comparison implies comparing the outcomes of retained and promoted students in different grades, which could be problematic for two reasons.

First, this approach would fail to capture one of the important mechanisms through which grade retention could improve student outcomes: the additional year of schooling that might better prepare students for more challenging course content. Second, student outcomes may not be comparable across grade levels. For example, in many settings, students take grade-specific tests, which makes it difficult for researchers to compare the test scores of retained and promoted students. Similarly, comparing student misbehavior across grades could be problematic as disciplinary enforcement could vary across grade levels (e.g., school administrators in middle school could be less lenient toward student misbehavior than those in elementary school).

Thus, studies on grade retention have typically relied on same-grade comparisons, looking at the outcomes of retained and promoted students when they reach the same grade (e.g., comparing the

8th grade test scores of the promoted students with the scores of retained students when they reach 8th grade). While this approach also has its shortcomings, it captures the effect of the additional year of instruction.[6]

Key finding #1: *Grade retention in elementary school (typically grades 3–5) usually improves student outcomes in middle and high school when implemented as part of a broader remediation effort.*

Test scores: Grade retention in elementary grades, when coupled with instructional support for students flagged for retention and those ultimately retained, significantly improves student test scores in elementary grades. Retained students (and those flagged for retention) also outperform their promoted peers when they reach the same grade levels in middle school.[7] For example, in Florida, retained students right below the retention cutoff score nearly 20% of a standard deviation higher on the 8th grade reading test than their promoted peers right above the cutoff.[8] Recent evidence also reveals that these benefits could be substantially larger for students with lower baseline achievement (whose test scores fall below and away from the cutoff), suggesting that regression discontinuity studies that focus on students with scores right below these cutoffs could underestimate the overall effect of these policies on student test scores.[9]

In addition to grade-specific scale scores, Florida reports vertically aligned developmental scale scores, which enable same-age comparisons between retained and promoted students. Similarly to the same-grade analysis, this exercise reveals sizable positive effects of retention in the short term (1–3 years), yet these effects fade entirely in middle and high school.[10] This discrepancy between the same-grade and same-age results is likely driven by the fact that the latter fail to account for the effect of the additional year of instruction on test scores.

Nontest outcomes: Evidence from New York City, Florida, Indiana, and Mississippi also suggests that grade retention in elementary school, when implemented as part of a broader remediation effort, reduces the students' likelihood of being retained in later grades, reduces the likelihood that students take remedial courses in middle and high school, and increases their likelihood of taking advanced courses in middle and high school.[11] These positive effects are likely driven by improved test scores, which are typically important determinants of grade promotion and course-taking in later grades.

What about measures of student disengagement from schooling such as disciplinary incidents and attendance? Only a few studies to date have looked at the effects of early grade retention coupled with instructional support on disciplinary outcomes, yielding mixed results. Evidence from Florida suggests a short-term increase in student suspensions that vanishes in middle school, another study from New York City reveals a similarly short-lived decline, and a recent study from Indiana finds no significant effect.[12] Findings also suggest no significant effect of these early grade retention policies on future absences.[13] Further, evidence from descriptive surveys indicates that students retained in elementary school reported a greater sense of school connectedness, lasting several years beyond retention, than comparable students who were promoted, providing evidence against the potential adverse effects of early grade retention on student disengagement.[14]

Key finding #2: *Emerging evidence suggests that the short-term benefits of early grade retention policies could extend beyond the students held back.*

There are several ways other than retention itself through which early grade retention policies could improve student outcomes. For example, as mentioned earlier, many early grade retention policies nationwide incorporate instructional support not only for the students ultimately retained but for

all students who fail to score above the retention cutoff. Further, higher expectations for student performance could lead to improved instructional quality and student effort. Thus, scoring below the cutoff could affect student outcomes through channels beyond its effect on the likelihood of retention.

Similarly, in many settings, states and school districts use multiple tests to identify which students are to be retained, with the first one flagging students for retention and the last one used for the retention decision. This “threat” of retention could affect student outcomes regardless of the ultimate retention decision if, for example, it sends a high-stakes signal to parents, students, and teachers about the students' educational needs.

In a new study, researchers examine these channels in Michigan schools and find that students who scored just below the retention cutoff on the first test significantly outperformed their peers right above the cutoff on standardized tests even though very few students in the state were eventually retained.[15] These findings suggest that the threat of retention coupled with instructional support provided to students flagged for retention could be important mediators behind the observed effects of retention policies, making it difficult to attribute these effects to retention alone. That said, it is plausible to expect that this threat effect could be less relevant in the long run if very few students end up being retained in practice.

Another recent study suggests that early grade retention policies could have positive spillover effects on the younger siblings of retained students.[16] In particular, linking student records with birth records in Florida (hence identifying siblings in administrative school records), the researchers find that the younger siblings of students whose third-grade reading scores fell right below the retention cutoff (and who were thus flagged for retention) scored significantly higher on standardized tests in elementary school than the younger siblings of students with scores right above the retention cutoff. These spillover effects are partly driven by improved educational resources for the younger siblings in subsequent years, suggesting that the high-stakes retention signal for the older siblings might inform parents and educators about the educational needs of the younger sibling and induce them to act.

Key finding #3: *Recent evidence suggests that the long-run cost of early grade retention to states could be considerably lower than an additional year of schooling.*

A potential criticism of grade retention policies is that they impose a substantial financial burden on states, as schools must offer an additional year of schooling to retained students. However, a study set in Florida illustrates how the long-run cost of early grade retention could be much smaller than the cost of an additional year of schooling for several reasons.[17]

First, several studies suggest that early grade retention policies could significantly reduce the likelihood of future remediation in middle and high school (e.g., grade retention or remedial course-taking).[18] In contrast, students flagged for retention but ultimately promoted often fail to graduate from high school on time.[19]

Aside from the costs and benefits for the retained students, another important factor to consider in assessments of the cost-effectiveness of retention policies is their spillover effects. For example, as noted above, the threat of retention could improve learning experiences for students in earlier grades (or during the grade of retention), reducing the share of students scoring below the retention cutoff in the grade of retention: In Florida, the share of first-time third graders scoring below the retention threshold dropped from 21% to 14% in the first five years of implementation of its

retention policy.[20] From a public policy perspective, these spillover effects are “free” and, as such, may have profound effects on the cost-effectiveness of early grade retention policies.

In short, the reduced cost of remediation across all affected students may substantially offset the higher costs for the very small number of students actually retained in grade.

Key finding #4: *Early evidence about the effects of grade retention in elementary school on postsecondary and labor market outcomes suggests zero or negative effects, yet more research is needed for these long-term effects to be better understood.*

While many studies to date have looked at the effects of early grade retention policies on middle and high school outcomes of students, we know relatively little about their effects on postsecondary and labor market outcomes. One study set in Florida examines the effects of test-based retention in third grade on postsecondary enrollment and finds precisely estimated zero effects (even though it also reveals sizable positive same-grade test score effects of retention).[21]

A more recent study examines the effects of Texas's third-grade retention policy using a regression discontinuity design.[22] The findings suggest that retention has no statistically significant effect on postsecondary enrollment or degree attainment, yet it significantly reduces annual earnings by age 26. In Texas, in contrast to several other settings, retention also reduces the likelihood of high school graduation and increases disciplinary incidents and juvenile crime in middle school, and while it improves test scores in the short run, these benefits fade (in both the same-age and same-grade comparisons).

These discrepancies could be driven by several features of Texas’s retention policy. In Texas, students are given three chances to score above the retention threshold (instead of one in Florida and two in several other settings such as New York City). Thus, very few students end up being retained: In the 2003–2004 school year, only 2% of third-grade students in Texas were retained (in Florida, this number was nearly 12% in the same school year), which could lead to higher levels of stigma for retained students.[23] More importantly, Texas is one of the few states that does not require school districts to provide instructional support for students after they are retained (although students are provided support before the final test retake), which could explain some of the more negative findings in this setting. (This is similar to Louisiana where research has revealed more negative effects of grade retention [24]).

Key finding #5: *Grade retention in middle and high school typically leads to worse educational outcomes, with little or no effect on academic achievement and higher levels of student disengagement.*

In contrast to the positive effects of grade retention in elementary school when it is coupled with instructional support, the research on retention in middle school grades finds negative effects, even in settings where similar instructional supports are provided to retained students. In Chicago, test-based grade retention in middle school reduces the likelihood of high school graduation, especially among African American girls.[25] In New York City, students retained in middle school are less likely to pass enough New York State Regents exams to graduate from high school and are likelier to drop out late in high school.[26] In Louisiana, retention in middle and high school leads to a reduced likelihood of high school graduation and college enrollment and increased likelihood of involvement in criminal activity in adulthood.[27]

While we need more research to better understand why grade retention policies that provide students with similar instructional support have different effects on student outcomes when implemented at different grade levels, there are several possible explanations. For example, one common argument against grade retention policies is that they place a significant emotional burden on students; such stigma effects might be more pronounced for older students, who might also have more difficulty adjusting to new peers. Further, students retained in earlier grades might have more time and opportunities to catch up with their peers, or instructional support might be more beneficial for students earlier in their schooling.

Conclusion

Despite the rapid growth in the literature on test-based retention policies over the past two decades, several important questions remain. First, we need more evidence on the long-term effects of early grade retention policies on postsecondary and labor market outcomes, especially in settings where retention is coupled with instructional support for low-performing students. These findings will help give a fuller picture about the benefits of early grade retention programs.

Second, an important limitation of regression discontinuity designs, on which most of this literature relies, is that they can say very little about the effects on students away from the cutoff who are flagged for retention. New research suggests that the test score effects of retention could be quite different for these students than for those who just missed the retention cutoff.[28] Thus, how retention policies affect other short- and long-term outcomes of students with lower baseline achievement is an important research question that remains.

Third, more research is needed on the spillover effects of grade retention policies. These include the effects of test-based grade retention policies on student learning in earlier grades, their effects on higher-performing students who experience a shift in peer composition after the retention grade, and the mechanisms behind such spillover effects.[29] Finally, we need more evidence about why the observed positive effects of early grade retention in some settings (e.g., same-grade effects on middle and high school test scores or advanced course-taking) do not necessarily translate into positive effects on postsecondary and labor market outcomes.

Endnotes and references

[1] Owings, William A., and Susan Magliaro. 1998. Grade Retention: A History of Failure. Educational Leadership & Workforce Development Faculty Publications 110.
[2] Brownstein, Ronald. 1997. Clinton Calls for End to "Social Promotion" in Schools. Los Angeles Times. October 29.
[3] https://reports.ecs.org/comparisons/state-k-3-policies-2023-14, accessed on 12/22/2025.
[4] Good-cause exemptions under Florida's policy include: (1) students who have been in the English learner program for less than two years, (2) students who have an individual education plan (IEP) that indicates that participation in the Florida Comprehensive Assessment Test (FCAT) program is not appropriate, (3) students who demonstrate an acceptable level of performance via a student portfolio or an alternative standardized reading assessment approved by the State Board of Education, (4) students who have certain disabilities and have been already retained once until third grade, and (5) students who have received intensive reading remediation for two years and have already been retained twice between kindergarten and third grade.
[5] https://www.chalkbeat.org/2023/6/12/23758532/grade-retention-social-promotion-studies-reading-research-mississippi/, accessed on 12/22/2025.
[6] Schwerdt, G., and M. R. West. 2013. The Effects of Test-Based Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida. CESifo Working Paper Series 4203. Schwerdt and West, 2013, provide a detailed discussion about the assumptions behind the same-age and same-grade comparisons in the context of grade retention.
[7] Greene, J.P., and M. A. Winters. 2007. Revisiting Grade Retention: An Evaluation of Florida's Test-Based Promotion Policy. Education Finance and Policy 2(4): 319–340; Greene, J.P., and M. A. Winters. 2012. The Medium-Run Effects of Florida's Test-Based Promotion Policy. Education Finance and Policy 7(3): 305–330; Mariano, L. T., and P. Martorell. 2013. The Academic Effects of Summer Instruction and Retention in New York City. Educational Evaluation and Policy Analysis 35(1): 96–117; Jacob, B. A., and L. Lefgren. 2004. Remedial Education and Student Achievement: A Regression-Discontinuity Analysis. The Review of Economics and Statistics 86(1): 226–244; Schwerdt, G., M. R. West, and M. A. Winters. 2017. The Effects of Test-Based Retention on Student Outcomes over Time: Regression Discontinuity Evidence from Florida. Journal of Public Economics 152: 154–169; Figlio, D., and U. Özek. 2020. An Extra Year to Learn English? Early Grade Retention and the Human Capital Development of English Learners. Journal of Public Economics, 186: 104184; Hwang, N., and C. Koedel. 2023. Helping or Hurting: The Effects of Retention in the Third Grade on Student Outcomes. Educational Evaluation and Policy Analysis. Prepublished October 12; Slungaard Mumma, K., and M. A. Winters. 2023. The Effect of Retention Under Mississippi's Test-Based Promotion Policy. Working Paper 2023-1. Wheelock Educational Policy Center. Available at wheelockpolicycenter.org; Zhong, J. Forthcoming. Early Grade Retention Harms Adult Earnings. American Economic Journal: Applied Economics.
[8] Schwerdt et al., 2017.
[9] Opper, I. M., and U. Özek. 2025. A Global Regression Discontinuity Design: Theory and Application to Grade Retention Policies. EdWorkingPaper: 23-798. Annenberg Institute at Brown University.
[10] Schwerdt et al., 2017; Figlio and Özek, 2020.
[11] Schwerdt et al., 2017; Figlio and Özek, 2020.
[12] Özek, U. 2015. Hold Back to Move Forward? Early Grade Retention and Student Misbehavior. Education Finance and Policy 10(3): 350–377; Martorell, P., and L. T. Mariano. 2018. The Causal Effects of Grade Retention on Behavioral Outcomes. Journal of Research on Educational Effectiveness 11(2): 192–216.
[13] Figlio and Özek, 2020; Slungaard Mumma and Winters, 2023; Hwang and Koedel, 2023.
[14] McCombs, J. S., S. N. Kirby, and L. T. Mariano, eds. 2009. Ending Social Promotion Without Leaving Children Behind: The Case of New York City. RAND Corporation, MG-894-NYCDOE.
[15] Berne, J. S., B. Jacob, C. Weiland, and K. O. Strunk. 2026. Assessing the Impact of Grade Retention: A Cautionary Tale of Exclusion Restriction Violations. Journal of Public Economics 253: 105534.
[16] Figlio, D., K. Karbownik, and U. Özek. 2023. Sibling Spillovers May Enhance the Efficacy of Targeted School Policies. NBER Working Paper 31406.
[17] Winters, M. A. 2022. The Cost of Retention Under a Test-Based Promotion Policy for Taxpayers and Students. Educational Evaluation and Policy Analysis 45(4): 688–695.
[18] Schwerdt et al., 2017; Figlio and Özek, 2020.

[19] Mariano, L., P. Martorell, and T. Berglund. 2024. The Effects of Grade Retention on High School Outcomes: Evidence from New York City Schools. Journal of Research on Educational Effectiveness 18(1): 1–31.
[20] LiCalsi, C., U. Özek, and D. Figlio. 2019. The Uneven Implementation of Universal School Policies: Maternal Education and Florida's Mandatory Grade Retention Policy. Education Finance and Policy 14(3): 383–413.
[21] Schwerdt et al., 2017.
[22] Zhong, forthcoming.
[23] https://tea.texas.gov/reports-and-data/school-performance/accountability-research/retention2003-04.pdf, accessed on January 7, 2026; LiCalsi et al., 2019.
[24] Larsen, M. F., and J. Valant. 2023. Fuzzy Difference-In-Discontinuities When the Confounding Variation Is Sharp: Evidence from Grade Retention Policies. Applied Economics Letters 30(15): 2009–2013; Larsen, M. F., and J. Valant. 2024. The Long-Term Effects of Grade Retention: Evidence on Persistence Through High School and College. Journal of Research on Educational Effectiveness 17(4): 615–646; Eren, O., B. Depew, and S. Barnes. 2017. Test-Based Promotion Policies, Dropping Out, and Juvenile Crime. Journal of Public Economics 152: 9–31; Eren, O., M. Lovenheim, and N. Mocan. 2022. The Effect of Grade Retention on Adult Crime: Evidence from a Test-Based Promotion Policy. Journal of Labor Economics 40(2): 361–395.
[25] Jacob, B. A., and L. Lefgren. 2009. The Effect of Grade Retention on High School Completion. American Economic Journal: Applied Economics 1(3): 33–58.
[26] Mariano et al., 2024.
[27] Larsen and Valant, 2023; Larsen and Valant, 2024; Eren et al., 2017; Eren et al., 2022.
[28] Opper and Özek, 2024.
[29] To the best of my knowledge, only one study to date has looked at the effects of grade retention policies on promoted students. Examining Indiana's retention policy, Hwang and Koedel, 2025, find limited effects on the educational outcomes of promoted students. Hwang, NaYoung, and Cory Koedel. 2025. The Peer Effects of Grade Retention. EdWorkingPaper: 25-1146. Annenberg Institute at Brown University.