Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1556

Estimating the Overall Effects of Statewide Early Literacy Policies: Evidence from Michigan

Seth Walker
Stanford University

Josep M. Nadal-Fernandez
Tulane University

Tara Kilbride
Michigan State University

This paper evaluates the impacts of Michigan's statewide early literacy reform using a synthetic difference-in-differences approach. We estimate the policy's effects on student achievement and grade retention and compare them to statewide literacy reforms in South Carolina and Mississippi. We find that the policy increased third-grade reading achievement by 8.7% of a standard deviation, with smaller gains in later grades and larger effects for cohorts exposed to the policy for longer. Importantly, these achievement gains remain positive and significant as students progress beyond third grade. We also find no evidence of spillover effects on math achievement or increases in retention. Compared to South Carolina, Michigan experienced larger and more consistent gains, though effects remained smaller than those in Mississippi.

VERSION: August 2026

Suggested citation: Walker, Seth, Josep M. Nadal-Fernandez, and Tara Kilbride. (2026). Estimating the Overall Effects of Statewide Early Literacy Policies: Evidence from Michigan. (EdWorkingPaper: 26-1556). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/5n45-9v54

Estimating the Overall Effects of Statewide Early Literacy Policies: Evidence from Michigan

Seth Walker[a,i,*], Josep M. Nadal-Fernandez[a,ii], Tara Kilbride[a]

[a] *Education Policy Innovation Collaborative (EPIC), Michigan State University, East Lansing, MI 48824, USA*
[i] *Present Address: SCALE Initiative, Stanford University, Stanford, CA 94305, USA*
[ii] *Present Address: Education Research Alliance for New Orleans, Tulane University, New Orleans, LA 70118, USA*
[*] *Corresponding author*
Email addresses: stw1034@stanford.edu (S. Walker); jnadalfernandez@tulane.edu (J.M. Nadal-Fernandez); kilbrid9@msu.edu (T. Kilbride)

ABSTRACT

This paper evaluates the impacts of Michigan's statewide early literacy reform using a synthetic difference-in-differences approach. We estimate the policy's effects on student achievement and grade retention and compare them to statewide literacy reforms in South Carolina and Mississippi. We find that the policy increased third-grade reading achievement by 8.7% of a standard deviation, with smaller gains in later grades and larger effects for cohorts exposed to the policy for longer. Importantly, these achievement gains remain positive and significant as students progress beyond third grade. We also find no evidence of spillover effects on math achievement or increases in retention. Compared to South Carolina, Michigan experienced larger and more consistent gains, though effects remained smaller than those in Mississippi.

JEL Classification: I21, I28, H75

Keywords: Early literacy; student achievement; grade retention; education policy; literacy interventions

1. INTRODUCTION

In the past two decades, over 40 U.S. states have adopted early literacy reforms (ExcelinEd, 2024), with a common focus on ensuring students reach reading proficiency by the end of third grade—a critical benchmark for future learning (Cunningham & Stanovich, 1997; Fiester & Smith, 2010; Hernandez, 2011; Sparks et al., 2014). These reforms have received renewed attention following declines in national fourth-grade reading achievement after the COVID-19 pandemic, with 2024 National Assessment of Educational Progress (NAEP) reading scores falling below 2019 levels and reaching their lowest levels on record (National Assessment of Educational Progress, 2024). However, evaluating the effectiveness of these policies is challenging due to variations in curricula, resources, and implementation across states. Most studies focus on third-grade retention policies, taking advantage of the clear eligibility cutoffs that determine which students are retained (Jacob & Lefgren, 2004; Greene & Winters, 2007; Winters & Greene, 2012; Mariano & Martorell, 2013; Özek, 2015; Roderick & Nagaoka, 2016; Schwerdt et al., 2017; Weiss et al., 2018; Figlio & Özek, 2020; Slungaard Mumma & Winters, 2023). While this approach provides clear insight into the effects of retention at these cutoffs, it does not capture the broader set of supports and interventions (e.g., early screening, literacy coaching, and home reading plans) included in many early literacy policies that are designed to improve reading proficiency for all students.

A smaller literature examines the overall effects of early literacy policies on student achievement across states. For example, Westall and Cummings (2023) evaluate early literacy reforms using cross-state variation in policy adoption and distinguish between key policy features such as retention requirements and broader supports. Although this provides useful evidence on how different types of policies relate to student outcomes, it does not isolate the effects of a specific policy within a state. More recently, Spencer (2024) studies Mississippi's Literacy-Based Promotion Act and finds large positive effects on fourth-grade reading and math achievement, but focuses on Mississippi alone. Relatedly, Mahmud (2025) examines state-mandated third-grade retention policies across several states and finds that retention mandates produce larger fourth-grade achievement gains when paired with broader literacy supports, though the analysis does not separately evaluate individual statewide reforms beyond Mississippi.

Michigan's Read by Grade Three (RBG3) law provides an opportunity to examine the effects of one such statewide literacy reform. Enacted in 2016, the law targets early literacy through a range of supports and interventions. The law created new early literacy funds to support these efforts, with the ambitious goal of ensuring that Michigan would become one of the top 10 most improved states in fourth-grade reading proficiency by the 2019 NAEP and rank among the top 10 states overall by 2025 (Michigan Legislature, 2016). In its first four years, Michigan spent $132.6 million on this initiative (Strunk et al., 2021), with approximately 91% of funding distributed as early literacy funds to intermediate school districts, school districts, and schools[i].

By studying Michigan's RBG3 law, this paper provides a policy-specific evaluation of the effects of a comprehensive early literacy reform on all students. Using a synthetic difference-in-differences approach, we find that the policy increased third grade reading achievement by 8.7% of a standard deviation, with smaller effects in later grades and no evidence of spillover effects on math achievement. These effects grow with exposure to the policy and remain positive across exposed

[i] Calculations based on *Michigan Department of Education Financial Data Files* (*https://www.mischooldata.org/financial-data-files/*). See the breakdown of grant distribution in Table A1.

cohorts as they progress through later grades. Effects are also larger in districts with lower baseline achievement, suggesting that the policy may have helped reduce pre-existing achievement disparities across districts. We also examine the impact of the policy on grade retention, finding no evidence of an increase in retention. This result indicates that improvements in reading achievement were driven by the broader set of supports and interventions embedded in the policy rather than retention alone. Finally, we compare Michigan's results to estimated effects from literacy reforms in Mississippi and South Carolina to examine how differences in policy design, state context, and implementation relate to policy effectiveness. In doing so, we show how this approach can be applied to evaluate other state-specific policies.

The remainder of this paper is organized as follows. Section 2 provides an overview of Michigan's RBG3 policy, outlining its key components, implementation timeline, and state context leading up to the reform. Section 3 describes the data sources used in this analysis, while Section 4 outlines the empirical methodology. Section 5 presents the main results, heterogeneity analyses, and cross-state comparisons. The paper concludes with a discussion of the findings and implications for early literacy policies.

2. POLICY BACKGROUND

2.1 Early Literacy Skills and Policy Efforts

The development of fundamental literacy skills by third grade is considered a key milestone in the learning process (Fiester & Smith, 2010), with some evidence showing that these early literacy skills persist and are predictive of reading ability even ten years later (Cunningham & Stanovic, 1997). Furthermore, literacy proficiency is associated with stronger employment and earnings outcomes among individuals with similar levels of educational attainment and may help reduce disparities in labor market outcomes across individuals with different levels of formal education (Lane & Conlon, 2016).

Despite the importance of early literacy skills, fewer than one-third of U.S. fourth-grade public school students perform at or above the NAEP Proficient level in reading (National Assessment of Educational Progress, 2024). In response, policymakers have introduced a range of reforms aimed at improving literacy outcomes. These include federal initiatives such as the Reading Excellence Act and Reading First, which expanded support for K-3 reading instruction and literacy interventions (U.S. Congress, 1998; U.S. Department of Education, Office of Elementary and Secondary Education, 2002). While evaluations of the latter found limited impacts on reading comprehension (Gamse et al., 2008), these initiatives helped spur the adoption of state-level literacy reforms. One prominent example is Florida's *Just Read, Florida!* initiative, implemented in 2002 and often viewed as a blueprint for many subsequent state literacy policies (Cummings et al., 2021). Evidence from Florida shows improvements in reading outcomes in both the short run (Greene & Winters, 2007) and medium term (Winters & Greene, 2012), providing support for the effectiveness of early state-level literacy reforms.

2.2 Michigan's Read by Grade Three Law

Michigan's RBG3 law built on earlier state efforts to address early literacy challenges. In the late 1990s and early 2000s, the state implemented a plan that provided parents with literacy kits at birth to encourage early reading (Cummings et al., 2021). However, it was not until the mid-2010s—amid

declining fourth-grade NAEP reading scores relative to other states[ii]—that Michigan pursued a more comprehensive legislative approach. In 2016, state legislators enacted the RBG3 law to improve literacy skills by the end of third grade through a coordinated set of literacy interventions and supports, including early screening and diagnostic reading assessments, the development of Individual Reading Intervention Plans (IRIPs), literacy coaching and professional development for teachers, and reading intervention programs.

This policy is implemented through a Multi-Tiered System of Support (MTSS)[iii] (Strunk et al., 2021) that provides increasingly intensive supports based on students' literacy needs. Under this framework, Tier I focuses on improvements to core literacy instruction provided to all students. As part of this effort, the law requires that Intermediate School Districts employ at least one literacy coach to facilitate professional development among teachers. In addition, school districts must administer a diagnostic assessment approved by the Michigan Department of Education (MDE) at least three times per year to all K-3 students to identify those in need of additional support. Students requiring further assistance receive Tier II interventions, including IRIPs, additional instructional time, small-group or one-on-one instruction, summer programming, and parental involvement activities, with approximately 33% of students receiving these supports (Education Policy Innovation Collaborative, 2019). Tier III builds on Tier II but is delivered with greater intensity and individualization for students with more substantial needs, covering approximately 4% of students (Strunk et al., 2022).

The law also included a third-grade retention provision for students scoring one grade level below on the English Language Arts (ELA) portion of the Michigan Student Test of Educational Progress (M-STEP). Originally scheduled to take effect in the 2019-20 academic year, implementation was delayed until 2020-21 due to the COVID-19 pandemic before being repealed for third-grade cohorts beginning in 2023-24. Because the retention component took effect after the period examined in this paper, our analyses do not capture any direct effects of mandatory retention.

3. DATA

3.1 Early Literacy Policy Data

To assess the impact of Michigan's law, this analysis draws on multiple data sources. First, to identify which states had early literacy policies during our sample period, we use a database compiled by ExcelinEd (2024) that documents the timing and characteristics of early literacy initiatives across all states. We supplement this dataset with information from state departments of education and legislative records to verify policy timing and document key components of each policy. Using this information, we construct indicators for whether and when each state adopted a statewide early literacy policy, as well as identify two states—Mississippi and South Carolina—that implemented early literacy policies around the same time as Michigan.

[ii] While Michigan was 3 points above the national public NAEP fourth-grade reading average in 1998, the state was five points below the average and ranked 41st in the nation by 2015.

[iii] See for example this guide illustrating how to implement this system at the school district level: https://resources.finalsite.net/images/v1568836530/resanet/drbszjpnchsgxle0u5cq/QuickguideforMTSSTheDistrictLevel.pdf.

3.2 Student Achievement Data

To analyze student achievement, we use data from the Stanford Education Data Archive (SEDA) 5.0. This dataset includes district-level test scores in grades 3-8 for both math and ELA from the 2008-09 to 2018-19 academic years.[iv] These scores are derived from a combination of state assessment data and NAEP scores, which are administered in odd-numbered years for fourth and eighth graders. SEDA uses linear interpolation and extrapolation to generate estimates for non-tested grades and years, thereby providing a complete grade-subject-year dataset (Fahle et al., 2024). We use the cohort-standardized (CS) scale, which expresses test scores in standard deviations (SD) relative to a common national baseline, allowing for comparisons across states, grades, and years.[v] To aid interpretation, 1 SD on the CS scale corresponds to approximately three grade levels (Fahle et al., 2024).

Additionally, for some districts, the SEDA dataset offers disaggregated test score data by gender, race/ethnicity, and socioeconomic status, which we use to examine heterogeneity in policy effects across student groups. The dataset also includes district-level characteristics such as locale, racial/ethnic composition, percentage of students eligible for free or reduced-price lunch, and unemployment rates.

Our analytical sample includes only districts represented in the SEDA data across all years, grades, and subjects to construct the balanced panel required for our analysis and avoid changes in sample composition. This restriction modestly reduces the number of in-sample districts in Michigan, with the resulting sample covering approximately three-quarters of tested students in the state. As shown in Table 1, these districts are broadly comparable to the full set of districts in terms of demographics, though they are slightly less rural and have smaller shares of Black and economically disadvantaged students. There are also differences in baseline third-grade ELA test scores, with in-sample districts exhibiting significantly higher average performance. Table A2 reports comparable statistics for South Carolina and Mississippi. While the sample restriction reduces coverage substantially more in Mississippi (20.3%) than in South Carolina (97.6%), the resulting samples in both states remain broadly reflective of the statewide testing population.

3.3 Retention Data

To examine retention outcomes, we use school-level retention counts by grade from the Civil Rights Data Collection (CRDC), administered by the Office for Civil Rights (OCR) within the U.S. Department of Education. Since the 2011-12 academic year, the CRDC has collected data biennially from public schools nationwide, apart from the 2019-20 collection, which was postponed due to the COVID-19

[iv] Although other versions of SEDA extend beyond the 2018-19 school year, we choose this analytic window due to the well-documented impact of the COVID-19 pandemic on student achievement (Kilbride et al., 2022; National Assessment of Educational Progress, 2024) that we cannot disentangle from policy impacts.

[v] The SEDA dataset has been widely used to estimate changes in district student achievement across the entire United States (e.g., Jang & Reardon, 2019; Chen & Harris, 2023) and validity analyses indicate high congruence between SEDA growth estimates and other benchmarks (Kuhfeld et al., 2019; Reardon et al., 2019).

pandemic. Consequently, our analysis includes data from the 2011-12, 2013-14, 2015-16, and 2017-18 academic years.[vi]

Table 1. Michigan SEDA Sample Comparison

(1)	(2)	(3)	(4)	(5)
Variable	**All Districts**	**Districts in Sample**	**Districts Not in Sample**	**t-test p value**
% Urban	0.071	0.087	0.048	0.059
% Town/Suburb	0.440	0.504	0.343	<0.001
% Rural	0.489	0.409	0.610	<0.001
% White	0.844	0.859	0.821	0.060
% Black	0.077	0.065	0.096	0.058
% Hispanic	0.046	0.047	0.045	0.843
% Asian	0.018	0.021	0.013	0.020
% Economically Disadvantaged	0.451	0.392	0.541	<0.001
% Free-or-reduced Price Lunch	0.462	0.400	0.555	<0.001
% English Learner	0.020	0.018	0.022	0.390
% Special Education	0.135	0.133	0.138	0.051
Average 2008 3rd Grade ELA Score	-0.037	0.028	-0.135	<0.001
2008 3rd Grade ELA Tested Students	113,290	85,244	28,046	–

Note: This table reports baseline means for a set of variables for all Michigan districts in SEDA 5.0 (column (2)), those included in the analytic sample (column (3)), and those excluded from the analysis (column (4)). Column (5) reports the p-value from a t-test of differences in means between included and excluded districts.

To construct retention outcomes, we combine the CRDC data with enrollment data from the National Center for Education Statistics **(**NCES) Common Core of Data (CCD) Membership files from the Public Elementary/Secondary School Universe Survey. These files provide school-level enrollment counts by grade, as well as enrollment by student race/ethnicity and free or reduced-price lunch (FRL) eligibility. We aggregate school-level retention counts and enrollment to the district level and define retention rates as the proportion of students in a given grade, district, and year who are retained. The resulting units correspond to administrative school districts (i.e., districts defined by the set of schools they operate), whereas the SEDA data are defined over geographic school districts (i.e., districts defined by

[vi] While the latest CRDC datasets also contain the 2020-21 and 2021-22 collections, we exclude them from our analysis as the results could be confounded by the COVID-19 pandemic.

school location within district service areas, regardless of governance). To maintain consistency across analyses, we restrict the sample to districts observed in all years and grades.

Table 2 shows that 90-95% of third-grade students are represented in our analytic sample across the states included in the analysis. Despite this high level of coverage, excluded districts differ from those included in the sample. In particular, they serve significantly higher shares of non-white and FRL-eligible students. For example, in Michigan, the percentage of non-white students is considerably higher in excluded districts (42.24%) than in included districts (25.69%), and the share of FRL-eligible students is also higher (67.31% vs. 54.53%, respectively). Similar patterns are observed in South Carolina and Mississippi. Retention rates are also higher in excluded districts, though these differences are not statistically significant.

Table 2. CRDC Sample Comparison

(1)	(2)	(3)	(4)	(5)
Variable	**All Districts**	**Districts in Sample**	**Districts Not in Sample**	**t-test p value**
		Panel A: Michigan		
% Non-White	28.684	25.812	41.820	<0.001
% FRL	56.841	54.582	67.175	<0.001
% Retained	1.242	1.205	1.412	0.586
2011 3rd Grade Enrollment	109,414	102,703	6,711	
		Panel B: South Carolina		
% Non-White	52.413	48.840	76.129	0.002
% FRL	68.004	66.231	79.770	0.028
% Retained	2.408	2.424	2.303	0.897
2011 3rd Grade Enrollment	51,429	49,349	2,080	
		Panel C: Mississippi		
% Non-White	57.631	52.372	80.027	<0.001
% FRL	80.381	78.215	89.608	0.001
% Retained	5.424	4.367	9.926	0.132
2011 3rd Grade Enrollment	36,107	32,532	3,575	–

Note: This table reports baseline means for a set of variables for all districts (column (2)) in CRDC, those included in the analytic sample (column (3)), and those excluded from this analysis (column (4)). Column (5) reports the p-value from a t-test of differences in means between included and excluded districts.

3.4 Policy Implementation Data

To explore whether differences in implementation led to different effects of Michigan's law, we draw on educator survey data collected in 2020. The survey, described in detail in Strunk et al. (2021), covered nearly 19,000 teachers, principals, district superintendents, and literacy coaches regarding policy understanding, perceptions of the law, instructional practices, literacy coaching, and capacity to implement the law, among other topics. We use exploratory factor analysis to construct composite measures of implementation based on survey responses from educators serving grades K-5.

To complement these implementation measures, we also compare financial resources available to support early literacy reforms across states. First, we use data from the National Public Education Financial Survey to examine annual K–12 per-pupil expenditures, providing a broader measure of education spending and fiscal capacity in each state. We then compile annual state appropriations dedicated specifically to early literacy initiatives from state budget documents for Michigan, Mississippi, and South Carolina. These appropriations are converted to annual K–3 per-pupil funding to facilitate comparisons in the level of financial support directly tied to each state's early literacy policy. Finally, to provide a more detailed picture of implementation in Michigan, we use district financial records to track expenditures financed through approved early literacy funds. We link these expenditures to object and function codes to characterize how districts allocated early literacy resources across spending categories.

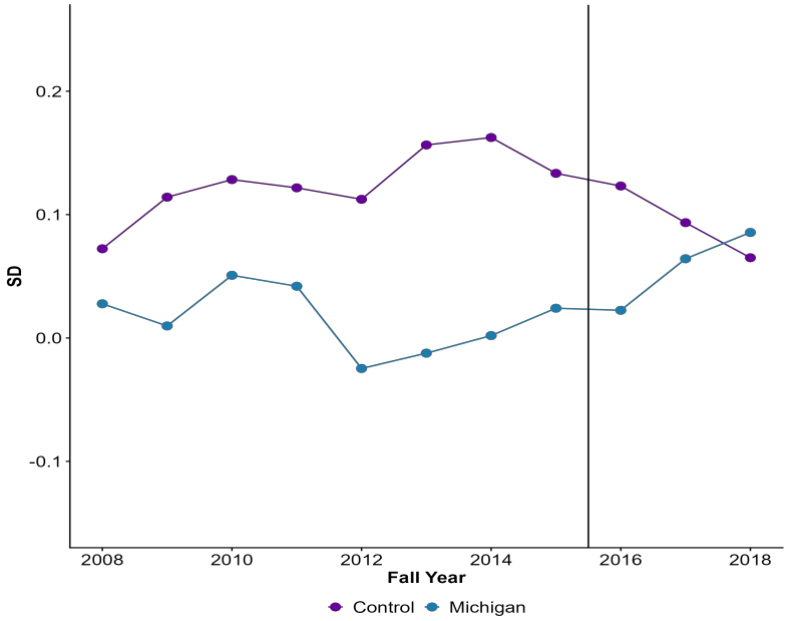
4. EMPIRICAL STRATEGY

Michigan's early literacy policy was implemented statewide and targeted all K-3 students, leaving no untreated comparison group within the state. As a result, identification must rely on comparisons with other states. We therefore construct a comparison group consisting of states without a statewide early literacy policy during our sample period. Table A3 lists the states without a statewide early literacy policy by fall 2018 and indicates which states are included in our pool of control states for the achievement and retention analyses after applying the balanced-panel requirements described above.

A natural approach is a difference-in-differences (DID) framework comparing Michigan to these states before and after the policy. However, this approach relies on a parallel trends assumption that is unlikely to hold in this context. As shown in Figure 1, Michigan and the comparison states exhibit different pre-policy trends in reading achievement, most notably in the years immediately preceding the policy (2012-15), when reading scores move in opposite directions. These differences in pre-policy trends raise concerns that a standard DID would attribute existing differences in trajectories to the policy.

An alternative approach would be to use a synthetic control method to construct a comparison unit that closely matches Michigan prior to the policy. However, this approach is also challenging as Figure 1 shows Michigan's pre-policy achievement levels are substantially lower than those of comparison states, with average gaps of up to 0.2 SDs in the years leading up to the policy. As a result, it is difficult to construct a synthetic control that simultaneously matches both the level and trajectory of Michigan's pre-policy outcomes.

Figure 1. Pre-Policy Trends in Third-Grade ELA Achievement



Note: This figure plots average third-grade ELA test scores in Michigan and comparison states from the 2008–09 through 2018–19 school years using SEDA CS SDs. Control states include states without a statewide early literacy policy during the sample period. The vertical line indicates the first year of implementation of Michigan's RBG3 law.

Given these concerns, one could instead select a subset of control states that appear to follow similar pre-policy trends, though such choices are inherently ad hoc. Instead, we adopt the synthetic difference-in-differences (SDID) estimator proposed by Arkhangelsky et al. (2021), which provides a systematic way to construct a comparison group from a balanced panel of control units. Like synthetic control methods (Abadie et al., 2010), SDID uses weighted combinations of control units to approximate the treated unit. However, SDID allows for differences in pre-policy levels while matching pre-policy trends, thereby addressing the challenges described above. This approach combines the strengths of both synthetic control and DID frameworks and has been used in related settings, including Spencer (2024) to study Mississippi's early literacy policy.[vii] Equation 1 presents the SDID estimator for the average treatment effect on the treated (ATT):

$$(\hat{\tau}^{SDID}, \hat{\mu}, \hat{\alpha}, \hat{\beta}) = arg\min_{\tau, \mu, \alpha_i, \beta_t} \left(\sum_{i=1}^{N} \sum_{t=1}^{T} (Y_{it} - \mu - \alpha_i - \beta_t - EL_{it}\,\tau)^2 \hat{\omega}_i^{SDID} \hat{\lambda}_t^{SDID} \right) \quad (1)$$

where EL_{it} indicates periods in which treated units are exposed to the early literacy policy. The estimated effect ($\hat{\tau}^{SDID}$) comes from a weighted two-way fixed effects regression with time (β_t) and unit (α_i) fixed effects. Estimated unit weights ($\hat{\omega}_i^{SDID}$) and time weights ($\hat{\lambda}_t^{SDID}$) are determined by solving the optimization problems described in Arkhangelsky et al. (2021).

Intuitively, SDID estimates treatment effects by reweighting control units and pre-treatment periods to better match treated units prior to policy adoption. The unit weights place greater weight on control units with pre-policy outcomes similar to those of treated units, while the time weights place greater

[vii] Spencer (2024) uses all states in their primary comparison pool, but considers a more restrictive donor pool of states with "minimal-to-no" early literacy policies, similar to our comparison group definition, as a robustness check.

weight on pre-policy periods that best align with post-treatment outcomes for the control group. We compute standard errors using the bootstrap procedure implemented in the *synthdid* package in R.

We estimate the SDID model separately for each grade to reflect the policy's focus on early elementary students and to allow treatment effects to vary across grades. The policy took effect beginning in the 2016-17 school year, initially affecting students in third grade and below. In subsequent years, treatment expands mechanically as these cohorts progress through higher grades (e.g., the 2016-17 third-grade cohort is treated in fourth grade in 2017-18), resulting in later grades becoming treated over time. The initially treated third-grade cohort reaches fifth grade by the end of our sample period, meaning students in grades 6 through 8 are never directly exposed to the policy. We nonetheless estimate effects for these grades as a placebo test to assess whether significant effects in treated grades reflect broader contemporaneous changes rather than the policy itself. For these grades, we assign treatment beginning in 2018-19—the final year of our sample period.

As a robustness check, we also re-estimate Equation 1 including covariates. In addition, we estimate subgroup-specific ATTs to examine heterogeneity in policy effects across student populations. Finally, we apply the SDID procedure to Mississippi's and South Carolina's early literacy policies to compare their effects to those of Michigan's policy.

Following Clarke et al. (2023), we present SDID event study estimates to illustrate the dynamic effects of the policy. Unlike in a standard DID framework, comparisons to pre-treatment outcomes are not made relative to a single baseline period (e.g.,$t = -1$), but instead to a synthetic baseline constructed as a weighted average of pre-treatment periods using the estimated time weights $\hat{\lambda}_t^{SDID}$. As a result, differences between the synthetic baseline and individual pre-treatment periods do not necessarily indicate violations of identifying assumptions, as these weights vary across periods.

5. RESULTS

5.1 Main Results

Table 3 and Figure 2 present the main results on student achievement. Table 3 reports estimated effects by grade, with columns (2) and (3) showing ELA results without and with covariates, respectively, and columns (4) and (5) showing the analogous math results. Figure 2 presents event studies for ELA achievement by grade level. Each point represents the estimated effect in a given year relative to the synthetic pre-policy baseline, with the gray shaded regions denoting 95% confidence intervals. The solid vertical line indicates the first year in which a cohort was exposed to the RBG3 law, and the red shaded area displays the period weights used to construct the synthetic baseline.

Consistent with the policy's focus on early elementary grades, we find positive and statistically significant effects on ELA achievement in grades 3 through 5. The largest effects occur in third grade, with an estimated increase of 8.7% of an SD. Effects decline but remain significant in grades 4 and 5 (5.3% and 2.7% of an SD, respectively). We also observe a 3.2% of an SD increase in sixth grade, despite these students not being directly targeted at the time of policy implementation, potentially reflecting spillovers from broader school-level changes in literacy instruction. In contrast, estimates for grades 7 and 8 are small and statistically insignificant, consistent with the absence of directly exposed cohorts in these grades and providing evidence that the estimated effects in lower grades are unlikely to reflect contemporaneous shocks affecting all grades.

Figure 2 illustrates how these effects evolve across cohorts and exposure periods. Across grades, pre-policy estimates are generally small and statistically indistinguishable from zero. Following implementation, positive effects emerge immediately for grades with directly exposed cohorts and become larger for students exposed to the policy for longer periods of time. For example, students in third grade in 2016-17 experienced an increase of approximately 2.5% of an SD, while those in third grade in 2018-19—who were exposed to the policy in grades 1 through 3—experienced an increase of nearly 15% of an SD. We observe a similar pattern in fourth-grade test scores, suggesting that the policy's effects increase with cumulative exposure.

Table 3. Effects on ELA and Math Achievement

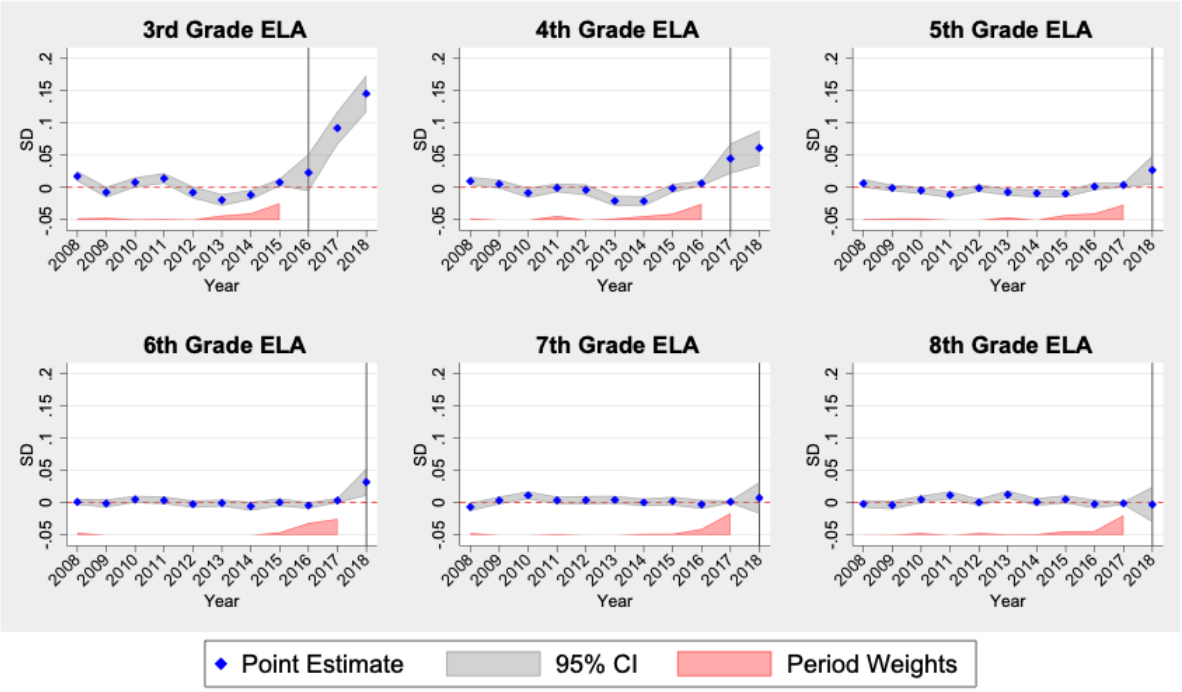
(1)	(2)	(3)	(4)	(5)
Grade	**ELA**		**Math**	
3rd	0.087***	0.093***	-0.010	0.008
	(0.0099)	(0.0104)	(0.0112)	(0.0141)
4th	0.053***	0.058***	0.002	0.016*
	(0.0113)	(0.0102)	(0.0109)	(0.009)
5th	0.027**	0.038***	-0.001	0.008
	(0.0115)	(0.0143)	(0.0119)	(0.0125)
6th	0.032***	0.028*	0.005	0.000
	(0.0102)	(0.0153)	(0.0110)	(0.0132)
7th	0.007	0.007	0.017	0.020**
	(0.0110)	(0.0124)	(0.0107)	(0.0099)
8th	-0.003	0.002	0.002	0.007
	(0.0140)	(0.0140)	(0.0120)	(0.0132)
Treated Districts	290	287	290	287
Untreated Districts	755	726	755	726
Covariates	No	Yes	No	Yes

*Note: This table reports estimated effects of the RBG3 law on ELA (columns (2)–(3)) and math (columns (4)–(5)) test scores, by grade. Each cell corresponds to a separate regression. Test scores are expressed in SEDA CS SDs. The double horizontal line indicates grades with cohorts directly exposed to the policy. Columns (3) and (5) include controls for time-varying district characteristics. Standard errors are reported in parentheses. *p<0.10; **p<0.05; ***p<0.01.*

However, this pattern may also reflect changes in the intensity of policy implementation over time. As shown in Table A1, the Michigan Department of Education provided nearly $13 million more in early literacy funds in 2018-19 than in 2016-17, potentially allowing for a broader and more intensive set of interventions over time. These channels are not mutually exclusive, so estimates for later cohorts may reflect both mechanisms.

A final takeaway from these figures is that effects persist as cohorts move beyond third grade. Following the 2016-17 third-grade cohort through grade 5 illustrates this pattern: point estimates increase from approximately 2.5% of an SD in third grade to nearly 5% of an SD in fourth grade and remain positive and significant in fifth grade (approximately 2.5% of an SD). Rather than fading after initial exposure, these gains in reading achievement appear to carry forward as students move to later grades.

Figure 2. Event Studies, 3rd-8th Grade ELA Achievement



Note: This figure plots SDID event study estimates of the effects of Michigan's RBG3 law on ELA test scores for grades 3 through 8. Test scores are expressed in SEDA CS SDs. Each point represents the estimated effect in a given year relative to the synthetic pre-policy baseline, with gray shaded regions denoting 95% confidence intervals. The vertical line indicates the first year in which a cohort was exposed to the policy, and the shaded red area displays the period weights used to construct the synthetic baseline.

Appendix Figure A1 reports the weights assigned to control states in the SDID estimation for third-grade ELA scores. The weights are concentrated among a small set of comparison states—Illinois, Pennsylvania, Georgia, and Alabama—with individual weights ranging from roughly 0.1 to 0.4 and accounting for approximately 95% of the total weight. For conciseness, we present weights only for the third-grade ELA analysis; similar patterns are observed for other grades, outcomes, subjects, and subgroup analyses.[viii]

Results are robust to the inclusion of covariates.[ix] This specification controls for changes in district characteristics, including demographic composition and economic disadvantage, to account for potential changes in district composition that could otherwise confound the estimated effects. Estimated effects remain similar to those reported in column (2), suggesting that changes in observable district characteristics are not driving the main results.

Columns (4) and (5) show little evidence that the early literacy policy meaningfully affected math test scores. In the few instances where estimates are marginally significant, point estimates are positive but small, suggesting that improvements in ELA achievement did not come at the expense of math performance. At the same time, they provide limited evidence of complementarities between early

[viii] Weights for all grades, outcomes, and subgroup analyses are available upon request.

[ix] This specification requires districts to have non-missing values for all covariates in each year, resulting in a slightly smaller sample than that used in column (2) of Table 3.

literacy interventions and math achievement, contrasting with findings from studies of grade retention interventions (Schwerdt et al., 2017) and broader policies in other states (Westall & Cummings, 2023; Spencer, 2024).

5.2 Heterogeneity

We next explore whether the effects of Michigan's RBG3 law differed across districts, student subgroups, and implementation contexts. We first examine heterogeneity by districts' baseline ELA achievement levels before estimating effects separately by student gender, race/ethnicity, and economically disadvantaged status. We then consider whether policy effects varied across districts with different reported implementation environments using educator survey-based implementation constructs. Each analysis is restricted to districts with available subgroup scores for all groups within a given comparison.

We begin by exploring whether effects differed by districts' baseline ELA achievement. We divide Michigan districts into quartiles based on third-grade ELA scores in 2015–16, with quartile cut points of –0.26, 0.01, and 0.24 SDs on the SEDA CS scale. We then apply these same cut points to districts in the control group. This approach produces roughly equal numbers of Michigan districts in each quartile, by construction, but relatively more control districts in the top quartile, reflecting the higher average baseline ELA achievement of control districts.

Table 4 reports estimated effects separately by quartile, while Figure A2 plots pairwise differences across quartiles. The table also contains average expenditure and implementation capacity levels by quartile. Estimated effects are generally smallest for districts in the top quartile (Q4) of baseline ELA achievement. In third grade, estimates range from 4.6% of an SD in Q4 to between 8.2% and 13.1% of an SD in the lower three quartiles, with differences between Q4 and both Q1 and Q3 statistically significant. We observe a similar pattern in fourth grade, where effects remain positive and significant in the three lowest-performing quartiles but are close to zero in Q4. While effects generally decay in later grades, estimated effects for Q3 districts remain positive and statistically significant through fifth grade. These results suggest that the policy may have helped reduce achievement disparities between lower- and higher-performing districts.

The relatively large and persistent effects observed in Q3 districts may initially appear surprising. Although these districts had higher initial achievement levels than Q1 and Q2 districts, many students likely remained below reading proficiency and therefore stood to benefit from the interventions required under the RBG3 law. At baseline, 69% of Michigan fourth-grade students scored below NAEP proficiency levels, suggesting that many students in Q3 districts were still likely to receive relatively intensive literacy supports under the policy. Consistent with this interpretation, expenditures on summer school programs—a targeted intervention for struggling readers—were similar across the first three quartiles and noticeably lower in Q4 districts. This pattern suggests that districts in Q3 continued to serve substantial populations of students requiring additional literacy support despite their relatively strong baseline performance.

Table 4. Effects on ELA Achievement by Baseline ELA Quartile

(1)	(2)	(3)	(4)	(5)
Grade	**Q1**	**Q2**	**Q3**	**Q4**
3rd	0.104***	0.082***	0.131***	0.046**
	(0.0215)	(0.0236)	(0.0272)	(0.0196)
4th	0.085***	0.046**	0.087***	0.007
	(0.0279)	(0.0203)	(0.0236)	(0.0197)
5th	0.027	0.029	0.052**	0.011
	(0.0261)	(0.0220)	(0.0260)	(0.0198)
Annual K-3 Per-Pupil Expenditure	$56.66	$53.79	$55.98	$55.16
Supported by RBG3 Law	(0.52)	(0.50)	(0.53)	(1.28)
Summer School	$3.98	$3.10	$4.73	$0.78
	(2.62)	(3.18)	(2.37)	(4.07)
Instructional Support	$23.62	$21.25	$23.38	$28.81
(ISD + District)	(14.81)	(11.90)	(13.34)	(14.27)
Instructional Support	$8.97	$6.18	$6.25	$14.68
(District)	(1.78)	(2.00)	(2.35)	(3.59)
Teacher Capacity to Implement	-0.058	-0.015	0.032	0.188
RBG3 Law	(0.261)	(0.294)	(0.373)	(0.329)
Treated Districts	73	73	72	72
Untreated Districts	160	167	144	284

*Note: This table reports estimated effects of the RBG3 law on ELA test scores by quartiles of baseline district ELA achievement and grade level. Quartiles are based on third-grade ELA scores in 2015–16, where Q1 represents the lowest-achieving districts and Q4 represents the highest-achieving districts. Mean K–3 per-pupil expenditures supported by the RBG3 law and mean teacher implementation capacity are reported for each quartile. Standard deviations are reported in parentheses below these means. Each cell corresponds to a separate regression. Test scores are expressed in SEDA CS SDs. Standard errors are reported in parentheses. *p<0.10; **p<0.05; ***p<0.01.*

The comparatively small effects observed in Q4 districts may reflect a different dynamic. While overall per-pupil expenditures supported by the RBG3 law were similar across quartiles, Q4 districts reported the highest levels of teacher capacity to implement the policy and allocated a larger share of instructional-support expenditures directly through districts rather than through ISDs. Despite these potentially favorable implementation conditions, students in Q4 districts had the highest achievement levels prior to the law and may therefore have received fewer interventions under the RBG3 framework. As a result, the scope for additional achievement gains may have been more limited than in lower-achieving districts.

Table A4 contains estimated effects by student race/ethnicity[x] in columns (2)-(4) and economically disadvantaged status in columns (5)-(7). The analysis by race/ethnicity is limited to the 25 Michigan districts that have subgroup scores for all grades and years. For this subset of districts, estimated third-grade effects are larger than those in the full sample, with increases of 14.5% and 11.6% of an SD for White and Black students, respectively. For White students, point estimates decline in each subsequent grade, similar to the pattern observed in the overall sample. Estimated effects for Black students remain statistically significant through fifth grade, including a 14.9% of an SD increase in fourth grade and an 8.2% increase in fifth grade. Differences between White and Black students, however, are not statistically significant in any grade.

[x] Our analysis is limited to White and Black students due to insufficient subgroup data from other racial/ethnic groups.

We next examine heterogeneity by economically disadvantaged status. Subgroup scores by economically disadvantaged status are available for a substantially larger set of districts, with 131 Michigan districts included in this analysis. Estimates for both economically disadvantaged and non-economically disadvantaged students follow a similar pattern, declining from approximately 11% of an SD in third grade to 4% of an SD in fifth grade. We again find no evidence of differential effects across these two groups.

Estimated effect by student gender are reported in Table A5, along with the difference in coefficients between the two groups in column (4). In third grade, we find similar improvements in ELA achievement for male and female students, with increases of 8.9% and 8.7% of an SD, respectively. Differences across groups emerge in later grades, however. In fourth grade, male students experienced an 8.7% of an SD increase, while the effect for female students was close to zero. These results suggest that the policy may have contributed to narrowing gender gaps in reading achievement.

To analyze relationships between implementation constructs and our district-level outcome variables, we aggregate each construct to district-level means and then group districts into terciles. We then estimate policy effects separately by tercile for third through fifth grade reading achievement using the full set of untreated control districts from the main analyses as the comparison group for each tercile, as survey data are only available for Michigan districts.

Unlike Table 4, we focus on an array of constructs that capture several dimensions of implementation that may plausibly shape the effectiveness of the reform. First, educator perceptions of the law may influence implementation fidelity and educator engagement with the policy. We therefore examine both principal and teacher perceptions of the RBG3 law, as school leaders may shape schoolwide implementation priorities while teachers are responsible for the day-to-day delivery of literacy instruction and interventions. Second, as teachers were the primary personnel responsible for implementing many of the law's supports, we examine measures of teacher capacity to implement the law and teachers' reported receipt of literacy coaching. These constructs capture the extent to which educators reported receiving the supports and resources necessary to carry out the policy as intended. Finally, because no single measure is likely to fully capture implementation intensity, we create a composite "Average Teacher Construct" that combines responses related to the receipt and perceived effectiveness of literacy coaching and literacy professional development, use of evidence-based instructional practices, perceptions of the RBG3 law, and capacity to implement the law. Principal- and superintendent-based measures are unavailable for districts without administrator survey responses, resulting in smaller samples for those constructs.Figure A3 presents estimated effects of the law on third-grade achievement by tercile for a selected set of implementation constructs, while Appendix Table A6 reports estimates across the full set of constructs for grades 3 through 5. Estimated effects are positive and statistically significant across all constructs and terciles, showing that the policy generated achievement gains across a range of implementation settings. At the same time, estimated effects are generally similar across terciles, with confidence intervals substantially overlapping in all individual and composite constructs. This pattern is also generally consistent across the additional constructs and grades shown in Table A6. As such, we find little evidence that district-level differences as measured by these implementation dimensions produced different effects. However, this should not be interpreted as evidence that implementation was

unimportant to the law's success. Rather, these results may reflect the fact that district-level averages mask meaningful variation in implementation at the school or classroom level. Further, because the law consisted of numerous complementary supports, variation in any single component may not fully capture broader differences in implementation quality across districts.

5.3 Comparison to Mississippi and South Carolina

We next compare Michigan's results to early literacy reforms implemented in Mississippi and South Carolina. These policies can be evaluated using the same empirical approach and data, providing useful benchmarks for contextualizing Michigan's results. While all three states adopted reforms aimed at improving literacy proficiency, the policies differed in their scope, retention provisions, the types of interventions and supports provided, and baseline levels of student achievement.

Mississippi's third-grade reading levels were over one-quarter of an SD below the national average in the 2012-13 school year. In response, Mississippi passed the Literacy-Based Promotion Act, implementing the policy beginning in the 2014-15 academic year to "improve kindergarten and first- through third-grade public school students' reading skills," with the goal of all third-grade students reading at grade level (Mississippi Legislature, 2013a). Starting in 2014-15, third-grade students scoring in the lowest achievement level on the state's reading assessment were eligible for retention, affecting around 15% of students (Mississippi Department of Education 2015a). Beginning in 2018-19, the retention threshold expanded to include students in the lowest two achievement levels, with approximately 25% of students falling below this threshold (Mississippi Department of Education, 2019). Beyond retention, the policy also included statewide teacher training, literacy coaches for low-performing schools, early screening, and targeted interventions for struggling readers (Spencer, 2024).

South Carolina was similarly around one-quarter of an SD below the national average prior to passing the Read to Succeed Act in June 2014. Unlike Michigan's and Mississippi's policies that focused entirely on K-3, South Carolina's act was implemented as a broader K-12 literacy reform, though many of its interventions targeted students in K-3 (South Carolina General Assembly, 2014). The policy was phased in gradually over several years, beginning in 2014-15 with requirements that schools administer early literacy and readiness assessments. In the subsequent year, districts were additionally required to provide interventions for struggling readers that aligned with the state's reading plan, which also included literacy coaching, teacher training requirements, and summer reading supports (South Carolina General Assembly, 2014). The retention component took effect in 2017-18 and required retention for third-grade students in the lowest achievement level of the state's reading assessment, impacting approximately one-quarter of students (South Carolina Department of Education, 2014). As both policies primarily targeted K-3 students and began implementation during the 2014-15 school year, grades three through seven are treated grades in the achievement analysis.

5.3.1 Achievement Effects Across States

Table 5 presents estimated effects of Mississippi's and South Carolina's early literacy policies on ELA and math achievement. Mississippi experienced large and statistically significant gains in both subjects across all grades. For ELA, estimates are largest in third grade at 21.9% of an SD and decline monotonically to 7.5% of an SD in eighth grade, consistent with prior work finding large achievement

gains from the Literacy-Based Promotion Act (Spencer, 2024; Mahmud, 2025). Similar to Michigan, Figure 4 shows that gains increased as cohorts were exposed to the policy longer, and that cohorts exposed in third grade continue to exhibit positive effects in later grades. Mississippi also experienced sizable positive spillovers onto math achievement, with effect estimates ranging from 11.9% to 23.3% of an SD across grades. The magnitude of these effects is comparable to those on ELA scores in early grades, but decline less in later grades.

Table 5. Effects in Mississippi and South Carolina

(1)	(2)	(3)	(4)	(5)
	Mississippi		South Carolina	
Grade	ELA	Math	ELA	Math
3rd	0.219***	0.233***	0.025	-0.009
	(0.0397)	(0.0350)	(0.0156)	(0.0175)
4th	0.188***	0.165***	-0.028*	-0.067***
	(0.0188)	(0.0386)	(0.0155)	(0.0192)
5th	0.163***	0.119***	-0.010	-0.041**
	(0.0238)	(0.0317)	(0.0180)	(0.0199)
6th	0.101***	0.147***	0.035**	0.007
	(0.0243)	(0.0399)	(0.0156)	(0.0152)
7th	0.082**	0.167***	0.038**	0.027
	(0.0405)	(0.0454)	(0.0173)	(0.0173)
8th	0.075**	0.126***	0.031**	0.023
	(0.0322)	(0.0315)	(0.0157)	(0.0227)
Treated Districts	21	21	72	72
Untreated Districts	755	755	755	755

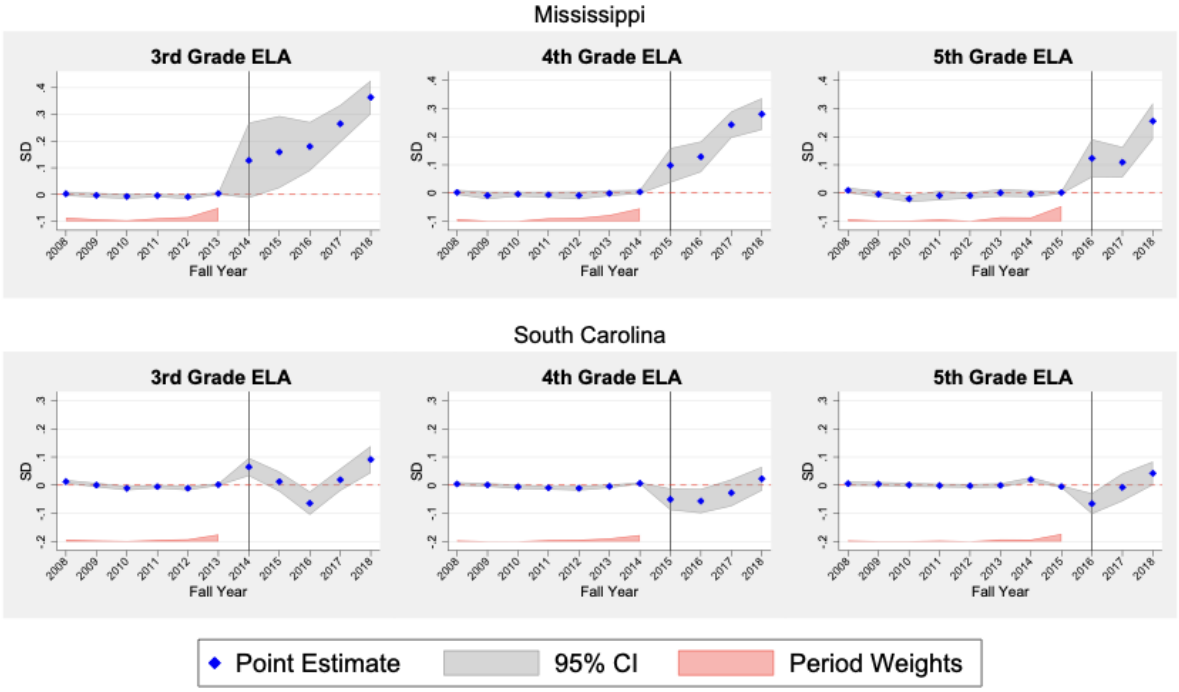
*Note: This table reports estimated effects of Mississippi's (columns (2) and (3)) and South Carolina's early literacy laws on ELA and math (columns (4) and (5)) test scores, by grade. Each cell corresponds to a separate regression. Test scores are expressed in SEDA CS SDs. The double horizontal line indicates grades with cohorts directly exposed to the policy. Standard errors are reported in parentheses. *p<0.10; **p<0.05; ***p<0.01.*

In contrast, South Carolina's policy produced more mixed effects across grades and cohorts. Figure 3 shows meaningful differences across cohorts that align with the policy's gradual phase-in. Cohorts exposed only to the earlier intervention and support components experienced modest declines in achievement, while those exposed to the retention provision exhibit positive estimates. There are also significant estimates in grades six through eight (3-4% of an SD) that likely reflect the broader K-12 scope of South Carolina's law, though these become statistically insignificant when including covariates (see Table A7). Unlike Mississippi, South Carolina experienced some declines in math achievement, suggesting that the increased emphasis on literacy instruction may have partially displaced time or resources from math in this context.

5.3.2 Explaining Cross State Differences

Table 6 summarizes several key features of the three states' literacy reforms that may help explain the variation in estimated achievement effects. While all three implemented comprehensive literacy policies centered on improving reading proficiency, they differed in overall per-pupil educational expenditure levels and growth, early literacy-specific funding, intervention intensity, retention provisions, and implementation.

Figure 3. Event Studies, 3rd-5th Grade ELA Achievement in Mississippi and South Carolina



Note: This figure plots SDID event study estimates of the effects of Mississippi's and South Carolina's early literacy policies on ELA test scores for grades 3 through 5. Test scores are expressed in SEDA CS SDs. Each point represents the estimated effect in a given year relative to the synthetic pre-policy baseline, with gray shaded regions denoting 95% confidence intervals. The vertical line indicates the first year in which a cohort was exposed to the policy, and the shaded red area displays the period weights used to construct the synthetic baseline.

First, overall educational spending and early literacy-specific investments provide one possible explanation for differences in policy effectiveness. As shown in the top graph of Figure 4, Michigan had the highest pre-law K-12 annual per-pupil expenditures at nearly $12,000, followed closely by South Carolina ($11,000), with Mississippi farther below ($9,000). Changes in expenditures after law passage followed a similar pattern, with South Carolina and Michigan increasing by approximately 4% per year compared to 2.4% in Mississippi. Even if returns to increased per-pupil spending are larger at lower levels, these figures reflect broader K-12 spending and therefore are unlikely to fully explain Mississippi's substantially larger achievement gains. More relevant may be the annual K-3 per-pupil funding directly tied to these laws. As shown in the bottom graph of Figure 4, South Carolina allocated the most at over $200 per pupil annually, followed by Mississippi at approximately $120 and Michigan at less than $100. While Mississippi's K–3 appropriations exceeded Michigan's, they remained well below South Carolina's, suggesting that differences in early literacy funding do not explain the magnitude of Mississippi's achievement gains.

Table 6. Comparison of Early Literacy Reform Components

	Michigan	Mississippi	South Carolina
Primary Grade Focus	• K-3 *(MI Leg., 2016)*	• K-3 *(MS Leg., 2013)*	• K-12 (major K-3 interventions) *(SC GA, 2014)*
Funding	• $12,000 in pre-law K-12 per-pupil expenditures • 3.8% annual increase in K-12 per-pupil expenditures post-law *(Strunk et al., 2021)*	• $9,000 in pre-law K-12 per-pupil expenditures • 2.4% annual increase in K-12 per-pupil expenditures post-law *(MS DOE, 2016a)*	• $11,000 in pre-law K-12 per-pupil expenditures • 4.0% annual increase in K-12 per-pupil expenditures post-law • Annual K-3 per-pupil funding unclear
Screening/Assessment	• K-3 screening 3x per year *(MI Leg., 2016)*	• K-3 screening	• PK-K screening within first 45 days of school year *(South Carolina General Assembly, 2014)*
Supports for Students with Reading Deficiencies	• IRIPs • Daily small-group or 1-on-1 instruction • Summer reading camps • ~33% of students *(MI Leg,, 2016; Strunk et al., 2022)*	• Intensive reading instruction • Transitional classes • ~15% of students (2014-15-2017-18); ~25% of students (2018-19) *(MS Leg., 2013; MS DOE, 2015a; 2019)*	• 30+ minutes of daily intensive in-class and supplemental intervention in small groups for PK-3 students • ~25% of students *(SC GA, 2014; SC DOE, 2014)*
Retention Policy?	• Not during study period *(MI Leg., 2016)*	• Yes, beginning in 2014-15 • Good cause exemptions • ~8% of students (2014-15-2017-18); ~15% of students (2018-19) *MS DOE 2015a; 2019)*	• Yes, beginning in 2017-18 • Good cause exemptions • ~25% of students (still received supports if received exemption) *(SC GA, 2013)*
Supports for Retained Students		• 90+ minutes of daily intensive instructional services in small group setting • Assigned high-performing teacher • Extended school day/week/year • Summer reading camp *(MS Leg., 2013)*	• 90+ minutes of daily intensive instructional services in small group setting • Assigned high-performing teacher • Summer reading camp *(SC GA, 2014; SC DOE, 2014; RMC Research Corporation, 2017)*
Parent Supports	• Parental notification • Read-at-home plans *(MI Leg., 2016)*	• Parental notification • Home reading strategies • Read-at-home plans *(MS Leg., 2013)*	• Parental notification • Offered supplemental tutoring outside of school day *(SC GA, 2014)*
Literacy Coaching	• 1+ literacy coach for each ISD • Additional literacy coaches hired by districts and schools *(Strunk et al., 2021)*	• Literacy coaches for low-performing schools ("Literacy Target Schools") *(MS DOE, 2015b)*	• Literacy specialists to provide support to coaches in districts with greatest need *(RMC Research Corporation, 2017).*
Teacher Training	• Targeted literacy professional development • Essential Instructional Practices in Early Literacy • Differentiated instruction *(MI Leg., 2016; Strunk et al., 2021)*	• Statewide science-of reading training • LETRS training *(MS DOE, 2015b; Spencer, 2024)*	• Required completion of Read to Succeed endorsement by 2020 *(RMC Research Corporation, 2017)*

Implementation Evidence	• 7 full-time employees to oversee implementation • 75% of K-3 teachers received some literacy professional development • Insufficient funding and time for implementation • Interventions not fully delivered to all qualifying students *(Strunk et al., 2021; Westall et al., 2023; Kilbride et al., 2024)*	• Statewide implementation guide • All Literacy Target Schools received literacy coaches • 10,000 K-3 teachers/principals completed LETRS training by end of 2014-15 *(MS DOE, 2015b; 2016b)*	• Less than 3 months between enactment and implementation • Office of Early Learning and Literacy + Read to Succeed Team fully formed by second year • Additional literacy specialists added *(RMC Research Corporation, 2017)*

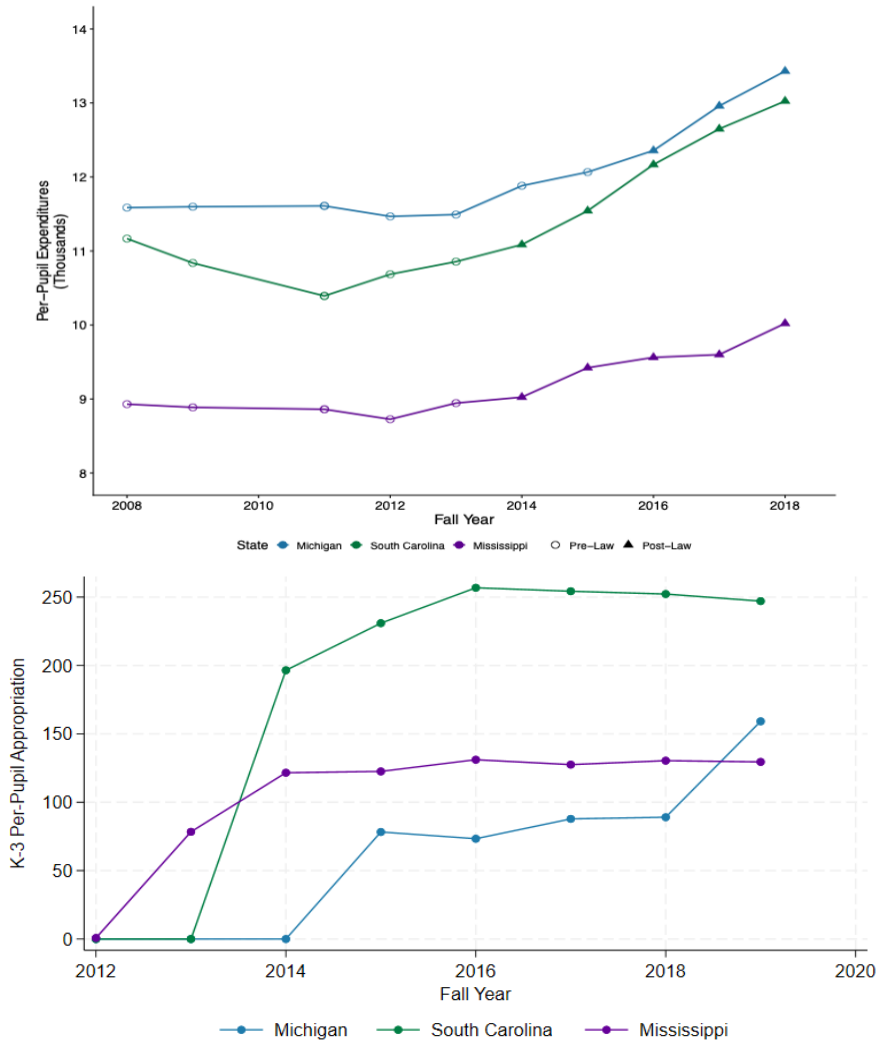
Notes: This table summarizes major components of statewide literacy reforms in Michigan, Mississippi, and South Carolina using legislative documents, state reports, and prior studies. K-12 per-pupil expenditure figures are derived from National Public Education Financial Survey data. Figures reported for funding, interventions, and retention policies are approximate. Citation abbreviations are used for conciseness: MI = Michigan; MS = Mississippi; SC = South Carolina; DOE = Department of Education; Leg. = Legislature; GA = General Assembly.

The three also differed in the intensity and structure of literacy interventions and supports provided to students. While each policy included common components such as literacy screening and intensive reading instruction for struggling readers, the scale and targeting of these supports varied considerably. Michigan provided Tier II supports to approximately 33% of K-3 students, larger than the share subject to similar interventions in Mississippi and South Carolina. However, Mississippi had a retention policy throughout the entire analysis period that provided additional support to retained students, while South Carolina implemented this component only during the final two years of the study period and Michigan's retention provision had not yet taken effect.

These differences are reflected in the retention event studies shown in Figure 5. While Michigan and South Carolina experienced no change in third-grade retention rates following implementation of their literacy reforms, Mississippi experienced a substantial immediate increase in retention. Table A8 similarly shows increases in retention across multiple early grades in Mississippi, including a roughly 1.2 percentage point increase in third-grade retention overall and a 2.5 percentage point increase (58%) in the first post-policy period observed. However, Mississippi's largest achievement gains emerge during later years when retention effects had begun to fade, suggesting that retention policies alone are unlikely to fully explain the cross-state variation in estimated achievement effects.

The states also varied in the support provided to teachers. Michigan adopted a broader coaching model that guaranteed at least one literacy coach for every intermediate school district (ISD), while Mississippi concentrated literacy coaches in low-performing schools, potentially allowing for more intensive support in schools serving students with the greatest literacy needs (Mississippi Department of Education, 2015b; Strunk et al., 2021). South Carolina instead relied on literacy specialists who worked through a larger statewide coaching network, making support for classroom teachers somewhat less direct (RMC Research Corporation, 2017). These approaches were paired with different teacher training models, ranging from Michigan's literacy-focused professional development to South Carolina's required Read to Succeed endorsement and Mississippi's large-scale Language Essentials for Teachers of Reading and Spelling (LETRS) science-of-reading training.

Figure 4. Per-Pupil Expenditures and Appropriations by State

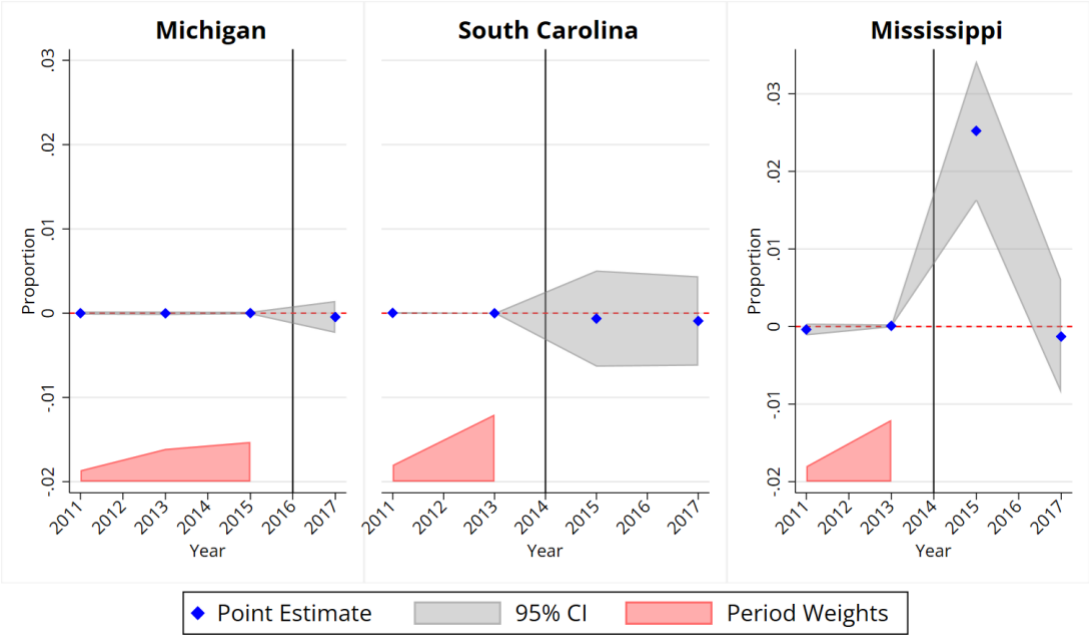


Notes: The top panel of this figure plots statewide K–12 per-pupil expenditures from 2008 to 2018 for Michigan, Mississippi, and South Carolina using National Public Education Financial Survey data. Expenditures are reported in thousands of dollars. Filled markers indicate years following implementation of each state's literacy reform. The bottom panel shows appropriation funds from state budgets per K-3 student devoted to support each states' early literacy policy. As legislation was enacted prior to implementation, states were able to appropriate funds before the law's effective date.

Finally, variation in state capacity, policy coordination, and rollout may have also contributed to differential policy effectiveness. Several studies in Michigan have documented both substantial statewide literacy efforts and challenges, including uneven delivery of interventions and resource disparities across districts, though these issues improved over time (Strunk et al., 2021; Westall et al., 2023; Kilbride et al., 2024). South Carolina also faced challenges during the early years of the Read to Succeed Act, in part because some policy components were introduced rapidly while others were phased in gradually over several years (RMC Research Corporation, 2017). In contrast, Mississippi adopted a more centralized literacy strategy that included formal implementation guides and large-scale teacher training completed within one year of the reform (Mississippi Department of Education, 2015b; 2016b). Mississippi also implemented several additional educational reforms during this

period, including teacher pay increases, performance-based incentives, expanded school accountability systems, growth in early childhood education access, and charter school expansion (Mississippi Department of Education, 2013; Mississippi Legislature, 2013b; Joint Legislative Committee on Performance Evaluation and Expenditure Review, 2020). Thus, the estimated effects likely capture the combined influence of the Literacy-Based Promotion Act and this broader reform environment. This more coordinated rollout and broader suite of reforms may help explain why Mississippi experienced larger and more persistent achievement gains despite lower overall educational spending levels.

Figure 5. Event Studies, 3rd Grade Retention



Note: This figure plots SDID event study estimates of the effects of early literacy policies on third-grade retention rates in Michigan, Mississippi, and South Carolina. Each point represents the estimated effect in a given year relative to the synthetic pre-policy baseline, with gray shaded regions denoting 95% confidence intervals. The vertical line indicates the first year in which a cohort was exposed to the policy. Retention rates are expressed as the share of students retained in third grade.

Collectively, these comparisons suggest that the effectiveness of statewide literacy reforms likely depends on the interaction of multiple policy dimensions rather than any single component alone. Funding levels, retention requirements, intervention intensity, teacher supports, policy coordination, and complementary educational reforms all appear relevant for understanding why estimated effects varied across Michigan, Mississippi, and South Carolina. Further, these results highlight the importance of considering both policy design and state context when evaluating comprehensive statewide literacy reforms.

6. DISCUSSION

This paper examines the overall effects of three comprehensive statewide early literacy policies on student achievement and grade retention. We find that Michigan's RBG3 law generated positive and statistically significant improvements in ELA achievement, particularly in third grade. Event study

estimates further indicate that cohorts exposed to the policy for longer periods experienced larger gains in third-grade ELA achievement. By the end of our sample period, the estimated third-grade gains approach roughly half of a grade level of learning, comparable to the magnitude of national post-pandemic reading losses documented in recent work (Dewey et al., 2025).[xi] Positive effects also persisted as exposed cohorts advanced into fourth and fifth grade. Estimated effects were smallest for the highest-performing districts, suggesting the policy may have helped narrow pre-existing achievement gaps. At the same time, we find no evidence of spillover effects on math achievement or increases in retention.

These patterns suggest that comprehensive literacy reforms may generate effects that both accumulate with longer exposure and extend beyond the initial year of treatment. One possible explanation is that earlier exposure allows students to benefit from literacy interventions during foundational stages of skill development. Later cohorts may also have benefited from more intensive implementation as districts expanded literacy expenditures and intervention efforts during the first several years of the policy. The persistence of positive effects in fourth and fifth grade suggests that the policy generated lasting improvements in ELA achievement for exposed cohorts. Comprehensive literacy reforms may therefore require several years before their full effects become apparent.

Comparisons across states further highlight how differences in policy design, implementation, and broader educational contexts may shape the effectiveness of early literacy reforms. While Michigan experienced meaningful improvements in ELA achievement, estimated effects were substantially larger in Mississippi and generally smaller in South Carolina. These differences coincide with meaningful variation in the structure, rollout, and environment surrounding each reform: Mississippi combined a more centralized implementation strategy with a broader set of complementary education reforms, South Carolina experienced a more gradual rollout and greater early implementation challenges, and Michigan generally fell between these two approaches. Together, these findings contribute to a growing body of research suggesting that achievement gains from early literacy reforms are not driven solely by grade retention, but also by the broader design, supports, and implementation environments within these policies (Berne et al., 2025).

Understanding the cost-effectiveness of statewide literacy reforms is also important for policy considerations. A cost analysis of Michigan's RBG3 law by Knight et al. (2026) estimates annual per-student costs of approximately $2,756 in grades K–3, though much of this reflects the reallocation of existing instructional resources rather than new spending. Using our smallest significant estimate (2.7% of an SD in fifth grade), the implied gain is approximately 0.98% of an SD per $1,000 if the full annual K–3 per-student cost is assigned to the cohort's initial year of exposure. By comparison, Jackson and Mackevicius (2024) estimate that a $1,000 annual increase in per-pupil spending over four years increases achievement by 3.2% of an SD, or 0.8% of an SD per $1,000 in a single year. While not a formal cost-effectiveness analysis, the estimated achievement effects identified in this study compare favorably in magnitude relative to existing education spending benchmarks.

[xi] The final third-grade estimate was 15% of an SD. As one SD corresponds to approximately three grade levels, this estimate is equivalent to roughly 0.45 grade levels.

Overall, these findings highlight the potential for comprehensive statewide literacy reforms to improve reading achievement at scale, while also demonstrating that policy effectiveness depends heavily on implementation and broader educational context. As more states continue to adopt and refine early literacy initiatives, future research should more closely examine how differences in implementation, intervention intensity, retention policies, and complementary literacy efforts contribute to policy effectiveness across states, districts, and student populations. Additionally, while this paper focuses on student achievement and grade retention, future work could examine whether early literacy reforms influence broader student outcomes, including student mobility, special education identification, and absenteeism.

DECLARATIONS

EPIC Disclaimer: The Education Policy Innovation Collaborative (EPIC) at Michigan State University is an independent, non-partisan research center that operates as the strategic research partner to the Michigan Department of Education (MDE) and the Center for Educational Performance and Information (CEPI). EPIC conducts original research using a variety of methods that include advanced statistical modeling, representative surveys, interviews, and case study approaches. This research result used data collected and maintained by EPIC. Results, information, and opinions solely represent the author(s) and are not endorsed by, nor reflect the views or positions of, grantors, MDE and CEPI, or any employee thereof. All errors are our own.

Funding: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305H190004 to Michigan State University. The opinions expressed are those of the authors and do not represent views of the Institute or the U.S. Department of Education.

Data Availability: Public data sources used in this study are cited in the manuscript. Survey data are not publicly available due to data-use restrictions.

Acknowledgments: We thank Amy Cummings and Soo Jeong Lee for their valuable research assistance. We are also indebted to Christina Weiland, Brian Jacob, Nicole Wagner, Jordy Berne, Tanya Wright, David Knight, and participants at the Southern Economics Association and AEFP conferences for their helpful comments.

REFERENCES

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, *105*(490), 493–505. https://doi.org/10.1198/jasa.2009.ap08746

Arkhangelsky, D., Athey, S., Hirshberg, D., Imbens, G., & Wager, S. (2021). Synthetic difference-in-differences. *American Economic Review*, *111*(12), 4088–4118. https://doi.org/10.1257/aer.20190159

Berne, J., Weiland, C., Jacob, B., & Strunk, K. (2025). The impacts of grade retention with minimal retention. EdWorkingPaper No. 25-1188. *Annenberg Institute for School Reform at Brown University.* https://doi.org/10.26300/tqrb-yh12

Chen, F., & Harris, D. (2023). The market-level effects of charter schools on student outcomes: A national analysis of school districts. *Journal of Public Economics*, *228*, 105015. https://doi.org/10.1016/j.jpubeco.2023.105015

Clarke, D., Pailañir, D., Athey, S., & Imbens, G. (2023). Synthetic difference in differences estimation. *arXiv Preprint arXiv:2301.11859*. https://doi.org/10.48550/arXiv.2301.11859

Cummings, A. (2024). Reading between the lines: Three essays on the development and implementation of Michigan's Read by Grade Three Law [Ph.D., Michigan State University]. In *ProQuest Dissertations and Theses* (3092202510). Dissertations & Theses @ Big Ten Academic Alliance; ProQuest Dissertations & Theses Global.

Cummings, A., Strunk, K., & De Voto, C. (2021). "A lot of states were doing it": The development of Michigan's Read by Grade Three Law. *Journal of Educational Change*, *24*(1), 107–132. https://doi.org/10.1007/s10833-021-09438-y

Cunningham, A., & Stanovich, K. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, *33*(6), 934–945. https://doi.org/10.1037/0012-1649.33.6.934

Dewey, D., Fahle, E., Kane, T., Reardon, S., & Staiger, D. (2025). Education Recovery Scorecard. February 2025: Pivoting from recovery to long-term reform. *Center for Education Policy Research, Harvard University, and The Educational Project Opportunity at Stanford University.* https://educationscore.wpenginepowered.com/wp-content/uploads/2025/02/Pivoting-from-Pandemic-Recovery-to-Long-Term-Reform-A-District-Level-Analysis.pdf (accessed 5 June 2026).

Education Policy Innovation Collaborative. (2019). Estimated Read by Grade 3 retention rates. A report prepared for the Michigan Department of Education. https://epicedpolicy.org/wp-content/uploads/2019/06/RBG3_estimated_cutpoint_analysis.pdf (accessed 5 June 2026).

ExcelinEd. (2024). Comprehensive Early Literacy Policy. A State-by-State Analysis of Fundamental Principles. https://excelined.org/wp-content/uploads/2021/10/ExcelinEd_PolicyToolkit_EarlyLiteracy_StatebyStateAnalysis_2021.pdf (accessed 5 June 2026).

Fahle, E., Saliba, J., Kalogrides, D., Shear, B., Reardon, S., & Ho, A. (2024). Stanford Education Data Archive: Technical documentation (Version 5.0).

https://stacks.stanford.edu/file/druid:cs829jn7849/SEDA_documentation_v5.0.pdf (accessed 5 June 2026).

Fiester, L., & Smith, R. (2010). Early warning! Why reading by the end of third grade matters (ED509795). ERIC. https://files.eric.ed.gov/fulltext/ED509795.pdf (accessed 5 June 2026).

Figlio, D., & Özek, U. (2020). An extra year to learn English? Early grade retention and the human capital development of English learners. *Journal of Public Economics*, *186*(104184), 1–16. https://doi.org/10.1016/j.jpubeco.2020.104184

Gamse, B., Jacob, R., Horst, M., Boulay, B., & Unlu, F. (2008). Reading First impact study. Final report. NCEE 2009-4038. *National Center for Education Evaluation and Regional Assistance*. https://files.eric.ed.gov/fulltext/ED503344.pdf (accessed 5 June 2026).

Greene, J., & Winters, M. (2007). Revisiting grade retention: An evaluation of Florida's test-based promotion policy. *Education Finance and Policy*, *2*(4), 319–340. https://doi.org/10.1162/edfp.2007.2.4.319

Hernandez, D. (2011). Double jeopardy: How third-grade reading skills and poverty influence high school graduation (ED518818). ERIC. https://files.eric.ed.gov/fulltext/ED518818.pdf (accessed 5 June 2026).

Jackson, C., & Mackevicius, C. (2024). What impacts can we expect from school spending policy? Evidence from evaluations in the United States. *American Economic Journal: Applied Economics*, *16*(1), 412–446. https://doi.org/10.1257/app.20220279

Jacob, B., & Lefgren, L. (2004). Remedial education and student achievement: A regression-discontinuity analysis. *The Review of Economics and Statistics*, *86*(1), 226–244. https://doi.org/10.3386/w8918

Jang, H., & Reardon, S. (2019). States as sites of educational (in)equality: State contexts and the socioeconomic achievement gradient. *AERA Open*, *5*(3), 2332858419872459. https://doi.org/10.1177/2332858419872459

Joint Legislative Committee on Performance Evaluation and Expenditure Review. (2020). FY 2020 annual report: Analysis of funding for Mississippi charter schools and the Charter School Authorizer Board. https://www.peer.ms.gov/sites/default/files/peer_publications/rpt645.pdf (accessed 5 June 2026).

Kilbride, T., Hopkins, B., Strunk, K., & Yu, D. (2022). Michigan's 2020-21 and 2021-22 benchmark assessments. *Education Policy Innovation Collaborative.* https://epicedpolicy.org/wp-content/uploads/2022/10/COVID_Benchmark_Assessments_Report_Oct2022.pdf (accessed 5 June 2026).

Kilbride, T., Nadal-Fernandez, J., & Walker, S. (2024). Retention and exemption decisions under the Read by Grade Three Law. *Education Policy Innovation Collaborative*. https://epicedpolicy.org/wp-content/uploads/2024/06/EPIC_RetentionDecisionReport_June2024.pdf (accessed 5 June 2026).

Knight, D., Cummings, A., Strunk, K., Westall, J., Wright, T., Mavrogordato, M., & Barlett, C. (2026). Cost analysis of a statewide early literacy reform: Examining implementation and resource allocation in Michigan's Read by Grade Three Law. Working Paper.

Kuhfeld, M., Domina, T., & Hanselman, P. (2019). Validating the SEDA measures of district educational opportunities via a common assessment. *AERA Open*, *5*(2), 2332858419858324. https://doi.org/10.1177/2332858419858324

Lane, M., & Conlon, G. (2016). The impact of literacy, numeracy and computer skills on earnings and employment outcomes. OECD Education Working Papers, No. 129. OECD Publishing, Paris. https://doi.org/10.1787/5jm2cv4t4gzs-en

Mahmud, S. (2025). Does state-mandate third-grade retention policy improve achievement? Evidence from a staggered-adoption difference-in-differences design. EdWorkingPaper No. 25-1352. *Annenberg Institute for School Reform at Brown University.* https://doi.org/10.26300/b0em-9a84

Mariano, L., & Martorell, P. (2013). The academic effects of summer instruction and retention in New York City. *Educational Evaluation and Policy Analysis*, *35*(1), 96–117. https://doi.org/10.3102/0162373712454327

Michigan Legislature. (2016). House Bill 4822 (Public Act 306 of 2016). https://www.legislature.mi.gov/documents/2015-2016/publicact/pdf/2016-PA-0306.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2013). Mississippi Public School Accountability Standards, 2012. https://www.mdek12.org/sites/default/files/Offices/MDE/OA/Accreditation/Commission%20on%20School%20Accreditation/Accreditation%20Accountability/2012-standards-8-8-13.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2015a). 2015-2015 Literacy-Based Promotion Act Annual Report of Performance and Student Retention. https://www.mdek12.org/sites/default/files/Offices/MDE/SSE/2014-2015-lbpa-report---all-districts_20160808091719_205160.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2015b). 2014-2015 Superintendent's Annual Report. https://mdek12.org/sites/default/files/Offices/MDE/SSE/2016-AnnRep/mde_super_report_14_15_digital.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2016a). 2015-2016 Superintendent's Annual Report. https://mdek12.org/wp-content/uploads/sites/29/2026/03/mde_superintendent_report_16_web_0_v1.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2016b). Mississippi Literacy-Based Promotion Act Implementation Guide (Rev. Ed.). https://www.mdek12.org/sites/default/files/documents/OAE/Literacy/ResourcesForAdmin/revised-11-09-16-lbpa-implementation-guide_feb17_20170223084950_172860.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2019). 2018-2019 Literacy-Based Promotion Act Annual Report of Performance and Student Retention. https://mdek12.org/sites/default/files/Offices/MDE/SSE/lbpa_summary_2018.pdf (accessed 5 June 2026).

Mississippi Department of Education. (2023). 2022-2023 Mississippi Academic Assessment Program (MAAP) Results: Executive Summary. https://www.mdek12.org/sites/default/files/Offices/MDE/OEA/OPR/2023/maap_2023_results_executive_summary_v1_final_8.14.pdf (accessed 5 June 2026).

Mississippi Legislature. (2013a). Senate Bill No. 2347, Literacy-Based Promotion Act. https://billstatus.ls.state.ms.us/documents/2013/pdf/SB/2300-2399/SB2347SG.pdf (accessed 5 June 2026).

Mississippi Legislature. (2013b). Senate Bill No. 2395. Early Learning Collaborative Act of 2013. https://billstatus.ls.state.ms.us/documents/2013/pdf/SB/2300-2399/SB2395SG.pdf (accessed 5 June 2026).

National Assessment of Educational Progress. (2024). NAEP report card: 2024 NAEP reading assessment. https://www.nationsreportcard.gov/reports/reading/2024/g4_8/?grade=4 (accessed 5 June 2026).

National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction [Reports of the subgroups]. *National Institute of Child Health and Human Development.* https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf (accessed 5 June 2026).

Özek, U. (2015). Hold back to move forward? Early grade retention and student misbehavior. *Education Finance and Policy*, *10*(3), 350–377. https://doi.org/10.1162/EDFP_a_00166

Özek, U. (2025). Does one plus one always equal two? Examining complementarities in educational interventions. *Journal of Policy Analysis and Management*, *44*(1), 45–72. https://doi.org/10.1002/pam.22594

Reardon, S., Papay, J., Kilbride, T., Strunk, K., Cowen, J., An, L., & Donohue, K. (2019). Can repeated aggregate cross-sectional data be used to measure average student learning rates? A validation study of learning rate measures in the Stanford Education Data Archive. CEPA Working Paper No. 19-08. *Stanford Center for Education Policy Analysis.* https://cepa.stanford.edu/sites/default/files/wp19-08-v201911.pdf (accessed 5 June 2026).

RMC Research Corporation. (2017). South Carolina Read to Succeed: An inside look. https://ed.sc.gov/data/reports/literacy/scde-literacy-reports/sc-read-to-succeed/sc-read-to-succeed-an-inside-look-2017-evaluation-by-rmc-research-corporation/ (accessed 5 June 2026).

Roderick, M., & Nagaoka, J.. (2016). Retention under Chicago's high-stakes testing program: Helpful, harmful, or harmless? *American Educational Research Journal*, *27*(4), 309–340. https://doi.org/10.3102/01623737027004309

Schwerdt, G., West, M., & Winters, M. (2017). The effects of test-based retention on student outcomes over time: Regression discontinuity evidence from Florida. *Journal of Public Economics*, *152*, 154–169. https://doi.org/10.1016/j.jpubeco.2017.06.004

South Carolina Department of Education. (2014). South Carolina Palmetto Assessment of State Standards (SCPASS). https://www.ed.sc.gov/data/test-scores/state-assessments/scpalmetto-assessment-of-state-standards-pass/2014/ (accessed 5 June 2026).

South Carolina General Assembly. (2014). Senate Bill 516. Read to Succeed Act. https://www.scstatehouse.gov/sess120_2013-2014/bills/516.htm (accessed 5 June 2026).

Slungaard Mumma, K., & Winters, M. (2023). The effect of retention under Mississippi's test-based promotion policy. Working Paper 2023–1. *Wheelock Educational Policy Center.* https://wheelockpolicycenter.org/high-quality-education/ms-read-by-grade-three/ (accessed 5 June 2026).

Sparks, R., Patton, J., & Murdoch, A. (2014). Early reading success and its relationship to reading achievement and reading volume: Replication of "10 years later." *Reading and Writing*, *27*, 189–211. https://doi.org/10.1007/S11145-013-9439-2

Spencer, N (2024). Comprehensive early literacy policy and the "Mississippi Miracle". *Economics of Education Review, 103,* 102598. https://doi.org/10.1016/j.econedurev.2024.102598

Strunk, K., Wright, T., Kilbride, T., Zhu, Q., Cummings, A., Westall, J., Turner, M., & De Voto, C. (2021). Michigan's Read by Grade Three Law: Year one report. *Education Policy Innovation Collaborative.* https://epicedpolicy.org/wp-content/uploads/2021/03/Year_One_RBG3_Report.pdf (accessed 5 June 2026)

Strunk, K., Wright, T., Westall, J., T., Zhu, Q., Kilbride, T., Cummings, A., Utter, A., & Mavrogordato, M. (2022). Michigan's Read by Grade Three Law: Year two report. *Education Policy Innovation Collaborative.* https://epicedpolicy.org/wp-content/uploads/2022/02/RBG3_Rpt_Yr2_Feb2022.pdf (accessed 5 June 2026).

U.S. Congress. (1998). Reading Excellence Act (H.R. 2614). https://www.congress.gov/bill/105th-congress/house-bill/2614 (accessed 5 June 2026).

U.S. Department of Education, Office of Elementary and Secondary Education. (2002). Reading First overview and guidance. Washington, DC. https://files.eric.ed.gov/fulltext/ED498687.pdf (accessed 5 June 2026).

Weiss, S., Stallings, D., & Porter, S. (2018). Is Read to Achieve making the grade? An assessment of North Carolina's elementary reading proficiency initiative (pp. 1–40). *North Carolina State University College of Education.* https://www-data.fi.ncsu.edu/wp-content/uploads/2021/02/01122551/RtA2018Reportv6.pdf (accessed 5 June 2026).

Westall, J., & Cummings, A. (2023). The effects of early literacy policies on student achievement. *Education Policy Innovation Collaborative.* https://files.eric.ed.gov/fulltext/ED672706.pdf (accessed 5 June 2026).

Westall, J., Strunk, K., & Utter, A. (2023). Challenges in Implementing Teacher-Student Assignment Policies: Evidence From Michigan's Read by Grade Three Law. *Education Policy Innovation Collaborative.* https://epicedpolicy.org/wp-content/uploads/2023/06/ChallengesImplementingTeacherAssignment_WorkingPaper_July2023.pdf (accessed 5 June 2026).

Winters, M., & Greene, J. (2012). The medium-run effects of Florida's test-based promotion policy. *Education Finance and Policy*, *7*(3), 305–330. https://doi.org/10.1162/EDFP_a_00069

APPENDIX

Table A1. Michigan Early Literacy Grant Funds, 2016-17 to 2018-19

Allocations	2016-17	2017-18	2018-19
Implementation Costs (MDE Staff)	$1 M	$1 M	$1 M
Michigan Education Corps	$1 M	$2.5 M	$3 M
Early Literacy Professional Learning Grant	$1.9 M	$ 0	$ 0
ISD Literacy Coaches	$3 M	$6 M	$7 M
Assessment Reimbursement	$4.5 M	$4 M	$9.2 M
Additional Instructional Time & Intervention	$17.5 M	$20.9 M	$19.9 M
Summer Reading Program Pilot			$500,000
Multisensory Structured Language Program Pilot			$250,000
Literacy Essentials Professional Learning			$1 M

Note: This table reports annual allocations for Michigan early literacy grant funds under the RBG3 law from 2016–17 through 2018–19. Amounts are based on Michigan Department of Education financial data files.

Table A2. South Carolina and Mississippi SEDA Sample Comparison

(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	South Carolina				Mississippi*			
	All Districts	Districts in Sample	Districts not in Sample	t-test p value	All Districts	Districts in Sample	Districts not in Sample	t-test p value
% Urban	0.064	0.072	0.000	0.001	0.034	0.000	0.040	0.017
% Town/Suburb	0.347	0.354	0.289	0.633	0.410	0.374	0.416	0.638
% Rural	0.589	0.574	0.711	0.324	0.556	0.626	0.543	0.361
% White	0.484	0.502	0.337	0.126	0.431	0.486	0.421	0.351
% Black	0.454	0.432	0.630	0.085	0.538	0.486	0.546	0.409
% Hispanic	0.047	0.050	0.026	0.024	0.020	0.019	0.020	0.905
% Asian	0.011	0.012	0.005	0.028	0.006	0.006	0.006	0.837
% Economically Disadvantaged	0.665	0.649	0.796	0.005	0.673	0.610	0.684	0.118
% FRL	0.647	0.631	0.775	0.010	0.774	0.706	0.786	0.071
% EL	0.032	0.034	0.015	0.014	0.010	0.012	0.010	0.708
% SPED	0.150	0.150	0.150	0.986	0.000	0.000	0.000	–
Average 2008 3rd Grade ELA Score	-0.248	-0.231	-0.386	0.102	-0.260	-0.238	-0.263	0.698
2008 3rd Grade ELA Tested Students	55,041	53,713	1,328	–	39,727	8,082	31,645	–

*Note: This table reports baseline means for a set of variables for all South Carolina (column (2)) and Mississippi (column (6)) districts in SEDA 5.0, those included in the analytic sample (columns (3) and (7)), and those excluded from the analysis (columns (4) and (8)). Columns (5) and (9) report the p-value from a t-test of differences in means between included and excluded states. *Mississippi did not report the percentage of enrolled students classified as receiving special education until fall 2009.*

Table A3. States Without an Early Literacy Policy by Fall 2018

State	Implemented	SEDA Control?	CRDC Control?
Alabama	2019-2020	Yes	Yes
California	2020-2021	No	No
Georgia	2025-2026	Yes	No
Hawaii	–	Yes	Yes
Illinois	–	Yes	Yes
Kansas	2022-2023	No	No
Louisiana	2021-2022	Yes	No
Montana	2024-2025	No	Yes
New Hampshire	–	Yes	Yes
North Dakota	–	No	Yes
Oregon	2023-2024	No	Yes
Pennsylvania	2024-2025	Yes	Yes
Rhode Island	–	No	No
South Dakota	–	No	Yes
Texas	2020-2021	No	No
West Virginia	2023-2024	No	No
Wyoming	2022-2023	No	No

Note: This table lists states without a statewide early literacy policy implemented by Fall 2018 and indicates whether each state is included in the pool of SEDA and CRDC comparisons. States that adopted policies after Fall 2018 are classified based on the first academic year of implementation. Policy implementation dates are based on ExcelinEd (2024), state legislation, and state department of education records.

Table A4. Effect on ELA Achievement by Race/Ethnicity and Economically Disadvantaged Status

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Grade	**White**	**Black**	**Difference**	**Econ. Dis.**	**Not Econ. Dis.**	**Difference**
3rd	0.145*** (0.0367)	0.116*** (0.0417)	0.030 (-0.079,0.139)	0.115*** (0.0172)	0.107*** (0.0193)	0.007 (-0.043,0.058)
4th	0.098** (0.0394)	0.149*** (0.0445)	-0.052 (-0.168,0.065)	0.103*** (0.0202)	0.068*** (0.0186)	0.035 (-0.019,0.089)
5th	0.054 (0.0442)	0.082** (0.0347)	-0.028 (-0.138,0.082)	0.037* (0.0211)	0.037* (0.0192)	-0.000 (-0.056,0.056)
Treated Districts	25	25	25	131	131	131
Untreated Districts	40	40	40	202	202	202

*Note: This table reports estimated effects of the RBG3 law on ELA test scores by race/ethnicity and economically disadvantaged status, by grade. Columns (2) and (3) report estimated effects for White and Black students, respectively, while column (4) reports the difference in coefficients between the two groups, with 95% confidence intervals reported in brackets. Columns (5) and (6) report estimated effects for economically disadvantaged and non-economically disadvantaged students, respectively, while column (7) reports the difference in coefficients between the two groups, with 95% confidence intervals reported in brackets. Each cell corresponds to a separate regression. Test scores are expressed in SEDA CS SDs. Standard errors are reported in parentheses for columns (2), (3), (5), and (6). *p<0.10; **p<0.05; ***p<0.01.*

Table A5. Effect on ELA Achievement by Gender

(1)	(2)	(3)	(4)
Grade	**Male**	**Female**	**Difference**
3rd	0.089***	0.087***	0.002
	(0.0176)	(0.0160)	(-0.044,0.049)
4th	0.087***	0.001	0.085***
	(0.0156)	(0.0168)	(0.040,0.130)
5th	0.023	0.002	0.021
	(0.0202)	(0.0164)	(-0.030,0.072)
Treated Districts	192	192	192
Untreated Districts	223	223	223

*Note: This table reports estimated effects of the RBG3 law on ELA test scores for male (column (2)) and female students (column (3)), by grade. Column (4) reports the difference in coefficients between groups, with 95% confidence intervals reported below in parentheses. Each cell corresponds to a separate regression. Test scores are expressed in SEDA CS SDs. Standard errors are reported in parentheses for columns (2) and (3). *p<0.10; **p<0.05; ***p<0.01.*

Table A6. Effect on ELA Achievement by Implementation Construct Tercile

(1)	(2)	(3)	(4)	(5)
Construct	**Grade**	**T1**	**T2**	**T3**
Panel A: Superintendent Constructs				
Superintendent Belief about Law Effectiveness	3rd	0.100*** (0.0286)	0.117*** (0.0284)	0.075*** (0.0287)
	4th	0.057* (0.0327)	0.072** (0.0323)	0.049 (0.0327)
	5th	0.043 (0.0327)	0.059* (0.0332)	0.054* (0.0328)
Treated Districts		23	24	23
Untreated Districts		726	726	726
Panel B: Principal Constructs				
Principal Perception of Law	3rd	0.092*** (0.0196)	0.093*** (0.0197)	0.102*** (0.0197)
	4th	0.055*** (0.0169)	0.049*** (0.0169)	0.068*** (0.0172)
	5th	0.030 (0.0226)	0.055** (0.0227)	0.019 (0.0222)
Treated Districts		65	64	66
Untreated Districts		726	726	726
Panel C: Teacher Constructs				
Teacher Perception of Law	3rd	0.094*** (0.0165)	0.094*** (0.0170)	0.091*** (0.0171)
	4th	0.048*** (0.0151)	0.059*** (0.0148)	0.067*** (0.0148)
	5th	0.042** (0.0178)	0.044** (0.0187)	0.029 (0.0187)
Teacher Capacity to Implement Law	3rd	0.087*** (0.0165)	0.083*** (0.0170)	0.110*** (0.0170)
	4th	0.063*** (0.0151)	0.053*** (0.0148)	0.067*** (0.0148)
	5th	0.045** (0.0178)	0.050*** (0.0188)	0.019 (0.0187)

Receipt of Literacy Coaching	3rd	0.090*** (0.0165)	0.101*** (0.0170)	0.089*** (0.0171)
	4th	0.063*** (0.0151)	0.058*** (0.0148)	0.053*** (0.0148)
	5th	0.054*** (0.0178)	0.024 (0.0181)	0.036* (0.0187)
Receipt of Professional Development on Evidence-Based Instructional Practices	3rd	0.091*** (0.0171)	0.095*** (0.0165)	0.093*** (0.0170)
	4th	0.065*** (0.0148)	0.046*** (0.0148)	0.062*** (0.0148)
	5th	0.046** (0.0182)	0.027 (0.0181)	0.041** (0.0188)
Receipt of Other Literacy Professional Development	3rd	0.097*** (0.0166)	0.097*** (0.0172)	0.086*** (0.0170)
	4th	0.064*** (0.0151)	0.062*** (0.0157)	0.049*** (0.0148)
	5th	0.047*** (0.0178)	0.046** (0.0193)	0.022 (0.0182)
Average Teacher Construct	3rd	0.086*** (0.0165)	0.107*** (0.0170)	0.087*** (0.0170)
	4th	0.047*** (0.0151)	0.072*** (0.0148)	0.055*** (0.0148)
	5th	0.032* (0.0178)	0.066*** (0.0187)	0.016 (0.0187)
Treated Districts		96	95	95
Untreated Districts		726	726	726

*Note: This table reports estimated effects of the RBG3 law on ELA test scores by terciles of various implementation constructs, by grade. T1 represents the lowest values of each implementation construct, while T3 represents the highest. Each cell corresponds to a separate regression. Sample sizes differ across panels as superintendent- and principal-based measures are unavailable for districts without administrator survey responses. Test scores are expressed in SEDA CS SDs. Standard errors are reported in parentheses. *p<0.10; **p<0.05; ***p<0.01.*

Table A7. Effects in Mississippi and South Carolina, With Covariates

(1)	(2)	(3)	(4)	(5)
	Mississippi		Math	
Grade	ELA	Math	ELA	Math
3rd	0.227***	0.247***	0.017	-0.007
	(0.0289)	(0.0367)	(0.0169)	(0.0212)
4th	0.191***	0.161***	-0.023	-0.065***
	(0.0348)	(0.0397)	(0.0163)	(0.0187)
5th	0.172***	0.118***	-0.020	-0.037**
	(0.0298)	(0.0340)	(0.0188)	(0.0182)
6th	0.106***	0.148***	0.032*	0.010
	(0.0355)	(0.0281)	(0.0188)	(0.0188)
7th	0.110***	0.158***	0.033	0.029
	(0.0405)	(0.0408)	(0.0263)	(0.0267)
8th	0.075	0.117***	0.020	0.026
	(0.0462)	(0.0451)	(0.0238)	(0.0233)
Treated Districts	21	21	72	72
Untreated Districts	726	726	726	726

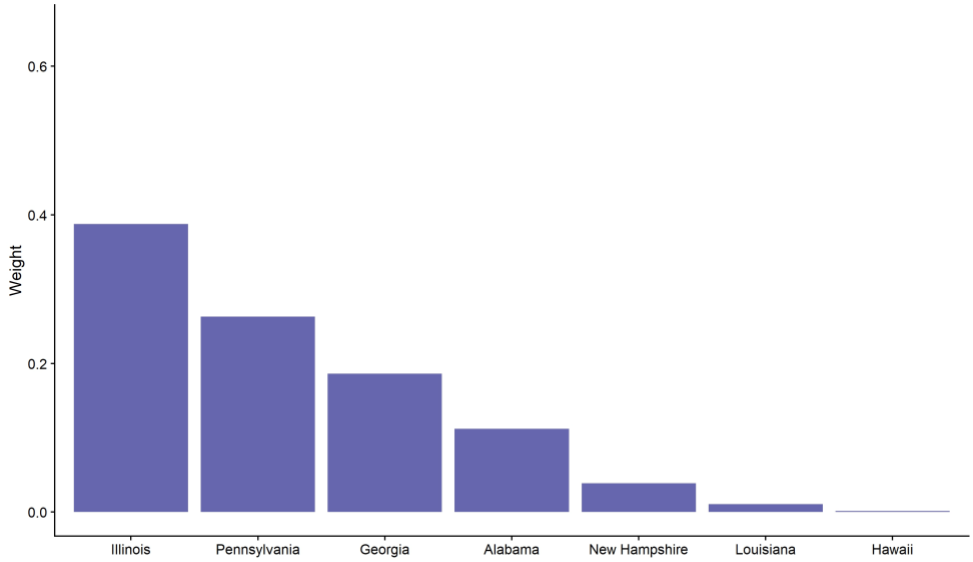
*Note: This table reports estimated effects of Mississippi's and South Carolina's early literacy laws on ELA and math test scores, by grade, including controls for time-varying district characteristics. Each cell corresponds to a separate regression. Test scores are expressed in SEDA CS SDs. The double horizontal line indicates grades with cohorts directly exposed to the policy during the sample period. Standard errors are reported in parentheses. *p<0.10; **p<0.05; ***p<0.01.*

Table A8. Effects on Grade Retention

(1)	(2)	(3)	(4)
	Michigan	**Mississippi**	**South Carolina**
Grade	**Estimate**	**Estimate**	**Estimate**
Kindergarten	-0.0041	0.0121**	-0.0018
	(0.0032)	(0.0048)	(0.0045)
1st	-0.0025	0.0082**	-0.0085
	(0.0017)	(0.0039)	(0.0060)
2nd	-0.0010	0.0047	-0.0046*
	(0.0010)	(0.0042)	(0.0025)
3rd	-0.0004	0.0121***	-0.0008
	(0.0010)	(0.0035)	(0.0025)
Treated Districts	613	73	115
Untreated Districts	2470	2470	2470

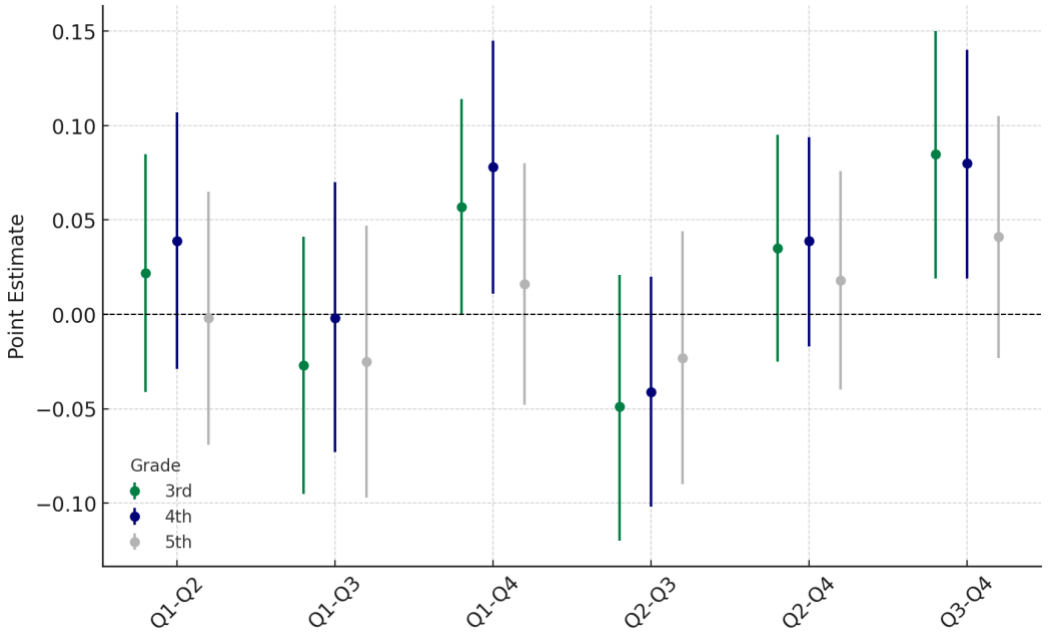
*Note: This table reports estimated effects of early literacy policies on K-3 retention rates in Michigan, Mississippi, and South Carolina. Each cell corresponds to a separate regression. Retention rates are expressed as the share of students retained in each grade. The double horizontal green line indicates years in which cohorts were directly exposed to the policy. Standard errors are reported in parentheses. *p<0.10; **p<0.05; ***p<0.01*

Figure A1. Control State Weights for 3rd Grade Achievement



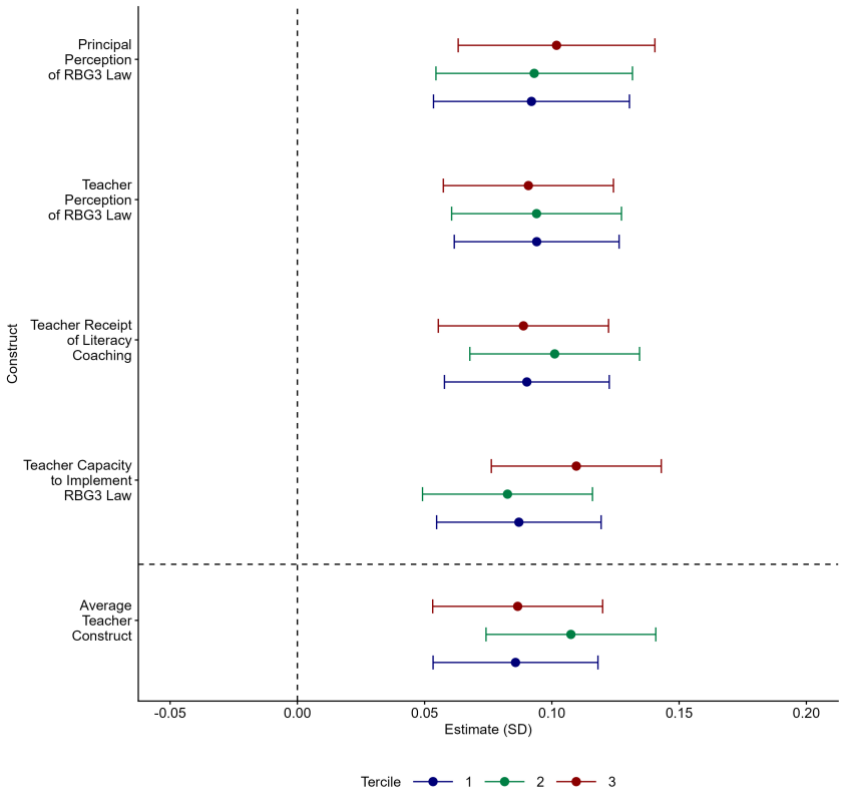
Note: This figure displays the SDID unit weights assigned to control states in the estimation of third-grade ELA achievement effects for Michigan's RBG3 law. Higher weights indicate states that more closely matched Michigan's pre-policy outcome trends and therefore contributed more heavily to the construction of the synthetic control group.

Figure A2. Differences in ELA Achievement Estimates by Baseline ELA Quartile



Note: This figure plots pairwise differences in estimated ELA achievement effects across quartiles of baseline district ELA achievement. Quartiles are based on third-grade ELA scores in 2015–16, where Q1 represents the lowest-achieving districts, and Q4 represents the highest-achieving districts. Points represent differences in SDID treatment effect estimates across quartiles, and vertical lines denote 95% confidence intervals. Test scores are expressed in SEDA CS SDs.

Figure A3. Effects on ELA Achievement by Implementation Construct Tercile



Note: This figure presents estimated effects of the RBG3 law on third-grade ELA achievement across terciles of educator survey-based implementation constructs. Districts are grouped into terciles for each construct, with T1 representing the lowest levels of the construct and T3 representing the highest. Error bars denote 95% confidence intervals.