Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 24-1007

Impact of States' Adoption of Response to Intervention (RTI) on Disability Identification

Zhiling M. Shea
University of Nevada, Las Vegas

Jade M. Jenkins
University of California, Irvine

Christopher Cleveland
Brown University

The rise of specific learning disability (SLD) identification from 1975 to the early 2000s prompted the 2004 reauthorization of the Individuals with Disabilities Education Act to permit states to use a student's response to intervention (RTI) in eligibility determination to improve identification. Single-state studies report that RTI adoption reduced SLD identification by 11% to 15%. We test whether that pattern holds nationally, using a state-year panel of all 50 states and the District of Columbia from 2000 to 2019 and a staggered-adoption difference-in-differences design that exploits variation in when states adopted RTI. In aggregate, we find no detectable effect on identification of SLD, speech or language impairment (SLI), or any disability. State-level interrupted-time-series analyses describe how individual states' trajectories moved around the national average; they do not, by themselves, attribute any state's movement to RTI. In both methods, identification changes are also unrelated to the strength of state policy: states that required RTI in eligibility determination are statistically indistinguishable from states that supported it. We discuss implications for interpreting single-state findings and for the distinction between requiring and supporting RTI policy.

VERSION: September 2026

Suggested citation: Shea, Zhiling M., Jade M. Jenkins, and Christopher Cleveland. (2026). Impact of States' Adoption of Response to Intervention (RTI) on Disability Identification. (EdWorkingPaper: 24-1007). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/se4t-2y83

Impact of States' Adoption of Response to Intervention (RTI) on Disability Identification

Zhiling M. Shea[1], Jade M. Jenkins[2], and Christopher Cleveland[3]

[1]Department of Early Childhood, Multilingual, and Special Education, University of Nevada, Las Vegas, zhiling.shea@unlv.edu,
https://orcid.org/0000-0003-1753-4319
[2]School of Education, University of California, Irvine, jvjenkin@uci.edu,
https://orcid.org/0000-0002-2000-3087.
[3]Department of Education, Brown University, christopher_cleveland1@brown.edu,
https://orcid.org/0000-0002-8619-6088.

The authors declare no conflicts of interest.

Correspondence concerning this article should be addressed to Zhiling M. Shea at the Department of Early Childhood, Multilingual, and Special Education, University of Nevada, Las Vegas. 4505 S. Maryland Pkwy, Las Vegas, NV 89154, USA. Tel: +1-702-895-3205. Email: zhiling.shea@unlv.edu.

Funding Statement

This research was supported by US Department of Education, Office of Special Education Programs (OSEP) grant award #H325D190031. This research was also supported by a grant from the American Educational Research Association which receives funds for its "AERA-NSF Grants Program" from the National Science Foundation under NSF award NSF-DRL #1749275. Opinions reflect those of the author and do not necessarily reflect those AERA or NSF.

Impact of States' Adoption of Response to Intervention (RTI) on Disability Identification

Abstract

The rise of specific learning disability (SLD) identification from 1975 to the early 2000s prompted the 2004 reauthorization of the Individuals with Disabilities Education Act to permit states to use a student's response to intervention (RTI) in eligibility determination to improve identification. Single-state studies report that RTI adoption reduced SLD identification by 11% to 15%. We test whether that pattern holds nationally, using a state-year panel of all 50 states and the District of Columbia from 2000 to 2019 and a staggered-adoption difference-in-differences design that exploits variation in when states adopted RTI. In aggregate, we find no detectable effect on identification of SLD, speech or language impairment (SLI), or any disability. State-level interrupted-time-series analyses describe how individual states’ trajectories moved around the national average; they do not, by themselves, attribute any state’s movement to RTI. In both methods, identification changes are also unrelated to the strength of state policy: states that required RTI in eligibility determination are statistically indistinguishable from states that supported it. We discuss implications for interpreting single-state findings and for the distinction between requiring and supporting RTI policy.

Keywords: response to intervention, specific learning disability, disability identification, state policy

Impact of States' RTI Adoption

Students identified with disabilities frequently encounter developmental, academic, and social challenges across their educational trajectories (Petersen, 2012). Federal policy shapes the services available to these students by setting eligibility criteria and regulating how and where services are delivered. The policy levers embedded in the Individuals with Disabilities Education Act (IDEA) can therefore either expand access to beneficial services or restrict it, depending on how identification and eligibility processes operate in practice.

Few of these levers have been altered as consequentially as the one introduced in the 2004 reauthorization of IDEA, which permitted states to use a student's response to scientifically based intervention as part of the process for determining eligibility under the category of specific learning disability (SLD; IDEA, 2004, § 614(b)(6)(B)). This provision made Response to Intervention (RTI) a permissible alternative to the ability–achievement discrepancy model that had governed SLD identification for nearly three decades. States responded unevenly. Fifteen states had adopted RTI by 2006; 38 had done so by 2014; and by 2018, 46 states had adopted RTI in some form, the last adoption occurring in 2017. Some states required RTI data as part of SLD eligibility determination, whereas others encouraged or supported RTI as instructional practice without specifying its role in identification (Education Commission of the States, 2020; Zirkel & Thomas, 2010a, 2010b).

National identification data suggest that this policy shift coincided with a change in longstanding trends. SLD identification rates rose steadily from the mid-1970s through the early 2000s before slowing and, in several states, reversing (Office of Special Education Programs, 2023). Whether that inflection reflects RTI's effect on identification practice, and whether any such effect was concentrated in states that made RTI consequential for eligibility, remains largely unexamined at the national level (Hughes & Dexter, 2022). The question matters because a decline in identification is not self-interpreting. It is the intended outcome if tiered instruction prevents academic difficulties from becoming disabilities, and a concerning one if students who require special education are prevented from receiving it.

RTI as Instruction and as an Identification Model

RTI originated as a multi-tiered prevention framework in reading and was subsequently extended to mathematics and behavior. In a three-tier configuration, all students receive high-quality core instruction in general education classrooms at Tier 1; students whose progress is inadequate receive small-group targeted instruction at Tier 2; and students who continue to make little progress receive intensive individualized intervention at Tier 3 (Berkeley et al., 2009; Fuchs & Vaughn, 2012; Fuchs et al., 2003; Jimerson et al., 2016). Tier 3 is typically the point at which students are most likely to be referred for special education evaluation, though states and districts vary in how they structure this progression.

The 2004 reauthorization gave RTI a second and distinct function as an SLD eligibility criterion. Under the existing discrepancy model, eligibility rested on a substantial gap between a student's measured ability (e.g., I.Q. assessment) and academic achievement, an approach criticized for requiring students to fail before receiving support and for its questionable construct and predictive validity for SLD identification (Fletcher et al., 2007; Fuchs et al., 2003). Permitting RTI as an eligibility criterion positioned it as a gatekeeping mechanism within the special education referral process rather than solely as an instructional framework.

RTI's dual role has direct implications for identification rates, and it is central to the present study. Two complementary mechanisms link RTI adoption to identification, and they differ in the timing and magnitude of response they imply. The instructional channel operates before referral: tiered instruction and progress monitoring may improve academic trajectories and reduce the number of students who reach evaluation, an effect that should emerge gradually. The eligibility channel operates on the decision rule itself: where RTI data are required, teams weigh documented non-response to intervention rather than an ability–achievement discrepancy, which can change identification quickly even if the underlying distribution of academic difficulty is unchanged. Which of these two mechanisms dominates is an empirical question, and it motivates our study's analytic approach; we separate states that required RTI in eligibility determination from states that supported RTI's use.

Variation in State RTI Policy

Federal law permitted states to use RTI in SLD identification but did not require it, and states exercised that discretion in different ways. Analyses of state statutes and guidance documents have found substantial variation in whether RTI is mandated or encouraged, whether the discrepancy model remains permissible alongside it, how many tiers are specified, what constitutes adequate progress, and what documentation eligibility teams must produce (Harr-Robins et al., 2009; Zirkel & Thomas, 2010a, 2010b). Some states codified RTI as a required component of SLD eligibility determination; others issued guidance encouraging tiered instruction while leaving identification procedures unchanged; and a small number established no statewide RTI policy at all.

Implementation varied as much as policy. National surveys conducted a decade apart documented persistent variability in tier structure, screening practice, progress-monitoring procedures, and decision rules governing movement between tiers, despite widespread formal adoption (Berkeley et al., 2009; Berkeley et al., 2020; Hoover et al., 2008). This gap between statutory language and school-level practice means that policy adoption cannot be assumed to indicate changed identification behavior, and it suggests that any national estimate may conceal considerable heterogeneity.

Evidence on RTI and Disability Identification

Efficacy trials and field studies of tiered intervention have mixed results, several report reductions in the number of students requiring special education services (VanDerHeyden et al., 2007; Wanzek & Vaughn, 2011). Reviews of this literature have consistently noted that the evidence base on RTI as an identification model remains thin relative to its policy footprint, and that most studies examine RTI as instruction rather than as an eligibility mechanism (Hughes & Dexter, 2022; Kovaleski et al., 2022). The national evaluation of RTI practices in elementary reading found that implementation quality varied widely across schools and that impacts did not consistently favor students receiving tiered intervention (Balu et al., 2015), and Wanzek and Vaughn (2011) found no significant relationship between a three-tier model and special education placement across elementary cohorts. VanDerHeyden and colleagues (2007), for example, observed identification rates declining from 6% to 3.5% across five elementary schools following a year of RTI implementation. Findings from scaled implementation have been more mixed. Hall-Mills (2021) analyzed Florida disability-identification prevalence data before and after the state adopted RTI in SLD eligibility determination, documenting declines in both language impairment and SLD identification. Gilmour et al. (2023) used statewide administrative data from Tennessee and found reductions in overall special education and SLD identification in grades K–5 following statewide adoption, with larger declines for Black and low-income students. Both studies rely on pre-post comparisons and therefore cannot rule out that concurrent factors drove the observed changes. Ainsworth et al. (2024) addressed this limitation by exploiting the staggered rollout of Oregon's Response to Instruction and Intervention initiative across districts, estimating that RTI adoption reduced special education identification by 1.4 percentage points (11%) and SLD identification by 0.5 percentage points (15%). Notably, they also found a decline in speech or language impairment identification of 0.7 percentage points (12%), a spillover onto a category RTI does not target. Oregon's initiative provided participating districts with technical assistance, training, and funding conditional on committing to RTI-based SLD identification, a level of state investment that is not characteristic of adoption nationally. Whether the identification response observed under such conditions describes what happened across the 46 states that adopted RTI in some form remains an open question, as does whether the response varies with how consequential a state made RTI for eligibility determination.

Three features of this evidence base limit what can be concluded from it about the national picture. First, ingle-state designs allow researchers to characterize policy and implementation with a precision national studies cannot match, but a state that implemented RTI as a required eligibility component with sustained state support may not be representative of the states that adopted RTI in some form, and an estimate recovered in that setting describes what RTI can do under favorable conditions rather than what it did on average. Second, the literature has not tested whether the identification response tracks the strength of the policy, leaving open whether the reported declines reflect RTI generally or the specific decision to make it consequential for eligibility. States differ from one another along the dimension that should matter most: whether RTI data were required or supported for eligibility determination or instructional practice. Third, existing work has given limited attention to whether observed declines represent reductions in identification or redistribution across categories, since a decline in SLD is interpretively ambiguous without knowing what happened to other categories.

Purpose of the Present Study

Impact of States' RTI Adoption

We examined how state adoption of RTI affected the rate at which students were identified with SLD, with SLI, and with any disability. We therefore address three research questions:

Research Question 1: What is the average effect of state RTI adoption on SLD, SLI, and overall disability identification??

Research Question 2: Do these effects vary by adoption cohort?

Research Question 3: To what extent do identification changes differ across states with varying RTI policies?

Our contribution is threefold. First, we use a national panel covering all states rather than a single state, allowing us to assess whether patterns reported in within-state studies generalize. Second, we exploit variation in the timing of state adoption within a difference-in-differences framework, which supports causal interpretation of the estimated effects. Third, we distinguish states that required RTI as part of SLD eligibility determination from states that supported RTI without altering identification procedures, which allows us to test whether the identification response tracks the policy's role in eligibility. We examine SLI alongside SLD and overall disability identification to assess whether any change in SLD reflected a reduction in identification or a redistribution across categories.

Method

Data

We integrate multiple sources at the state–year level. Sources include: states Departments of Education (RTI adoption coding); reports from the Institute of Education Sciences and its Regional Educational Laboratories (Detgen et al., 2011; Harr-Robins, Shambaugh, & Parrish, 2009; Sawyer et al., 2008; Stepanek & Peixotto, 2009); Berkeley et al. (2009) and Zirkel &

Thomas (2010a, 2010b) for adoption timing and strength; the Office of Special Education Programs (OSEP) IDEA Section 618 child-count data for disability counts, cross-checked against the compiled Section 618 series distributed by the Unlocking Potential Data Center (UPDC), which republishes the OSEP counts in panel form and does not add cases; CDC WONDER for child population counts; and the American Community Survey (ACS) for demographic and socioeconomic covariates. We merged on Federal Information Processing Standard (FIPS) state codes and year.

Measures

State RTI adoption. RTI status is a binary variable that equals 1 in the years a state adopted RTI and 0 otherwise. For example, Colorado adopted RTI in 2009. Colorado thus has a value of 1 from 2009 to 2019 and 0 for the years before 2009. The variable information is primarily drawn from sources on states' Department of Education websites, reports from the Institute of Education Sciences (Detgen et al., 2011; Harr-Robins et al., 2009; Sawyer et al., 2008; Stepanek & Peixotto, 2009), and papers (Berkeley et al., 2009; Zirkel & Thomas, 2010).

Adoption timing and strength. We coded two distinct constructs, reported separately in Appendix A of Online Supplemental Materials because they do not partition the states identically. Adoption timing (Appendix A1) records whether and when a state put RTI in place in any form by statute or regulation, or through state guidance, initiatives, or technical assistance, set to 1 in the first year of adoption and thereafter; it distinguishes twenty-three early adopters (2006–2009), twenty-three later adopters (2010–2017), and five states that never adopted. Adoption strength (Appendix A2) records whether RTI was established in law, sorting states into RTI-required (mandated as an SLD-identification model), RTI-supported (promoted through guidance or funding but not mandated, with the discrepancy model often retained), and non-RTI (never adopted or not in law): seventeen required, twenty-six supported, and eight without a mandate. The two codings diverge only for three states—Arizona (2006), Alabama (2011), and South Carolina (2010)—that adopted RTI in practice but never codified it, which is why five states are never-adopters under timing but eight fall in the non-RTI-in-law category (5 + 3 = 8). Our estimates use the Appendix A1 adoption year for treatment timing and the Appendix A2 category for policy strength. Figure 1 shows the distribution of policy types over the panel, rising sharply after the 2004 IDEA reauthorization and continuing through the mid-2010s.

Coding protocol and reliability. Because a state's RTI adoption must be inferred from statutes, regulations, administrative codes, and state guidance documents, coding requires judgment, and we therefore document the protocol and assess its reliability. A graduate researcher and the lead author independently coded each state's adoption year and adoption strength from state Department of Education websites and eligibility guidance, IES/REL reports (Detgen et al., 2011; Harr-Robins et al., 2009; Sawyer et al., 2008; Stepanek & Peixotto, 2009), and the state surveys in Berkeley et al. (2009) and Zirkel & Thomas (2010a, 2010b). Both coders worked from a written rule set and did not confer while coding. The decision rules mapped the "required / permitted / encouraged" language used in the anchor sources onto our two variables. Where sources disagreed, the rule set gave precedence to the state's own statutory or regulatory language over secondary characterizations, and to the earliest year for which documentary evidence of statewide adoption could be established. Intercoder agreement was assessed on all 51 units, and the two coders' assignments were nearly identical. Coders assigned the same adoption year for 48 of 51 units and the same adoption strength for 46 of 51 units. For the units coded differently, the coders jointly revisited the source documents and reconciled to consensus; the consensus codes are those reported in Appendix A.

Outcome Variables. We constructed three state-year rates: the proportion of students identified with SLD, with SLI, and with any disability. Each numerator is the aggregated OSEP Section 618 child count for the relevant category, which covers students ages 6 through 21. Our primary specification divides each numerator by the state–year K–12 enrolled population, and we use the same K–12 denominator for all three outcomes so that the three rates are directly comparable. Because the numerator's age band and the K–12 denominator do not align exactly, we treat these as identification rates per enrolled student rather than as prevalence estimates for a fixed population, and we return to this in the Limitations. A secondary set of outcomes restricts both numerator and denominator to elementary-age children, ages 6 through 11, dividing the Section 618 count for that age band by the CDC WONDER population for the same ages. These are the one construction in which the two populations align, and we report them as a robustness check in Appendix D in Online Supplemental Materials. Figure 2 provides a descriptive overview of the SLD trends for all 50 states: RTI-required states first, then RTI-supported states, then non-RTI states, ordered within each group by the steepness of decline between 2006 and 2019.

Covariates. Table 1 presents all the covariates included in the study. Per-pupil expenditure, pupil–teacher ratio, and English learner share are from the NCES Common Core of Data; Supplemental Nutrition Assistance Program and Temporary Assistance for Needy Families receipt are from the University of Kentucky Center for Poverty Research (2022) National Welfare Data; the remaining covariates are from the American Community Survey. Table 1 reports baseline characteristics measured in 2006, comparing the fifteen states that had adopted RTI by that year with the remaining thirty-six states. The two groups are similar in nearly every dimension.

Empirical Framework

Callaway–Sant'Anna Difference-in-Differences (CSDID). We estimate the effect of RTI adoption using a staggered-adoption difference-in-differences design (DID). This design compares adopting states with states that had not (or not yet) adopted RTI, and each state's post-adoption trajectory with its own pre-adoption trend. We report two-way fixed effects (TWFE) only as a robustness check (Appendix C in Online Supplemental Materials).

Our primary estimand is the group-time average treatment effect on the treated, ATT(g, t), the effect in calendar year t for the cohort of states first adopting RTI in year g:

$$\text{ATT}(g, t) = \text{E}[\, Y\_t(g) - Y\_t(0) \mid G = g \,] \text{ ,}$$

where Y_t(g) is the identification rate (% of students with SLD, SLI, disability) under RTI adoption states in year g, Y_t(0) the identification rate from never-adopted states, and G the state's adoption cohort. We aggregated ATT(g, t) into (i) a single overall/simple ATT, (ii) a dynamic event-study path in event time (t − g), and (iii) group-specific and calendar-time summaries. It was estimated with the doubly robust estimator of Callaway and Sant'Anna (2021), implemented in Stata by Rios-Avila et al. (2021) using not-yet-treated states as the comparison group. Only five states never adopted RTI, and, as we report in the Robustness section, the never-treated comparison does not satisfy the parallel-trends requirement. We included state-level controls for the proportion of the population receiving Supplemental Nutrition Assistance Program (SNAP) benefits, the proportion receiving Temporary Assistance for Needy Families (TANF), the share of students classified as English language learners, gross state product per capita, and CPI-adjusted annual personal income per capita, which together capture broad differences in economic conditions across states.

Interrupted Time Series. To address our RQ3, whether disability identification changes differ across states with varying RTI adoption policies, we supplement the CS estimates with state-level interrupted time series (ITS) analyses for all three outcomes. For each state, we fitted a linear pre-trend using the years before RTI adoption and extrapolate it forward, which allowed us to assess whether a state's post-adoption trajectory breaks from its own pre-existing trend.

Interrupted Time Series (ITS). The ITS specification is

$$Y_{st} = \pi_{0s} + \pi_{1s}t + u_{st}, t < m_s \; Y_{st} = \psi_{0s} + \psi_{1s}t + v_{st}, t \geq m_s$$

where Y_{st} is the identification rate in state s and year t and m_s is the adoption year. We fit the pre-adoption segment on the pre-adoption years and extrapolate it across the panel to give each state's counterfactual trajectory; the vertical gap between this projection and the observed post-adoption path is the state's departure from its own trend. Each state's level and slope change is estimated with Newey–West standard errors (lag 2) and used only to classify that state's break relative to its own pre-trend. Because the fits are state-specific and unpooled, the ITS is a descriptive complement to the CS estimates, not a source of pooled inference.

Results

Callaway Sant'anna DID Estimates

To answer RQ2, Figure 3 shows the dynamic CS estimates for the K–12 proportion of students identified with SLD, SLI, and any disability. For SLD (Figure 3, panel A), the pre-treatment estimates for years −5 through −1 are all close to zero and their confidence intervals include zero, supporting the parallel trends assumption underlying the design. In the post-treatment period the estimates drift mildly positive as the horizon lengthens, but they are not distinguishable from zero at any point from year 0 through year 8. Table 2 Panel A shows the aggregate ATT is 0.157 percentage points (SE = 0.187, *p* = .401, 95% CI [−0.210, 0.524]) against a 2006 base rate of 5.42%. The average, however, conceals variation across cohorts. A joint Wald test of the hypothesis that all ten cohort ATTs are equal is rejected, $\chi^2(9) = 387.06$, $p < .001$. Two cohort estimates are individually distinguishable from zero, and each rests on a single state: 2008 Missouri, at −0.829 percentage points (SE = 0.126, $p < .001$), and 2013 Mississippi, at 0.483 (SE = 0.138, $p < .001$). Because single-state cohorts carry small standard errors and can dominate a homogeneity test, we re-ran it on the eight cohorts holding two to fifteen states each. Heterogeneity survives, $\chi^2(7) = 17.23$, $p = .016$, and dropping the two single-state cohorts leaves the aggregate unchanged at 0.157. Table 2, Panel C reports all ten cohorts with their sizes. We find no evidence that state RTI adoption reduced SLD identification on average, and we return to what the cohort-level pattern does and does not support in the Discussion.

For SLI (Figure 3, panel B), both pre- and post-treatment confidence intervals include zero throughout the full −5 to 8 year window. Table 2 Panel A shows the aggregate ATT is 0.005 percentage points (SE = 0.066, p = .940, 95% CI [−0.124, 0.134]) against a base rate of 2.41%. RTI adoption does not appear to have shifted SLI identification rates in either direction. Unlike SLD, SLI shows no cohort heterogeneity once the single-state cohorts are set aside: the joint test across all ten cohorts is rejected, $\chi^2(9) = 246.34$, $p < .001$, but across the eight multi-state cohorts it is not, $\chi^2(7) = 5.71$, $p = .574$.

For overall disability identification (Figure 3, panel C), confidence intervals are wide throughout both the pre- and post-treatment periods and include zero at every point. Table 2 Panel A shows the aggregate ATT is 0.114 percentage points (SE = 0.233, p = .624, 95% CI [−0.343, 0.571]). Overall disability behaves like SLD rather than SLI on the cohort dimension: the joint homogeneity test is rejected across all ten cohorts, $\chi^2(9) = 65.17$, $p < .001$, and remains so across the eight multi-state cohorts, $\chi^2(7) = 15.44$, $p = .031$. We also estimated effects on the non-SLD residual, all disability identification other than SLD, which is the direct test of whether students were redistributed across categories rather than identified at lower rates. That estimate is likewise null, at −0.043 percentage points (SE = 0.169, $p = .799$). These results are consistent with the calendar-year patterns in Figure 2 and the event-time patterns in Appendix B in Online Supplemental Materials, neither of which shows a visible difference in disability trends around adoption.

Turning to RQ3, we estimated group-specific ATTs by adoption strength within the same framework. Each policy group is estimated on the subsample comprising that group together with the five never-adopting jurisdictions, retaining the not-yet-treated comparison used for the primary aggregate estimates. Restricting the sample this way holds the contrast within a single policy group while keeping the never-adopters available as controls in cohort-periods where no within-group state is still untreated. Estimating the same contrast with never-treated comparisons only returns a difference of −0.157 percentage points (SE = 0.372, $p = .674$), which is less precise and does not change the conclusion. States that required RTI in eligibility determination return an ATT of −0.008 percentage points (SE = 0.258, $p = .976$); states that supported RTI without requiring it return 0.175 (SE = 0.198, $p = .378$). The difference between them is −0.182 percentage points (SE = 0.326, $p = .575$, 95% CI [−0.820, 0.456]); its minimum detectable effect at 80% power is 0.913 percentage points, roughly the size of the Ainsworth et al. (2024) SLD estimate. This contrast does not rule out a required-versus-supported difference of single-state magnitude. We test the influence of the strength coding directly in the Robustness Checks section; no alternative coding produces a detectable required-versus-supported difference.

ITS Estimates

Figure 4 plots each state's observed SLD series against an ITS projection, grouped by policy type. We also estimated each state's level and slope change with Newey–West standard errors at lag 2. Among the 43 states assigned to a policy group, a significant negative post-adoption slope change appears in 2 of the 17 required states and 3 of the 26 supported states, shares that are indistinguishable (Fisher's exact $p = 1.000$). Mean slope change is 0.196 percentage points per year in required states (SD = 0.236) and 0.091 in supported states (SD = 0.143); the difference of 0.105 is not significant (Welch t = 1.64, $p = .114$). The state-level trajectories show substantial heterogeneity, but that heterogeneity is not organized by the strength of state RTI policy.

Robustness Checks

Figure 3 also overlays a doubly robust specification adjusting for covariates. The two series track one another closely at every event time, and the covariate-adjusted aggregate estimates are likewise null for all three outcomes (SLD −0.008, $p = .942$; SLI 0.020, $p = .732$). We report the specification without covariates as primary because the covariates did not make a difference.

As a robustness check on the CS estimates, we also estimated conventional TWFE regressions, reported in Appendix C in Online Supplemental Materials. Table C1 presents results for four state groupings: all states, required versus supported states, required versus non-RTI states, and supported versus non-RTI states. Across all states the covariate-adjusted TWFE estimate for SLD is a reduction of 0.19 percentage points. That estimate is not robust, and we report it with two qualifications. It is produced on a sample of 604 state-years rather than the full 1,020, because the English learner covariate has no coverage before 2004 or after 2018 and covers only 23 to 25 states between 2015 and 2018. Re-estimating the same specification without covariates on the full 2000–2019 panel gives −0.036 percentage points (p = .766), and adding state-specific linear trends to the covariate-adjusted specification gives −0.050 (p = .418). SLI and overall disability show no significant association in any grouping. We therefore read the TWFE results as consistent with the CS estimates in showing no robust effect.

We also examined whether the results are specific to the K–12 population by re-estimating the models with outcomes restricted to elementary-age children, ages 6 through 11, the range in which RTI-based instruction and identification decisions are most concentrated. Table D1 in Online Supplemental Materials reports the estimates: 0.145 percentage points for SLD (p = .346), −0.092 for SLI (p = .389), and −0.058 for overall disability (p = .804). Figures D1 through D3 show the per-state trajectories across the elementary band for SLD, SLI, and overall disability. The pattern matches the K–12 results, which weighs against the possibility that the null reflects dilution across the wider K–12 population.

Two further checks support this result. Table 2 Panel B reports the SLD ATT under all four combinations of not-yet-treated versus never-treated comparison states, crossed with and without covariates, over a common 2004–2019 window. Changing only the comparison group (not-yet-treated to never-treated, holding covariates fixed at none) moves the estimate from 0.157 to −0.010; changing only the covariates (holding the not-yet-treated comparison) moves it from 0.157 to −0.008. Neither change alone materially moves the estimate away from the primary result. Three alternative staggered-adoption estimators that use clean comparisons by construction return the same null for SLD: the CS estimate of 0.157, the Borusyak et al. (2024) imputation estimate of 0.162 (SE = 0.201), and the de Chaisemartin and D'Haultfœuille (2020) average total effect of 0.062 (SE = 0.083), whose joint test of event-study effects returns p = .493 and whose joint placebo test returns p = .195. Allowing one or two years of anticipation, as would be appropriate if states changed practice before codifying policy, leaves the estimate unchanged at 0.136 (p = .535) and 0.094 (p = .683). The descriptive trends in Figure 2 are consistent with this: SLD identification declines in most states over the period, but the decline appears in states with and without RTI policies alike and shows no comparable changes around adoption.

Lastly, although our strength coding gives precedence to states' own statutory and regulatory language, published 50-state reviews classify a number of states differently, so we re-estimated the contrast under two alternative codings drawn from those reviews (Hauerwas et al., 2013; Zirkel, 2017). The first reassigns fourteen jurisdictions and yields 19 required and 24 supported states; the second retains only the states both codings assign identically, leaving 10 required and 18 supported and setting misclassification to zero by construction. Appendix E in Online Supplemental Materials reports all three. The point estimate is not stable across them, and no coding yields a detectable difference. Because the direction reverses, we draw no inference from the sign of the strength contrast; because the precision does not improve, we read the contrast as uninformative about a required-versus-supported gap below roughly one percentage point.

Discussion

We examined nationally whether state RTI policy changed the rate at which students are identified with SLD, and whether any change was specific to the category RTI was designed to affect. Across a national state-year panel we find no detectable effect of RTI adoption on identification of SLD, SLI, or any disability. The estimates are precise at the 95% confidence interval for SLD spans −0.21 to +0.52 percentage points, or −3.9% to +9.7% of the 2006 base rate and excludes declines of the magnitude reported in prior single-state studies. Identification with SLI did not change either, or neither did the non-SLD residual, so we find no evidence of redistribution across eligibility categories.

Specificity of the SLD Identification Effect

A central issue in the SLD identification literature is whether replacing or supplementing traditional identification approaches changes which students are identified or instead changes only the label assigned to students who would have received special education services under another eligibility framework (Fletcher & Miciak, 2019; Kavale & Spaulding, 2008). Prior evidence on RTI and identification comes from single states. Gilmour et al. (2023) found reductions in SLD identification following Tennessee's RTI adoption; Hall-Mills (2021) documented declines in SLD alongside changes in language-impairment identification in Florida; and Ainsworth et al. (2024), using staggered district-level implementation in Oregon, found declines in SLD, overall special education, and speech or language impairment identification. Prior research has also documented category shifts alongside RTI implementation: Hall-Mills (2021) reported changing prevalence patterns for SLD and language impairment during Florida's RTI policy period, and Ainsworth et al. (2024) found that RTI adoption in Oregon reduced identification in both SLD and SLI. Ainsworth et al. (2024) estimated that RTI adoption in Oregon reduced overall special education identification by 1.4 percentage points (11%), SLD identification by 0.5 points (15%), and SLI identification by 0.7 points (12%).

Each of these lies outside the corresponding national confidence interval we estimate. Our minimum detectable effects at 80% power are 0.525 percentage points for SLD, 0.184 for SLI, and 0.653 for overall disability. For SLI and overall disability we are therefore comfortably powered to detect effects of the reported size and do not find them. For SLD the K–12 detection boundary of 0.525 points sits just above the 0.5-point effect reported, and the elementary-age estimates in Appendix D in Online Supplemental Materials provide an age-band check rather than a resolution of that ambiguity. Restricting both numerator and denominator to ages 6 through 11, the population closest to the one Ainsworth et al. studied, gives a minimum detectable effect of 0.432 percentage points for SLD, below the reported effect, alongside 0.300 for SLI and 0.650 for overall disability. On that band we are powered to detect all three Oregon effects, and we detect none of them: the estimates are 0.145 points for SLD (95% CI [−0.157, 0.447]), −0.092 for SLI ([−0.301, 0.117]), and −0.058 for overall disability ([−0.513, 0.398]). Because SLI is not a target of RTI's tiered academic intervention model, a rise in SLI identification alongside a fall in SLD would have indicated relabeling rather than reduced identification. We observe neither movement: the SLI estimate is 0.005 percentage points with a confidence interval of −0.124 to 0.134, roughly ±5% of the base rate. The non-SLD residual provides the more comprehensive test, aggregating every disability category other than SLD, and it is likewise null at −0.043 percentage points. Taken together, the absence of movement in SLD, in SLI, and in the residual indicates that state RTI adoption did not redistribute students across eligibility categories nationally because it did not detectably move identification at all.

Two differences between the studies may help explain this divergence. First, the studies examine different populations. Our primary outcome is a state-year identification rate across the broader K–12 population, whereas Ainsworth et al. (2024) examine elementary students and exploit variation in RTI implementation across Oregon districts. A spillover of RTI's effect onto SLI identification, if concentrated in the early elementary grades, could therefore be attenuated in a K–12 aggregate. Our analysis provides little support for this explanation, however. Restricting the sample to elementary-age children, ages 6 through 11, raises the mean SLI identification rate from 2.41 percent of K–12 enrollment to 4.22 percent of the age-6-through-11 population, so the restriction does concentrate the category in which Ainsworth et al. report a spillover. It does not produce a corresponding decline. The estimate is −0.092 percentage points, and the interval excludes the 0.7-point reduction they report. Post-adoption SLI trajectories within this age range (Figure D2) depart from projected pre-adoption trends in several states, but the departures are more often characterized by slower declines than projected rather than by the steeper declines observed in Oregon.

A second and potentially more consequential difference concerns the treatment itself. Ainsworth et al. (2024) studied an intensive district-level implementation initiative in which participating districts received training, technical assistance, and financial support and committed to implementing RTI district-wide, including its use in SLD identification. Our treatment is state policy adoption. A state statute or administrative rule establishes a policy environment and may authorize or require the use of RTI, but it does not necessarily ensure that schools change screening, progress monitoring, intervention, referral, or eligibility practices. This distinction matters because spillover effects on SLI may arise from broader changes in referral and evaluation behavior rather than from changes in the SLD eligibility rule alone.

Three non-exclusive readings of this contrast are worth distinguishing. First, adoption may differ from implementation: Oregon's initiative paired adoption with sustained technical assistance and funding that state statutes alone do not guarantee, so a policy-adoption indicator may understate the effect of RTI where it is genuinely implemented with fidelity. Second, our treatment may be mistimed or mismeasured relative to the Table A2 strength categories: coding a state's earliest documentary evidence of adoption, rather than the date an eligibility mandate or discrepancy-model prohibition took effect, could attenuate a true effect if identification practice changed later than the coded year. Third, the single-state estimates themselves may not generalize: Oregon and Tennessee are, respectively, one district-level initiative and one state, and each may have implemented RTI under conditions unusually favorable to detecting an effect. Our data cannot adjudicate among these readings, and they are not mutually exclusive; the first is consistent with implementation research showing substantial variation in how states and districts operationalize RTI (Berkeley et al., 2009; Berkeley et al., 2020).

No Effects of RTI on Students with SLI and Disabilities

The null effect in SLI identification does not mean that no students were reclassified into other disability categories. It means only that SLI does not appear to be the destination category for any decline in SLD identification. This distinction is important because other health impairment and autism identification increased over time. Our design cannot determine whether any students who might previously have been identified with SLD were instead identified under these or other categories. Similarly, the absence of an effect on overall disability identification should not be interpreted as evidence that RTI had no effect on special education participation. Rather, we do not detect a statistically significant change in the aggregate disability identification rate corresponding to RTI adoption.

Variation by State RTI Policy

The third research question examined whether the magnitude of the SLD identification response differed according to the strength of state RTI policy. We cannot distinguish it: the required-versus-supported contrast reported above is imprecise rather than precisely null, with a minimum detectable effect of 0.913 percentage points. The state-level interrupted time series estimates are consistent with this reading, with an important caveat. A significant steepening of the post-adoption decline in SLD identification appears in 2 of the 17 required states and 3 of the 26 supported states, shares that are indistinguishable (Fisher's exact $p = 1.000$). Significant movement in the opposite direction, a slowing of the pre-existing decline, is far more common and is likewise unrelated to policy strength, appearing in 13 of 17 required and 17 of 26 supported states. With groups of this size the comparison has limited power to detect small differences, but neither pattern is ordered as a dose-response account would predict. States that required RTI as part of SLD eligibility determination return an ATT of −0.008 percentage points and states that supported or permitted RTI without mandating it return 0.175, a difference of −0.182 percentage points that is not statistically significant ($SE = 0.326$, $p = .575$, 95% CI [−0.820, 0.456]). The post-adoption response was far from uniform within either policy group. The state-level slope changes bear this out. They average 0.196 percentage points per year in required states ($SD = 0.236$) and 0.091 in supported states ($SD = 0.143$), a difference of 0.105 that is not statistically significant (Welch $t = -1.64$, $p = .114$), and the two distributions overlap across nearly their entire range. Figure 4 plots each state's own trajectory. These checks indicate that a national deceleration in SLD identification, present in adopting and non-adopting states alike, accounts for a substantial part of the raw “slowing” pattern. We treat the ITS accordingly as descriptive of the timing of the national inflection relative to adoption, not as evidence that adoption caused it.

Limitations

Several limitations bear on how these findings should be interpreted, most of which stem from the aggregate nature of state-level identification data. First, because our data are state-year aggregates, we observe identification but cannot determine whether students no longer identified with SLD are receiving effective tiered support, are receiving inadequate support without the procedural protections that identification confers, or are being identified later in their school careers. This last possibility deserves emphasis, because concerns that RTI delays identification have been raised throughout its adoption (Reynolds & Shaywitz, 2009). Aggregate rates are also silent on disproportionality. A decline in the overall identification rate is consistent with either a narrowing or a widening of the racial, ethnic, and linguistic disparities that have long characterized SLD identification (Artiles, 2015; Morgan et al., 2015).

Second, adoption dates are coded from state documentation and published sources and may be imprecise at the district level: districts within a state may adopt at different times, and the interval between formal adoption and meaningful implementation varies. The construct matters for interpretation. As indicated in Method, the timing variable records the earliest documentary evidence that a state put RTI in place in any form, including guidance and technical-assistance initiatives may not took legal effect. The two dates diverge in some states: several of our coded years precede the year an eligibility mandate took legal effect, where a state's regulation set a later phase-in deadline. Fifteen states carry a coded year of 2006, the year the IDEA 2004 regulations took effect, and ten of those were still classified as permitting rather than requiring RTI three years later. Because a coded year can therefore precede the year identification practice was legally obliged to change, our estimates are attenuated relative to a design dated on mandate effective dates, and cohort membership throughout, including the cohort estimates in Table 2 (Panel C), would shift under that alternative dating. Several early mandates also bound only part of the grade range, which a state-year binary treats as full treatment from the first year and which attenuates further. Our results may understate the response in states that implemented RTI with fidelity. Third, our study period overlaps with other policies affecting identification, including federal accountability requirements and state special education finance reforms shown to influence identification rates (Cullen, 2003; Dhuey & Lipscomb, 2011). The absence of pre-adoption trends offers some reassurance that our estimates do not simply reflect broader secular change, but we cannot fully isolate the effect of RTI from concurrent shifts in policy and practice. Relatedly, the elementary-age robustness check reported in Appendix D of the Online Supplemental Materials is bounded by data availability: OSEP Section 618 reports age-band detail only from 2005, so that check spans 2005 through 2019 rather than the full study period. Seventeen of the forty-six states with a coded adoption year have fewer than three usable pre-adoption observations inside that window, fifteen of them because they adopted in 2006 and are therefore observed only in 2005.

Last, we estimate effects on identification, not on learning. A decline in SLD identification is the intended result if tiered instruction prevents academic difficulties, and a troubling one if students who need specialized instruction remain in intervention cycles without receiving it. Our data cannot distinguish these.

Implications for Policy and Practice

For state policymakers, the central implication is that adopting RTI is not equivalent to changing identification practice. Nationally, we do not detect an average change in identification associated with RTI adoption, and the required-versus-supported contrast is imprecise: our confidence interval for that difference does not rule out an effect the size of the single-state estimates in the literature. The difference between policy and practice is well documented: national surveys conducted a decade apart found persistent variability in how districts structured tiers, monitored progress, and applied decision rules, despite widespread formal adoption (Berkeley et al., 2009; 2020), and state statutes and guidance documents themselves differ substantially in what they specify (Zirkel & Thomas, 2010a, 2010b). Implementation surveys point to preparation as part of the explanation, finding that personnel frequently reported insufficient training in the assessment and data-based decision-making practices RTI requires

(Berkeley et al., 2020; Hoover et al., 2008). States revising RTI or multi-tiered systems of support policy would benefit by attending to implementation infrastructure and the specific role RTI data play in eligibility determination (Balu et al., 2015; Hoover, 2010; Kovaleski et al., 2022).

For district personnel and educator preparation programs, our findings indicate that identification rates responded to RTI in some states and not others, which places weight on local decision rules. How schools operationalize varies considerably, and different decision rules can classify the same student differently (Fuchs et al., 2003; Hoover, 2010; VanDerHeyden & Witt, 2005). Eligibility teams operating under RTI therefore need explicit criteria for when insufficient response to intervention warrants referral, and explicit timelines that prevent extended intervention from functioning as indefinite delay, a concern raised since RTI's earliest adoption, that a model intended to replace waiting for failure could instead postpone identification (Kavale & Spaulding, 2008; Kovaleski et al., 2022; Reynolds & Shaywitz, 2009).

Directions for Future Research

The pressing need is to link identification effects to student outcomes. Whether declining SLD identification represents successful prevention or delayed service cannot be answered from identification counts, and the field's interest in RTI ultimately rests on that distinction. Studies linking state RTI implementation to achievement, and to the timing of identification within students' school careers, would address it directly. Second, disproportionality warrants dedicated attention. Because our aggregate data cannot reveal whether declines were evenly distributed across student groups, research using disaggregated identification counts or individual administrative records is needed to determine whether RTI narrowed, widened, or left unchanged the disparities it was partly intended to address. Third, implementation should be measured rather than inferred. Direct measures of screening practice, progress-monitoring fidelity, and the operational role of RTI data in eligibility decisions would permit a sharper test of whether the state variation we document reflects differences in how RTI was carried out.

Conclusion

Extending previous literature about RTI’s impacts on special education identification to a national state-year panel, we do not find positive or negative effects. Our confidence intervals exclude declines of the magnitude previous studies report, for each corresponding outcome. Each of those states paired adoption with sustained implementation support, for example, Oregon's initiative provided participating districts with technical assistance, training, and funding conditional on committing to RTI-based identification, and that is not characteristic of adoption nationally.

References

Ainsworth, N., Cleveland, C., & Penner, A. (2024). The Effects of Response to Intervention on Disability Identification and Achievement. (EdWorkingPaper: 24-1010). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/72jm-2m09

Artiles, A. J. (2015). Beyond responsiveness to identity badges: Future research on culture in disability and implications for Response to Intervention. *Educational Review*, 67, 1–22.

Artiles, A. J., Kozleski, E. B., Dorn, S., & Christensen, C. (2006). Learning in inclusive education research: Re-mediating theory and methods with a transformative agenda. *Review of Research in Education*, 30(1), 65–108. https://doi:10.3102/0091732x030001065

Baker, A. C., Larcker, D. F., & Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates?. *Journal of Financial Economics*, 144(2), 370-395.

Balu, R., Zhu, P., Doolittle, F., Schiller, E., Jenkins, J., & Gersten, R. (2015). Evaluation of Response to Intervention practices for elementary school reading. *National Center for Education Evaluation and Regional Assistance*, *November*, 1–308. https://ies.ed.gov/pubsearch/pubsinfo.asp?pubid=NCEE20164000

Berkeley, S., Bender, W. N., Gregg Peaster, L., & Saunders, L. (2009). Implementation of response to intervention: A snapshot of progress. *Journal of Learning Disabilities*, *42*(1), 85–95. https://doi.org/10.1177/0022219408326214

Berkeley, S., Scanlon, D., Bailey, T. R., Sutton, J. C., & Sacco, D. M. (2020). A Snapshot of RTI Implementation a Decade Later: New Picture, Same Story. *Journal of Learning Disabilities*, 53(5), 332–342. https://doi.org/10.1177/0022219420915867

Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285. https://doi.org/10.1093/restud/rdae007

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Cullen, J. B. (2003). The impact of fiscal incentives on student disability rates. *Journal of Public Economics*, *87*(7–8), 1557–1589. https://doi.org/10.1016/S0047-2727(01)00203-1

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996. https://doi.org/10.1257/aer.20181169

Detgen, A., Yamashita, M., Davis, B., & Wraight, S. (2011). State policies and procedures on response to intervention in the Midwest Region (Issues & Answers Report, REL 2011–No. 116). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Midwest. https://files.eric.ed.gov/fulltext/ED520327.pdf

Dhuey, E., & Lipscomb, S. (2011). Funding Special Education by Capitation: Evidence from State Finance Refor. *Education Finance and Policy*, *6*(2), 168–201.

Education Commission of the States. (2020). Response to Intervention (RTI) policies and state guidance. Education Commission of the States.

Fletcher, J. M., & Miciak, J. (2019). The Identification of Specific Learning Disabilities: A Summary of Research on Best Practices. *Grantee Submission.*

Fletcher, J. M., Coulter, W. A., Reschly, D. J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities: Some questions and answers. Annals of Dyslexia, 54(2), 304–331. https://doi.org/10.1007/s11881-004-0015-y

Fletcher, J. M., Lyon, G. R., Fuchs, L. S., & Barnes, M. A. (2007). *Learning disabilities: From identification to intervention*. Guilford Press.

Fletcher, J. M., & Vaughn, S. (2009). Response to Intervention: Preventing and Remediating Academic Difficulties. *Child Development Perspectives*, *3*(1), 30–37. https://doi.org/10.1111/j.1750-8606.2008.00072.x

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003). Responsiveness-to-Intervention: Definitions, Evidence, and Implications for the Learning Disabilities Construct. *Learning Disabilities Research and Practice*, *18*(3), 157–171. https://doi.org/10.1111/1540-5826.00072

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The Prevention, Identification, and Cognitive Determinants of Math Difficulty. *Journal of Educational Psychology*, *97*(3), 493–513. https://doi.org/10.1037/0022-0663.97.3.493

Fuchs, L. S., Fuchs, D., & Hollenbeck, K. N. (2007). Extending Responsiveness to Intervention to Mathematics at First and Third Grades. *Learning Disabilities Research & Practice*, *22*(1), 13. http://www.census.gov/Press-Release/

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-Intervention. *Journal of Learning Disabilities*, 45(3), 195–203. https://doi.org/10.1177/0022219412442150.

Gilmour, A. F., Harper, J., Lloyd, B., & Van Camp, A. (2023). Response to Intervention and Specific Learning Disability Identification: Evidence From Tennessee. *Journal of Learning Disabilities*, 00222194231215013. https://doi.org/10.1177/00222194231215013

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Hall-Mills, S. (2021). Shifting Prevalence Patterns for Special Educational Needs in the Era of Response-to-Intervention Policy. *Frontiers in Education*, 6. https://doi.org/10.3389/feduc.2021.676646

Harr-Robins, J. J., Shambaugh, L. S., & Parrish, T. (2009). The status of state-level response to intervention policies and procedures in the West Region states and five other states. *Evaluation*, *077*. http://ies.ed.gov/ncee/edlabs.

Hauerwas, L. B., Brown, R., & Scott, A. N. (2013). Specific learning disability and Response to Intervention: State-level guidance. Exceptional Children, 80(1), 101–120.

Hoover, J. J., Baca, L., Wexler-Love, E., & Saenz, L. (2008). *National Implementation of RTI: Research Summary*.

Hoover, J. J. (2010). Special education eligibility decision making in response to intervention models. *Theory Into Practice*, 49(4), 289-296.

Hughes, C. A., & Dexter, D. D. (2022). *The Use of RTI to Identify Students With Learning Disabilities: A Review of the Research*. RTI Network, The National Center for Learning Disabilities. http://www.rtinetwork.org/learn/research/use-rti-identify-students-learning-disabilities-review-research

Jimerson, S. R., Burns, M. K., & VanDerHeyden, A. M. (Eds.). (2016). Handbook of response to intervention: The science and practice of multi-tiered systems of support (2nd ed.). *New York, NY: Springer Science.*

Kavale, K. A., & Spaulding, L. S. (2008). Is Response to Intervention Good Policy for Specific Learning Disability? *Learning Disabilities Research & Practice*, 23(4), 169–179. https://doi.org/10.1111/j.1540-5826.2008.00274.x

Kovaleski, J. F., VanDerHeyden, A. M., Runge, T. J., Zirkel, P. A., & Shapiro, E. S. (2022). *The RTI approach to evaluating learning disabilities*. Guilford Publications.

Morgan, H. (2020). Misunderstood and mistreated: Students of color in special education. *Voices of Reform*, 3(2), 71-81. Retrieved from https://www.voicesofreform.com/article/18595-misunderstood-and-mistreated-students-ofcolor-in-special-education doi: 10.32623/3.100005

Morgan, Paul L., George Farkas, Marianne M. Hillemeier, and Steve Maczuga. (2012). Are minority children disproportionately represented in early intervention and early childhood special education? *Educational Researcher*, 41(9), 339-351

Morgan, Paul L., George Farkas, Marianne M. Hillemeier, Richard Mattison, Steve Maczuga, Hui Li, and Michael Cook. (2015). Minorities are disproportionately underrepresented in special education: Longitudinal evidence across five disability conditions. *Educational Researcher*, 44(5), 278-292.

Morgan, Paul L., George Farkas, Michael Cook, Natasha M. Strassfeld, Marianne M. Hillemeier, Wik Hung Pun, and Deborah L. Schussler. (2017). Are Black children disproportionately overrepresented in special education? A best-evidence synthesis. *Exceptional Children*, 83(2), 181-198.

Morgan, P. L., Staff, J., Hillemeier, M. M., Farkas, G., & Maczuga, S. (2013). Racial and ethnic disparities in adhd diagnosis from kindergarten to eighth grade. *Pediatrics*, *132*(1), 85–93. https://doi.org/10.1542/peds.2012-2390

Office of Special Education Programs. (2023). *IDEA Section 618 data: Annual reports to Congress on the implementation of IDEA*. U.S. Department of Education.

Petersen, D. (2012). Policy: Its history, intentions, and consequences for children with special health care needs. In *Handbook of Children with Special Health Care Needs* (pp. 1–22). Springer New York. https://doi.org/10.1007/978-1-4614-2335-5_1

Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. Review of Economic Studies, 90(5), 2555–2591. https://doi.org/10.1093/restud/rdad018

Rios-Avila, Fernando (2021). drdid and csdid: Doubly robust DID with multiple time periods, *Economics Virtual Symposium*, 7, *Stata Users Group*.

Reynolds, C. R., & Shaywitz, S. E. (2009). Response to Intervention: Ready or not? Or, from wait-to-fail to watch-them-fail. *School Psychology Quarterly*, *24*(2), 130–145. https://doi.org/10.1037/a0016158

Sawyer, R., Holland, D., & Detgen, A. (2008). State policies and procedures and selected local implementation practices in response to intervention in the six Southeast Region states (Issues & Answers Report, REL 2008–No. 063). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Southeast.

Stepanek, J., & Peixotto, K. (2009). Models of response to intervention in the Northwest Region states (Issues & Answers Report, REL 2009–No. 079). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northwest. https://files.eric.ed.gov/fulltext/ED506348.pdf

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

University of Kentucky Center for Poverty Research. (2022). *UKCPR National Welfare Data, 1980-2020*. Available at http://ukcpr.org/resources/national-welfare-data (Accessed April 2, 2022)

VanDerHeyden, A. M., & Witt, J. C. (2005). Quantifying the context of assessment: Capturing the effect of base rates on teacher referral and a problem-solving model of identification. *School Psychology Review*, 34,161−183.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a Response to Intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, *45*(2), 225–256. https://doi.org/10.1016/j.jsp.2006.11.004

Zirkel, P. A. (2017). RTI and other approaches to SLD identification under the IDEA: A legal update. Learning Disability Quarterly, 40(3), 165–173.

Wanzek, J., & Vaughn, S. (2007). Research-Based Implications From Extensive Early Reading Interventions. *School Psychology Review*, *36*(4), 541–561. https://doi.org/10.1080/02796015.2007.12087917

Wanzek, J., & Vaughn, S. (2011). Is a three-tier reading intervention model associated with reduced placement in special education? *Remedial and Special Education*, *32*(2), 167–175. https://doi.org/10.1177/0741932510361267

Zirkel, P. A., & Thomas, L. B. (2010a). State Laws and Guidelines for Implementing RTI. *TEACHING Exceptional Children*, *43*(1), 60–73. https://doi.org/10.1177/004005991004300107

Zirkel, P. A., & Thomas, L. B. (2010b). State Laws for RTI: An Updated Snapshot. *TEACHING Exceptional Children*, *42*(3), 56–63. https://doi.org/10.1177/004005991004200306

Tables

Table 1
Baseline Descriptive Statistics in 2006

	RTI sample		Non-RTI sample		
	Mean	SD	Mean	SD	Mean *Diff*
Outcome Measures					
% of students with SLD	0.06	0.01	0.05	0.01	0.01
% of students with SLI	0.02	0.01	0.02	0.01	0.00
% of students with disabilities	0.13	0.02	0.13	0.02	0.00
Covariates					
% White	0.78	0.11	0.74	0.17	0.04
% Hispanic	0.11	0.09	0.11	0.12	0.00
% Black	0.11	0.10	0.11	0.13	0.00
% Other	0.11	0.06	0.16	0.15	-0.04
% In school	0.83	0.01	0.83	0.02	0.00
% Female	0.49	0.01	0.49	0.01	0.00
% Enrolled in public school	0.70	0.03	0.70	0.04	0.00
% Enrolled in pre-K and K-12	0.87	0.02	0.87	0.03	0.00
% High-school dropout	0.01	0.00	0.01	0.00	0.00
Personal income, ACS (2018 dollars)	6413.35	680.24	6551.94	920.00	-138.59
Family income (2018 dollars)	84122.37	8372.84	89963.47	18677.43	-5841.10
% of students in poverty	0.12	0.03	0.12	0.03	-0.01
% Rural	0.23	0.18	0.22	0.20	0.01
Total enrollment	996944.53	742447.46	954490.94	1288960.69	42453.59
Pupil-teacher ratio	16.02	2.82	14.76	2.10	1.26
Per pupil expenditure (2018 dollars)	13177.14	1962.92	14655.28	3659.34	-1478.14
% of students received free- or reduced-price lunch	0.40	0.09	0.39	0.11	0.01
% ELL students	0.07	0.04	0.07	0.05	0.00
State population	6222432.80	4928055.34	5695650.56	7212950.48	526782.24
% SNAP recipients	0.10	0.04	0.09	0.03	0.01
Gross state product per capita (millions, 2018 dollars)	0.04	0.01	0.05	0.02	0.00
Unemployment rate	4.29	0.79	4.52	1.07	-0.24
Poverty Rate	11.98	2.44	11.80	3.48	0.18
State *N*	15		36		

Notes. Baseline characteristics measured in 2006, comparing the fifteen jurisdictions that had adopted RTI by that year with the remaining thirty-six. Outcome measures are proportions of enrolled students. Group differences were tested with two-sample t-tests; none is statistically significant at $p < .05$. The largest difference, in the pupil-teacher ratio, has $p = .084$.

Table 2

Panel A. CS Average Treatment Effects of State RTI Adoption on Identification

Outcome	ATT	SE	p	95% CI
% SLD	0.157	0.187	.401	[−0.210, 0.524]
% SLI	0.005	0.066	.940	[−0.124, 0.134]
% any disability	0.114	0.233	.624	[−0.343, 0.571]
% non-SLD disability	−0.043	0.169	.799	[−0.374, 0.288]

Panel B. SLD Estimates by Comparison Group and Covariate Set

Specification	ATT	SE	p	N
Not-yet-treated, no covariates (primary)	0.157	0.187	.401	816
Not-yet-treated, covariates	−0.008	0.114	.942	791
Never-treated, no covariates	−0.010	0.239	.965	816
Never-treated, covariates	−0.416	0.111	< .001	756

Panel C. SLD Estimates, Cohort-Specific Group-Time Average Treatment Effects

Cohort	States	ATT	SE	p	95% CI
2006	15	0.310	0.209	.139	[−0.100, 0.720]
2007	2	−0.603	0.392	.124	[−1.371, 0.164]
2008	1	−0.829	0.126	< .001	[−1.076, −0.583]
2009	5	0.548	0.361	.129	[−0.160, 1.256]
2010	5	0.027	0.235	.908	[−0.433, 0.487]
2011	4	−0.125	0.241	.605	[−0.596, 0.347]
2012	5	0.228	0.195	.241	[−0.153, 0.610]
2013	1	0.483	0.138	< .001	[0.214, 0.753]
2015	6	−0.149	0.161	.355	[−0.465, 0.167]
2017	2	−0.037	0.097	.702	[−0.227, 0.153]
Group average	46	0.123	0.165	.458	[−0.201, 0.447]

Panel A *Notes.* Estimates in percentage points. CS estimator with not-yet-treated comparison states, no covariates, 2000–2019.

Panel B *Notes.* Estimates in percentage points; CS estimator, SLD outcome, 2004–2019. Covariates in the adjusted rows are SNAP and TANF receipt, gross state product per capita, and CPI-adjusted personal income per capita; the English learner share (used in Table C1) is excluded because its limited coverage (604 of 816 state-years) would make the cells non-comparable, and its inclusion leaves the pattern unchanged. N varies across cells because adjusted rows drop state-years with missing covariates, and differs from Table C1 because the CS and OLS estimators define the sample differently.

Panel C *Notes.* Estimates in percentage points; CS estimator, not-yet-treated comparison, no covariates, 2000–2019. Cohorts are defined by adoption year; all ten in the panel are shown (none adopted in 2014 or 2016), with sizes in parentheses summing to 46 adopting jurisdictions. The 2008 and 2013 cohorts each contain one state (Missouri and Mississippi), so those estimates describe a single state. The group average (0.123) is the unweighted mean across cohorts and differs from the state-weighted simple ATT in Table 2 Panel A (0.157).

Figures

Figure 1
Number of States by RTI Policy Type over Years

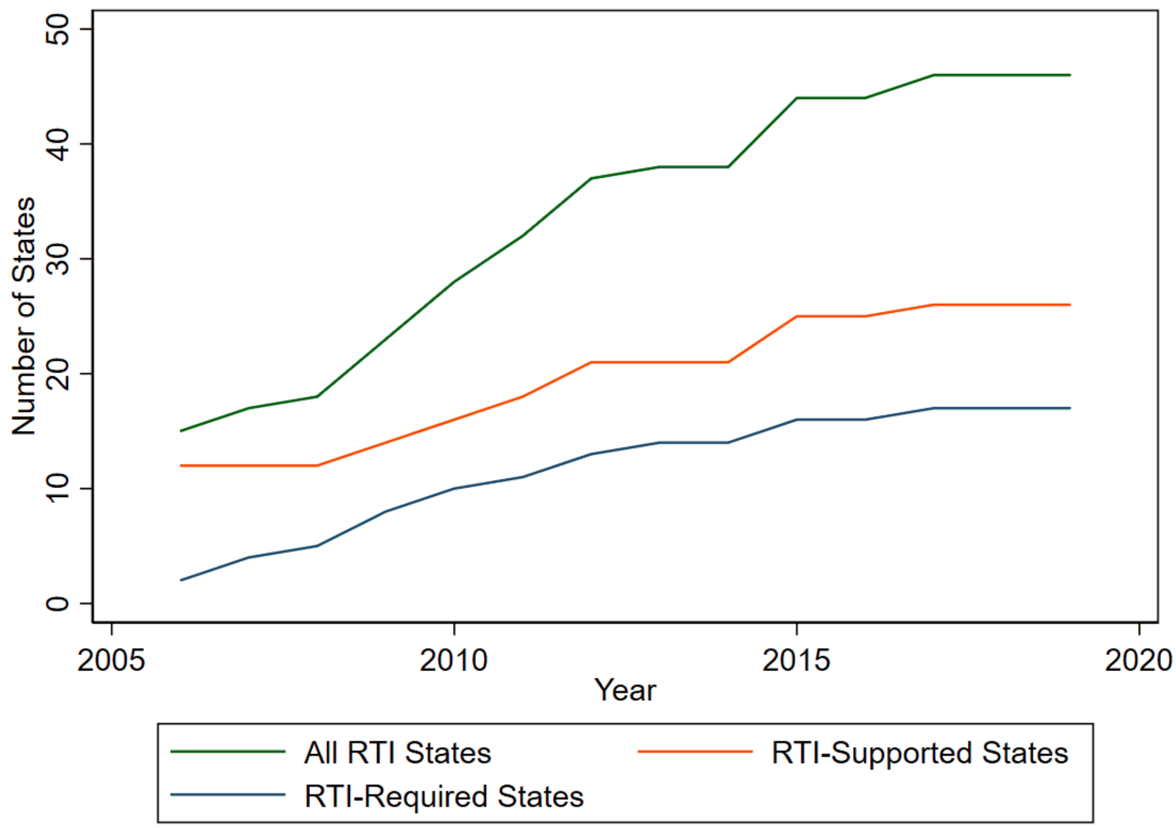
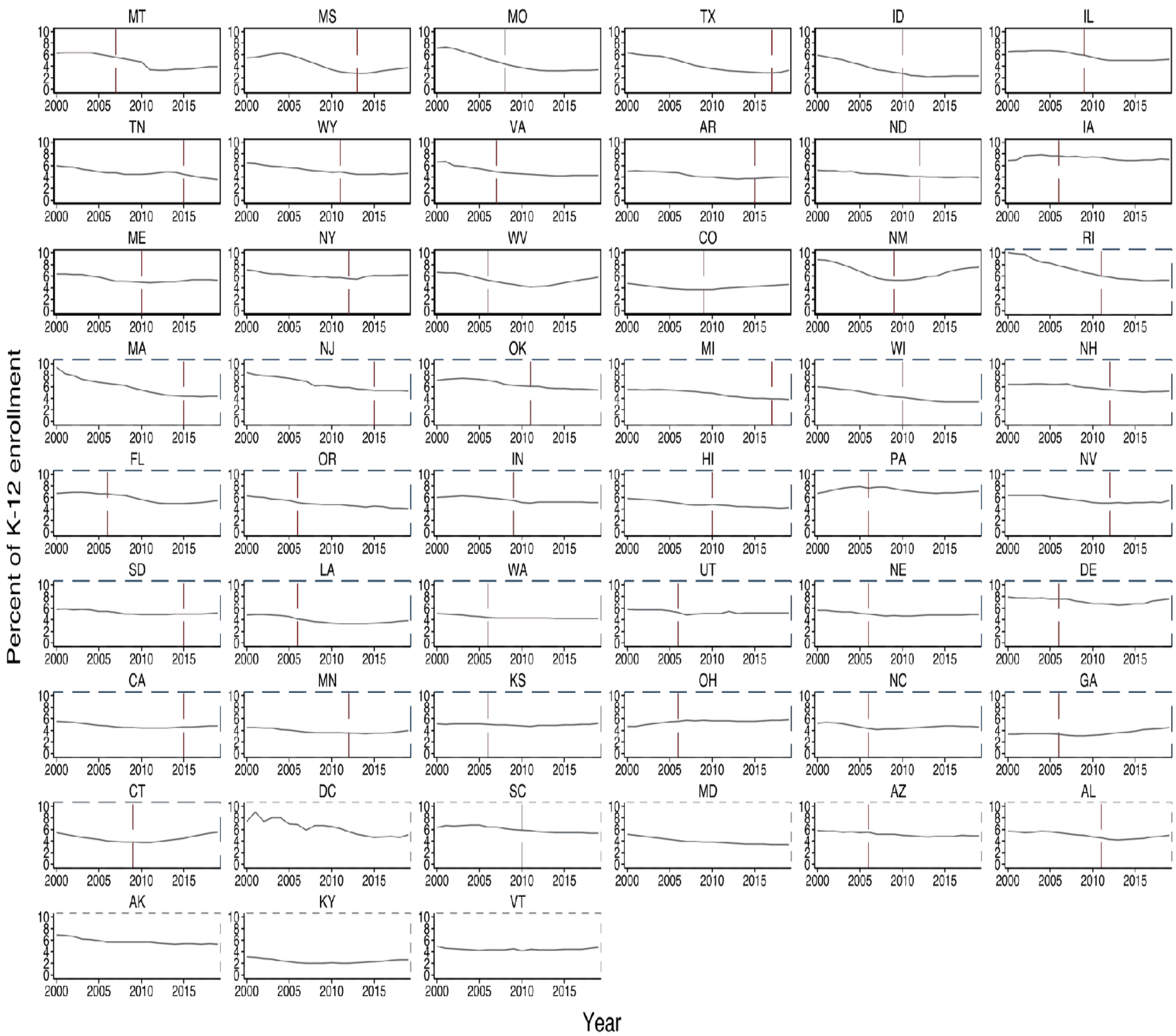


Figure 2

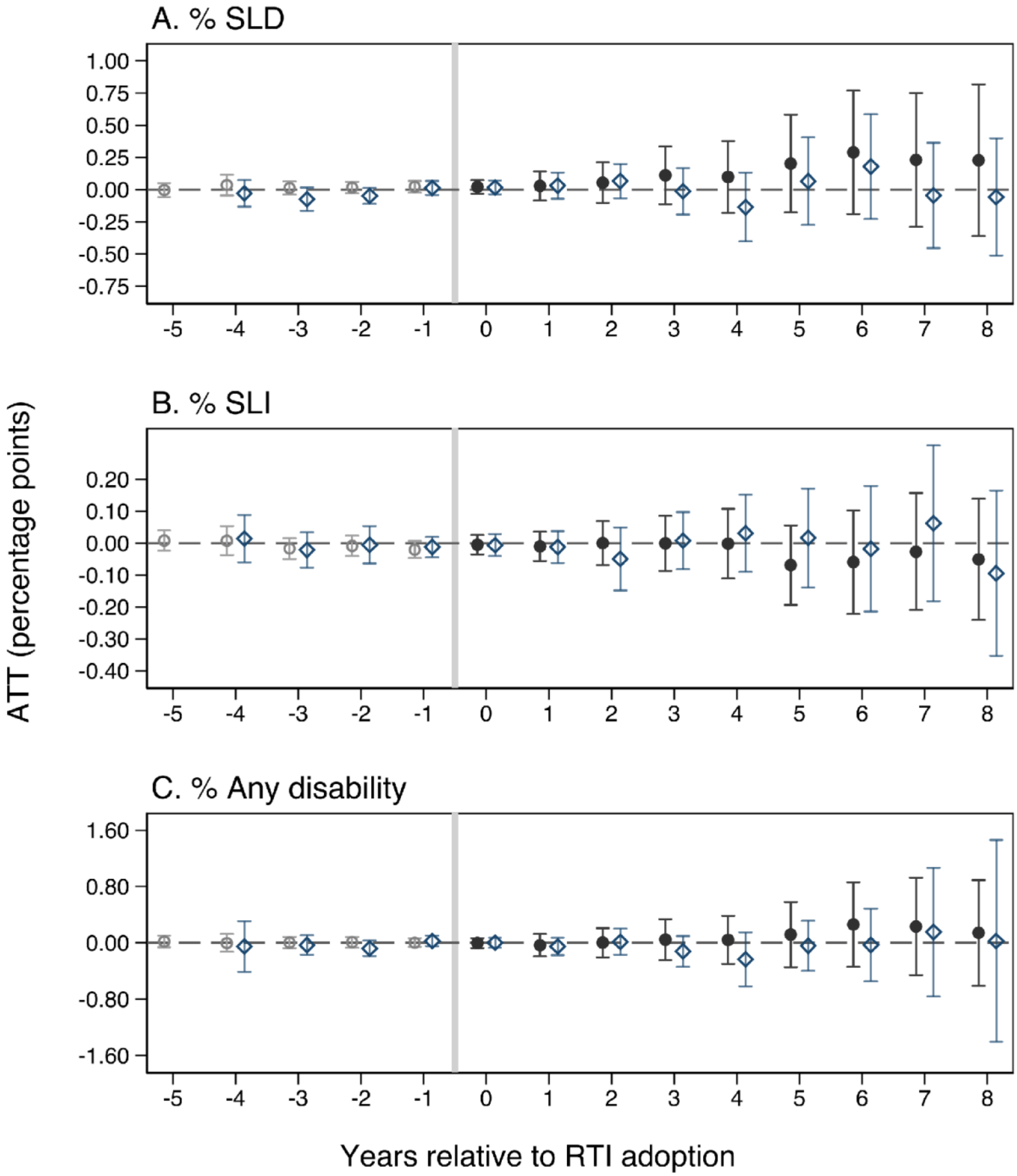
Percentage of Students Identified with SLD (K–12) by State and RTI Policy Type, 2000–2019



Notes. Solid black = RTI-required, dashed blue = RTI-supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Red dashed vertical line marks RTI adoption; non-RTI states have no adoption line. Series are observed values only; no trend is fitted or extrapolated in this figure.

Figure 3

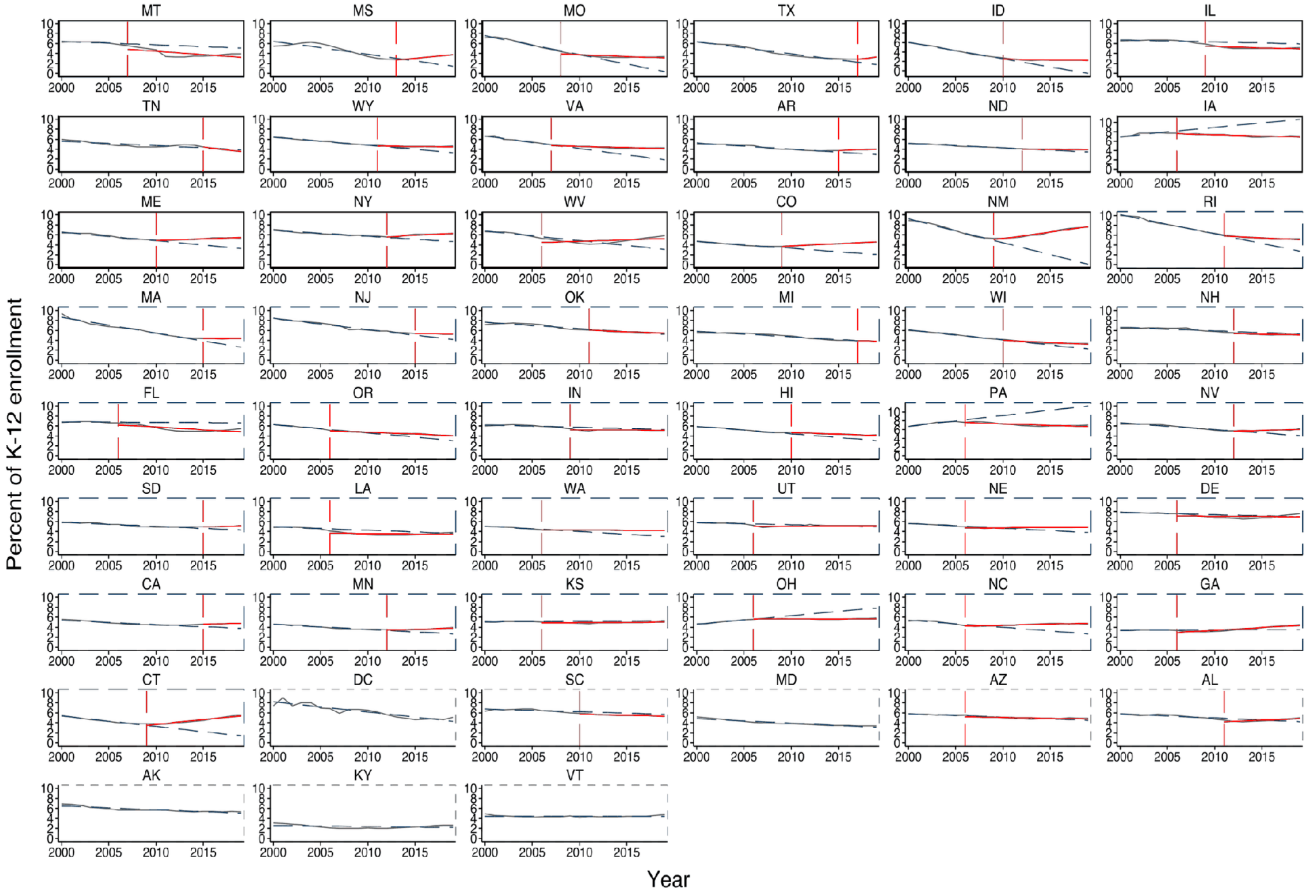
Dynamic Effects of RTI Adoption on Identification, With and Without Covariate Adjustment



Notes. CS estimator with not-yet-treated comparison states. Grey circles = no covariates, 2000–2019; navy diamonds = covariate-adjusted, 2004–2019. Hollow markers are pre-adoption placebo estimates. Bars are 95% confidence intervals, clustered by state. Panels use different vertical scales.

Figure 4

Interrupted Time Series Estimates of % SLD by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI-supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Interior lines: grey = observed, blue dashed = pre-adoption trend extrapolated, red = post-adoption trend. Red dashed vertical line marks RTI adoption.

Impact of States' Adoption of Response to Intervention (RTI) on Disability Identification

Online Supplemental Materials

These materials accompany the article "Impact of States' Adoption of Response to Intervention (RTI) on Disability Identification."

Appendix A

Table A1

RTI Start Year by State and Switching Status between 2006 and 2017

Early Switched States	**Year**	**Later Switched States**	**Year**	**Never-on States**
Arizona	2006	Hawaii	2010	Alaska
Delaware	2006	Maine	2010	Kentucky
Florida	2006	South Carolina	2010	DC
Georgia	2006	Wisconsin	2010	Maryland
Iowa	2006	Idaho	2010	Vermont
Kansas	2006	Alabama	2011	
Nebraska	2006	Rhode Island	2011	
North Carolina	2006	Oklahoma	2011	
Ohio	2006	Wyoming	2011	
Pennsylvania	2006	Minnesota	2012	
Louisiana	2006	Nevada	2012	
Oregon	2006	North Dakota	2012	
Washington	2006	New York	2012	
West Virginia	2006	New Hampshire	2012	
Utah	2006	Mississippi	2013	
Montana	2007	Arkansas	2015	
Virginia	2007	New Jersey	2015	
Connecticut	2009	Tennessee	2015	
Missouri	2009	California	2015	
New Mexico	2009	Massachusetts	2015	
Colorado	2009	South Dakota	2015	
Illinois	2009	Michigan	2017	
Indiana	2009	Texas	2017	
Total N = 23		Total N = 23		Total N = 5

Notes. Most information for RTI adoption is from papers of Berkeley et al. (2009) and Zirkel & Thomas (2010). Since these two papers only had RTI information during the study year, we checked each state's website of the Department of Education for details and double-checked the information with the two papers.

Table A2

RTI Policy Adoption Categorizations

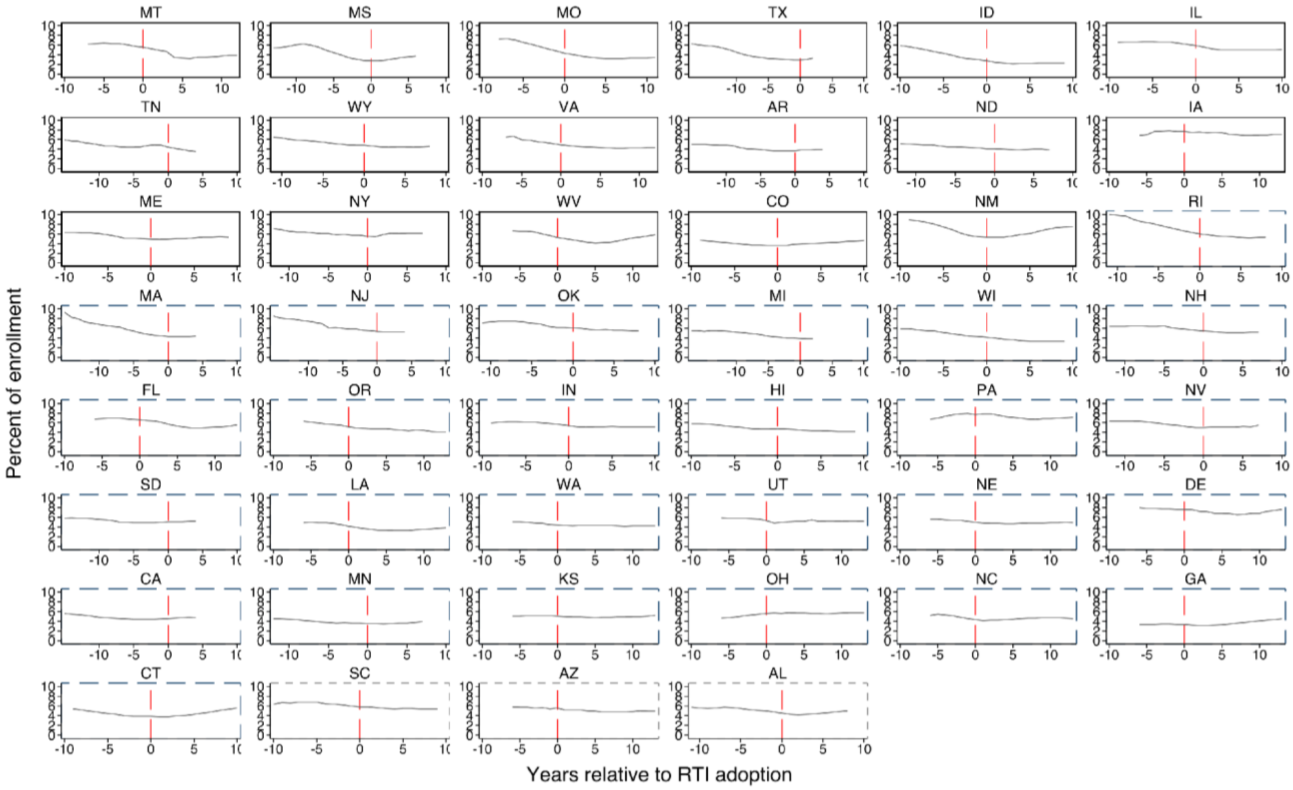
RTI-Supported States	**RTI-Required States**	**Non-RTI States**
California	Arkansas	Alabama
Connecticut	Colorado	Alaska
Delaware	Idaho	Arizona
Florida	Illinois	DC
Georgia	Iowa	Kentucky
Hawaii	Maine	Maryland
Indiana	Mississippi	South Carolina
Kansas	Missouri	Vermont
Louisiana	Montana	
Massachusetts	New Mexico	
Michigan	New York	
Minnesota	North Dakota	
Nebraska	Tennessee	
Nevada	Texas	
New Hampshire	Virginia	
New Jersey	West Virginia	
North Carolina	Wyoming	
Ohio		
Oklahoma		
Oregon		
Pennsylvania		
Rhode Island		
South Dakota		
Utah		
Washington		
Wisconsin		
Total N = 26	Total N = 17	Total N = 8

Notes. The categorization is different from Appendix Table 1, specifically for non-RTI states, which include states that did not have RTI in law.

Appendix B

Figure B1

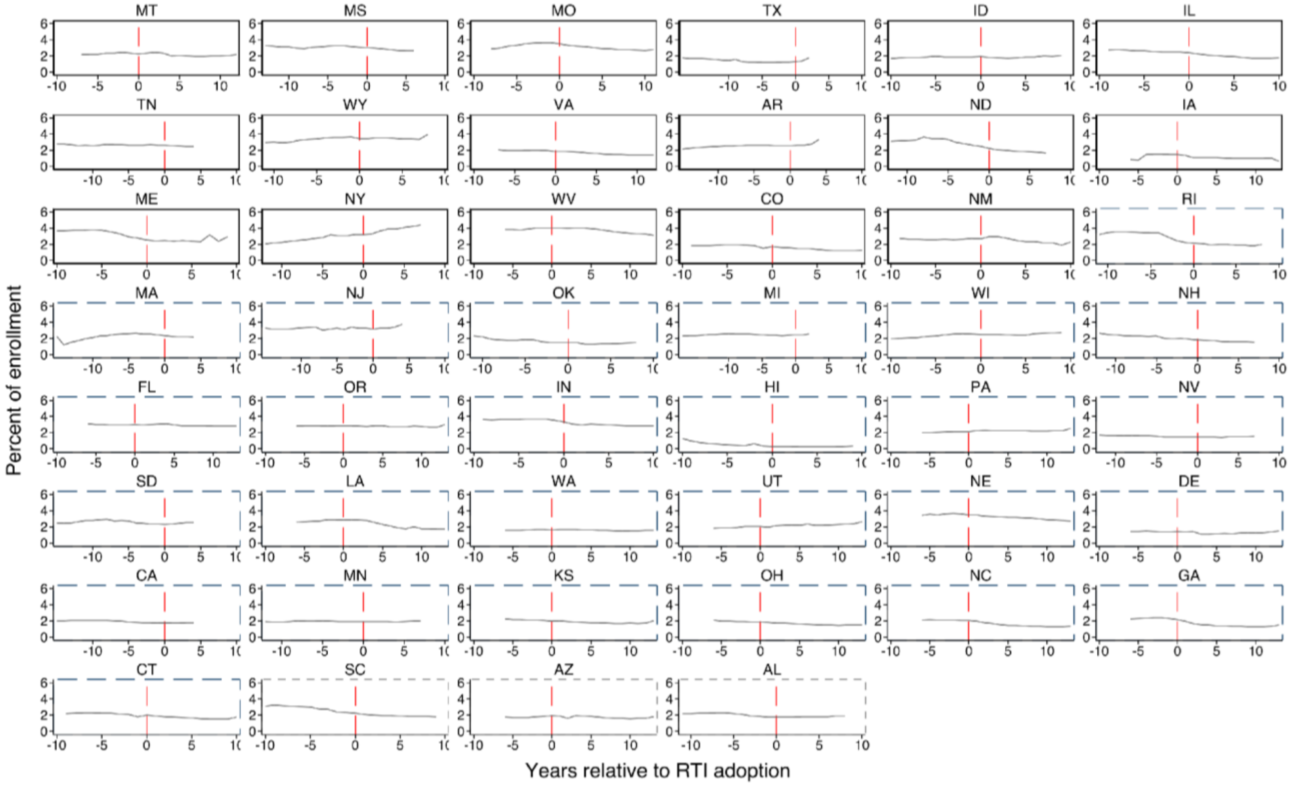
Percentage of Students Identified with SLD (K–12) in Event Time by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Red dashed vertical line marks RTI adoption year.

Figure B2

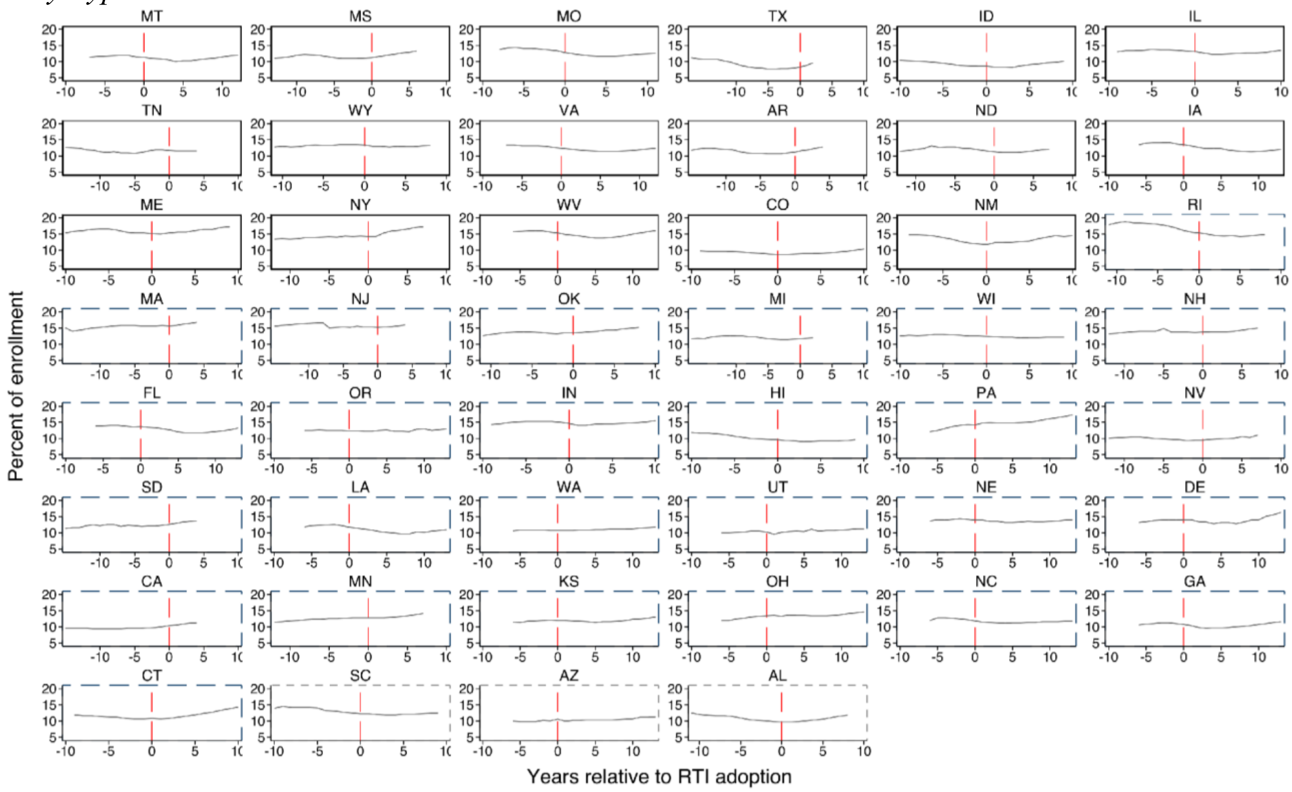
Percentage of Students Identified with SLI (K–12) in Event Time by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Red dashed vertical line marks RTI adoption year.

Figure B3

Percentage of Students Identified with Any Disability (K–12) in Event Time by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Red dashed vertical line marks RTI adoption year.

Appendix C

Two-Way Fixed Effects Estimates

As a robustness check on the CS estimates in the main text, we also estimated conventional TWFE regressions.

Two-Way Fixed Effects (TWFE). The TWFE specification is

$$Y_{st} = \delta\, RTI_{st} + X_{st}'\,\beta + \mu_s + \lambda_t + \varepsilon_{st}\,,$$

where Y_{st} is the identification rate (SLD, SLI, or any disability) in state s and year t; RTI_{st} is the adoption indicator; X_{st} is the covariate vector; μ_s and λ_t are state and year fixed effects and ε_{st} is the error term. Standard errors are clustered by state. With staggered adoption and heterogeneous, dynamic treatment effects, the estimate of δ can be biased by the use of already-treated states as controls (Baker et al., 2022; Borusyak et al., 2022; Goodman-Bacon, 2021; Sun & Abraham, 2021).

The below table presents the TWFE results examining the impact of RTI adoption on SLD, SLI, and overall disability identification rates. The analyses were conducted separately for four sets of state groupings: (1) all states, (2) states that supported RTI versus states that required RTI, (3) states that required RTI versus those without RTI, and (4) states that supported RTI versus those without RTI. These results suggest considerable heterogeneity in states' post-RTI trajectories, potentially reflecting differences in policy implementation (e.g., identification model) influencing identification practices.

Table C1.

Overall and Differential Effects of RTI Adoption on SLD, SLI, & Disability Identification Rates.

	All States	RTI-Required vs Supported States	RTI-Required vs None States	RTI-Supported vs None States
Panel A: % of students with SLD				
RTI adopted	-0.00188**	-0.00122	-0.00278*	-0.00203**
	(0.000861)	(0.000879)	(0.00141)	(0.000916)
Panel B: % of students with SLI				
RTI adopted	-0.000365	-0.000569	0.000266	-0.000690
	(0.000559)	(0.000500)	(0.000904)	(0.000668)
Panel C: % of students with all disabilities				
RTI adopted	-0.00168	-0.00147	-0.00125	-0.00232
	(0.00124)	(0.00127)	(0.00131)	(0.00154)
N	604	507	292	409

Notes. Models are estimated separately for four state groupings per outcome: all states; RTI-supported vs. RTI-required; RTI-required vs. non-RTI; and RTI-supported vs. non-RTI. All include five covariates (SNAP receipt, TANF receipt, English learner share, gross state product per capita, CPI-adjusted personal income per capita). Coefficients are proportions (×100 for percentage points); state-clustered standard errors in parentheses. Sample sizes are limited by English-learner coverage (unavailable before 2004 and after 2018; 23–25 states in 2015–2018), which reduces the all-states panel from 1,020 to 604 state-years. * $p < .10$, ** $p < .05$.

Appendix D

Age-Band Robustness: Elementary Ages (6–11)

As robustness check, we re-estimated the models with both restricted to ages 6 through 11, approximately grades 1 through 6, taking the numerator from the Section 618 counts for that band and the denominator from CDC WONDER population estimates for the same ages. Measures note that our primary outcomes divide a Section 618 child count covering ages 6 through 21 by K–12 enrollment, so numerator and denominator describe different populations. Appendix D removes that mismatch. This is the one construction in the paper in which the two populations coincide, and it is also the closest available match to the single-state evidence: Gilmour et al. (2023) estimated Tennessee's post-adoption declines on grades K through 5. Because Section 618 reports age-band detail only from 2005, the check spans 2005 through 2019 rather than the full study period. The elementary-age estimates are also more precise than the K–12 estimates for SLD: the minimum detectable effect at 80% power is 0.432 percentage points, against 0.525 in the primary specification. Figures D1 through D3 show the per-state ITS trajectories across the elementary band for SLD, SLI, and overall disability respectively. Seventeen of the forty-six jurisdictions with a coded adoption year have fewer than three usable pre-adoption observations inside the 2005 through 2019 window, and no pre-adoption trend is fitted for them; their panels show the post-adoption segment alone.

Table D1

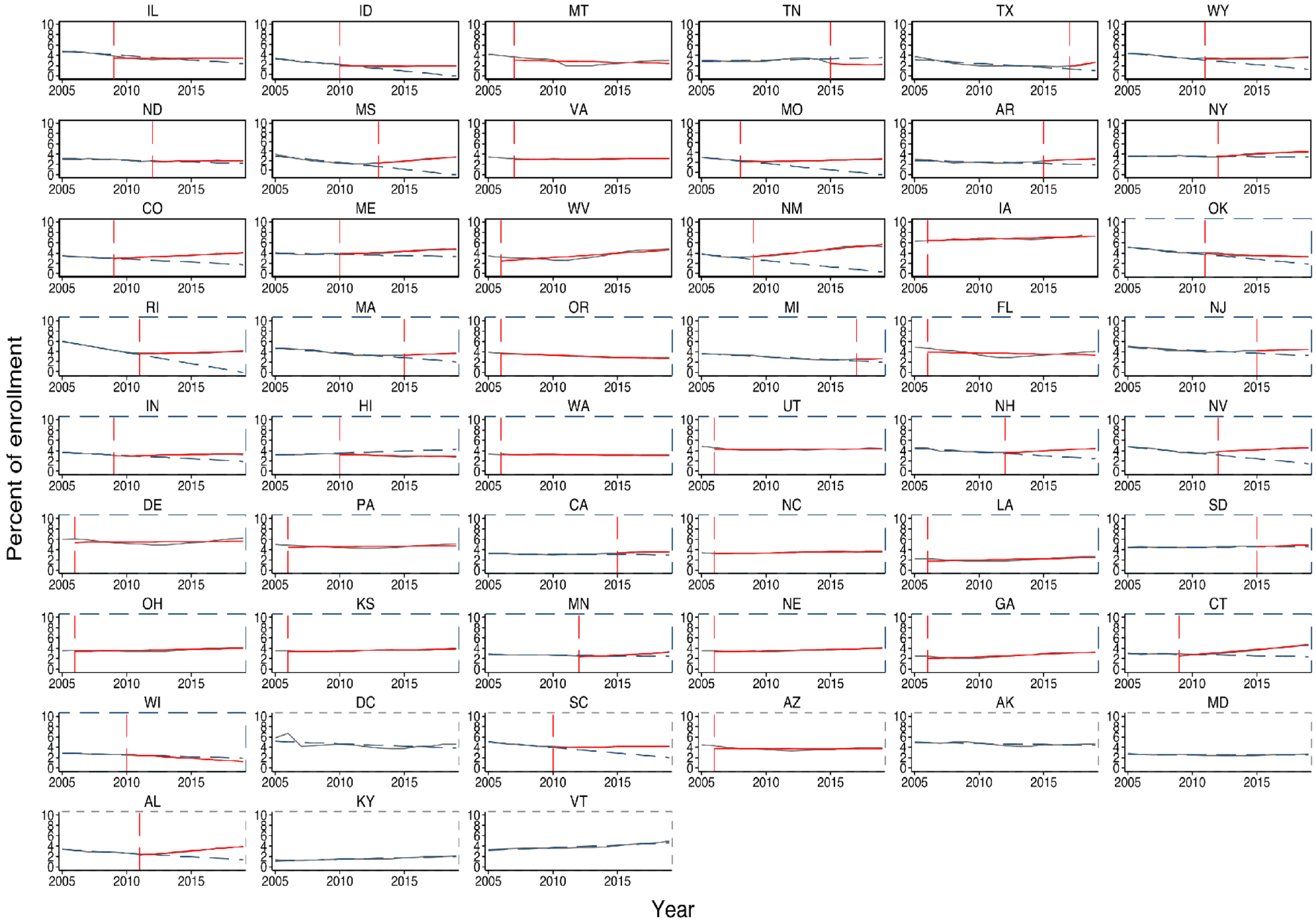
CS Average Treatment Effects of State RTI Adoption on Identification, Ages 6 Through 11

Outcome	ATT	SE	p	95% CI
% SLD, ages 6–11	0.145	0.154	.346	[−0.157, 0.447]
% SLI, ages 6–11	−0.092	0.107	.389	[−0.301, 0.117]
% any disability, ages 6–11	−0.058	0.232	.804	[−0.513, 0.398]

Notes. Estimates in percentage points. CS estimator with not-yet-treated comparison states, no covariates, 2005–2019, standard errors clustered by state. Numerator and denominator are restricted to the same age band. Minimum detectable effects at 80% power are 0.432 for SLD, 0.300 for SLI, and 0.650 for any disability at ages 6–11, against 0.525, 0.184, and 0.653 in the K–12 specification

Figure D1.

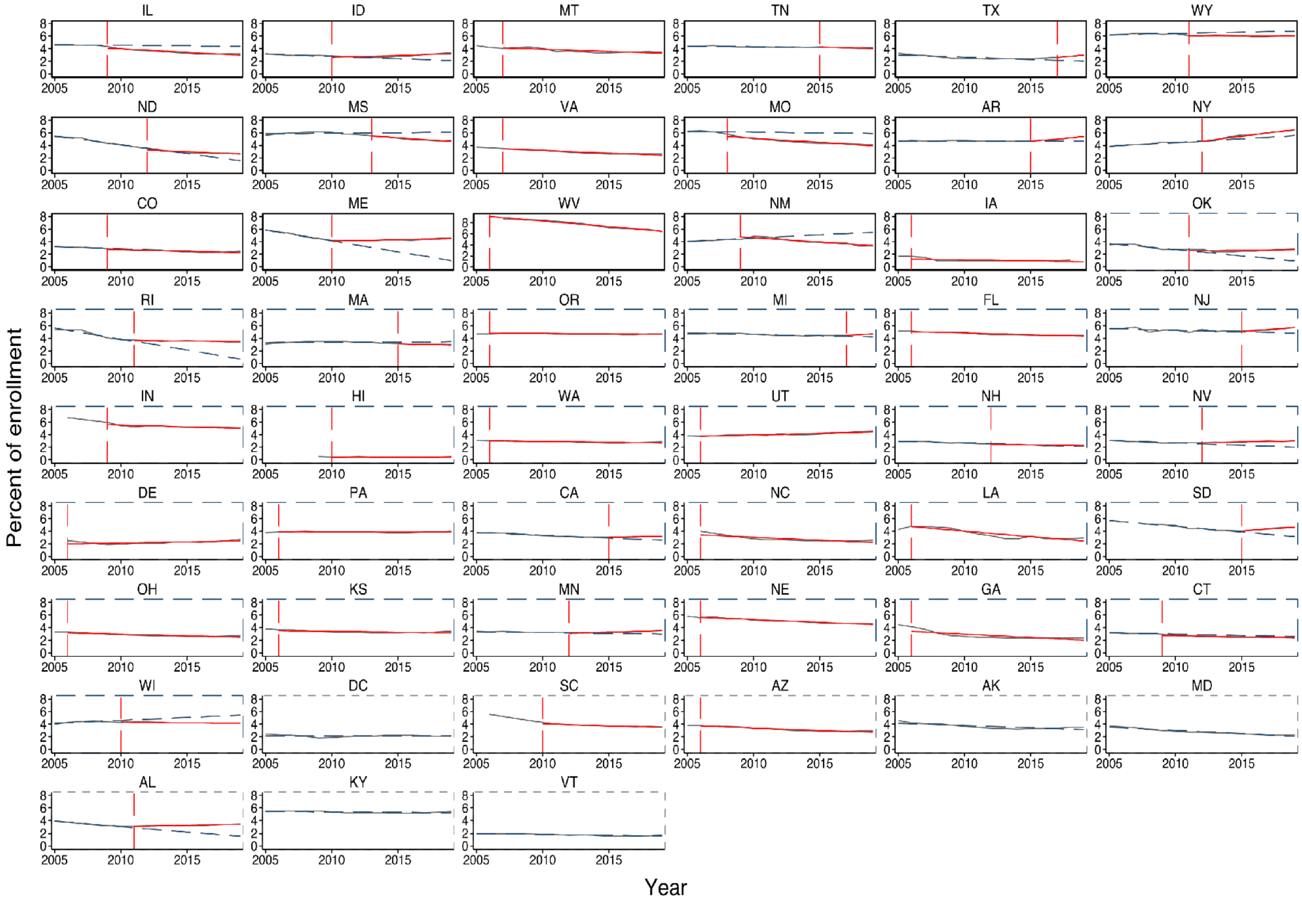
Interrupted Time Series Estimates of % SLD (Ages 6–11) by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI-supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Interior lines: grey = observed, blue dashed = pre-adoption trend extrapolated, red = post-adoption trend. Red dashed vertical line marks RTI adoption.

Figure D2

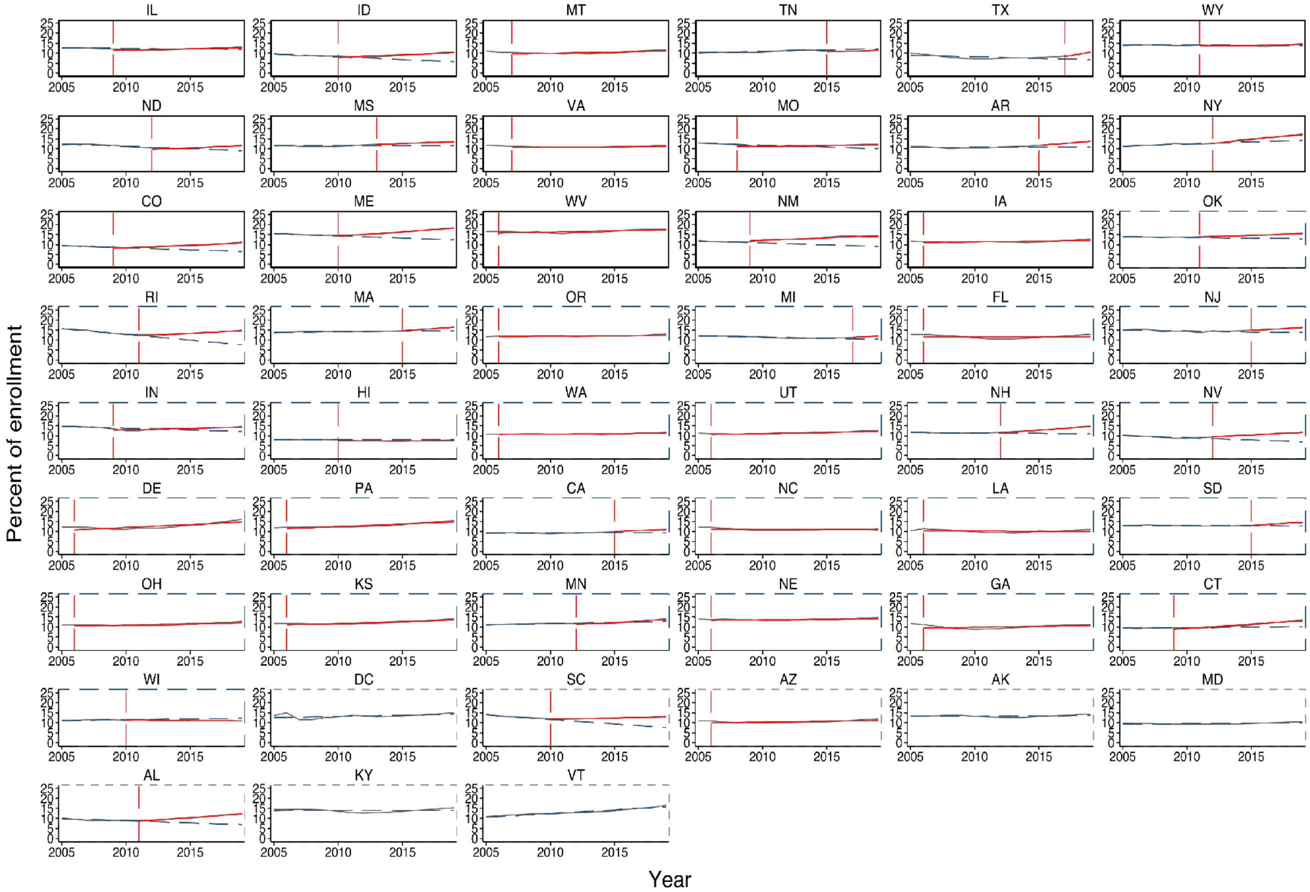
Interrupted Time Series Estimates of % SLI (Ages 6–11) by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI-supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Interior lines: grey = observed, blue dashed = pre-adoption trend extrapolated, red = post-adoption trend. Red dashed vertical line marks RTI adoption. Seventeen states have fewer than three usable pre-adoption observations in the 2005–2019 window and show the post-adoption segment alone.

Figure D3

Interrupted Time Series Estimates of % Any Disability (Ages 6–11) by State and RTI Policy Type



Notes. Solid black = RTI-required, dashed blue = RTI-supported, dotted grey = non-RTI states. States are grouped by policy type, ordered within group by steepness of decline. Interior lines: grey = observed, blue dashed = pre-adoption trend extrapolated, red = post-adoption trend. Red dashed vertical line marks RTI adoption. Seventeen states have fewer than three usable pre-adoption observations in the 2005–2019 window and show the post-adoption segment alone.

Appendix E

Adoption-Strength Coding Sensitivity

The required-versus-supported contrast reported in Results depends on the adoption-strength coding in Table A2, which two coders built independently from state statutes, regulations, administrative codes and eligibility guidance, giving precedence to states' own statutory language over secondary characterizations. Published 50-state reviews nonetheless classify a number of states differently, and because a reader may compare Table A2 against them, we re-estimated the contrast under two codings drawn from those reviews. Under the reconciled coding, required states are the seventeen that Hauerwas et al. (2013, Table 2) place in a row requiring RTI as of October 2011, plus North Carolina and Tennessee, which Zirkel (2017) names as the most recent additions to the mandate group; all other adopting states are supported and the non-RTI group is unchanged. Fourteen jurisdictions change group, in both directions. The consensus coding retains only the states the two codings assign identically, dropping the fifteen disputed cases, so misclassification within the contrast is zero by construction. The specification is otherwise unchanged across the three rows of Table E1: Callaway–Sant'Anna group-time ATT on percent SLD, 2000–2019, not-yet-treated comparisons within a sample restricted to the focal strength group plus the five never-adopting jurisdictions, standard errors clustered by state. Only group membership differs. The point estimate is not stable — it changes sign between the published coding and both alternatives — so we draw no inference from its direction. The standard error is almost invariant across the three, and no coding produces a detectable difference, including the consensus coding in which misclassification cannot be responsible. We therefore report the strength contrast as uninformative about a gap below roughly one percentage point rather than as evidence that no gap exists.

Table E1

RQ3 Required-Versus-Supported Contrast Under Three Adoption-Strength Codings

Coding	Required / supported	Difference (SE)	p	MDE at 80% power
Published (Table A2)	17 / 26	−0.182 (0.326)	.575	0.912
Reconciled	19 / 24	0.084 (0.327)	.797	0.916
Consensus	10 / 18	0.095 (0.350)	.787	0.981

Notes. Estimates in percentage points. Callaway–Sant'Anna group-time ATT on percent SLD, not-yet-treated comparisons, no covariates, 2000–2019, standard errors clustered by state. Group ATTs, with standard errors in parentheses, are −0.008 (0.258) for required and 0.175 (0.198) for supported under the published coding; 0.148 (0.264) and 0.064 (0.193) under the reconciled coding; and 0.267 (0.296) and 0.172 (0.187) under the consensus coding. The difference is required minus supported; its standard error treats the two group ATTs as independent, which is approximate because they share the never-adopter controls. MDE is the minimum detectable effect on the difference at 80% power (2.802 × SE). The published coding is Table A2; the reconciled coding follows Hauerwas et al. (2013) and Zirkel (2017); the consensus coding retains