Balancing Federalism: The Impact of Decentralizing School Decision Making*

Eric A. Hanushek, Patricia Saenz-Armstrong, and Alejandra Salazar

August 8, 2024

Abstract

Education policy in the United States, while primarily the responsibility of the state governments, involves complicated decision making at the local, state, and federal levels. Federal involvement dramatically increased under the No Child Left Behind Act of 2001 (NCLB). But, reflecting resistance to various parts of this law, the involvement of federal policy making was substantially reduced when Congress passed the Every Student Succeeds Act (ESSA) in 2015. This change in policy allows estimation of the impact of altered federalism. By looking at how states reacted to their enhanced decision-making role, we see a retreat from the use of output-based policy toward teachers, and this retreat was associated with significantly lower student achievement growth. The snapshot of federalism impacts here is a lower bound on the effects as more states will very likely react to the flexibility of ESSA and as more school districts change their teacher force.

* We benefited from the overall support of the National Council on Teacher Quality and the assistance of Kelli Lakis and Lisa Staresina in developing the cross-walk between the elements of teacher policies and the provisions of NCLB and ESSA .

Introduction

Most discussions of the rise and fall of the *No Child Left Behind Act of 2001* (NCLB) focus on how federal policy affected school operations and student performance, but there is a deeper aspect to this legislative change. The United States like a number of other democracies has historically maintained a complicated system of school governance with individual states assuming primary responsibility of both funding and operation of the schools. But NCLB moved important aspects of school policy up from the states to the federal level. The broad issue of how the federal structure affects educational outcomes has never been analyzed. Importantly, the replacement of NCLB with *Every Student Succeeds Act* (ESSA) provides a window through which various implications of federalism itself can be measured and assessed.

NCLB combined both more rigorous test-based accountability for schools and a substantial increase in the role of the federal government in education policy. While many states had already introduced some form of test-based school accountability, NCLB created a rigid structure that applied to all states regardless of whether they had their own approach. Over time, NCLB became more and more unpopular, leading the U.S. Congress to change the federal role substantially in 2016 (McGuinn (2016)). ESSA retained the general idea of test-based accountability but sent responsibility for its design and implementation back to the states. This reversion to more decentralized policies matches the overall primacy of states in educational policy but leads to natural questions about the implications for student learning.

Subsidiarity indicates that policy responsibility should go to the lowest level of government that can accomplish the intended purpose (Oates (1972, 1999)). This decentralization principle suggests that the move from federal to state decision making could enhance efficiency and provide better educational outputs. But the application of this principle becomes complicated when concerns such as externalities and political preferences are introduced.

An early hint of the impact of this policy change comes from tracking the pattern of student scores over time. Much attention has been given to the learning losses from the COVID pandemic, which can readily be seen by comparing reading and math scores between 2020 and 2022 (Figure 1). But it is instructive to trace the pattern of scores on the National Assessment of Educational Progress (NAEP) over a longer period. While the pattern of scores before 2000 differed between reading and math and has no common explanation, scores in both achievement domains rise

significantly from 2000-2012 before falling significantly prior to the pandemic period. The overall pattern after 2000 is strikingly coincidental with the implementation of NCLB and the subsequent move to ESSA.

Previous evaluations of NCLB have addressed the general question of whether test-based accountability is effective in raising student outcomes. Fully evaluating the impact of federal accountability policy is difficult because both NCLB and ESSA were implemented across all states at the same time. The most persuasive analyses of NCLB either investigate the impact of state use of test-based accountability before NCLB (Carnoy and Loeb (2002), Hanushek and Raymond (2005)) or compare the results across states that already had test-based accountability at the introduction of NCLB to those that did not have such a system (Dee and Jacob (2011)).

These studies find a net-positive impact of the test-based accountability surrounding NCLB, but that is not the issue when looking at the move from NCLB to ESSA. ESSA retains the idea of test-based accountability including the regular testing of students in grades 3-8. The largest change with ESSA is allowing states much more flexibility in how they structure the provision of schooling, i.e., a rebalancing from federal policy to state policy in the design and implementation of local schooling. And the impact of this altered federalist structure depends both on the behavioral decisions of states about school regulations and incentives and on the impacts of these decisions on the operations of local schools.

We pursue a two-stage approach to our evaluation of the impact of altered decision making embodied in the move from NCLB to ESSA. In the first stage we focus on the key area of teacher policy and assess changes that states made with the flexibility offered by the altered federal statute. We then evaluate how these changes affect student output. The combination of these two components provides a lower-bound estimate of the impact of this major change in federalism.

Following this evaluation strategy for an exhaustive set of potential state education policies is not feasible. NCLB, ESSA, and associated policies altered many elements of federal educational requirements and undoubtedly played out in a variety of specific state regulations and laws designed to achieve the desired student learning outputs.[1]

[1] NCLB and ESSA are much broader than just the accountability sections, being just a part of the main components of federal involvement in education. Moreover, they are the result of successive re-

We pursue a more modest goal of evaluating the interaction of federal statutes with a broad suite of teacher policies including teacher certification, evaluations, and incentives. Past research has pinpointed teacher quality as the most important aspect of schools and one having strong long term learning impacts (e.g., Hanushek and Rivkin (2010), Chetty, Friedman, and Rockoff (2014), Koedel, Mihaly, and Rockoff (2015)). This suite of policies was arguably the locus of the most consequential changes in going from NCLB to ESSA in terms of loosening the pressure on states and schools.

Issues of teacher quality received increasing emphasis over time under NCLB. While NCLB originally had requirements about "highly qualified teachers" (HQT) that focused on background characteristics of teachers and not their performance, this was modified over time. It was, for example, central to the 2009 federal incentive program of "Race-to-the-top" that offered substantial grants to states that promised to institute certain teacher accountability policies.[2] It was also explicitly included in the granting of flexibility under NCLB as pressures against the rigidity of NCLB arose.[3] But use of student achievement in teacher evaluations also became increasingly controversial (e.g., Baker et al. (2010)), Koretz (2009)). Dissatisfaction with the teacher evaluation components of student testing both provided motivation for ESSA and contributed to explicitly pulling back from the set of federal requirements related to test-based teacher evaluation, turning these issues back to the states.

We first identify a set of output-based policies related to teacher accountability that were included in the NCLB legislation but that were relaxed in the ESSA legislation. We also identify a set of input-based teacher policies that were unchanged by the legislation and that act as "control policies" that could have happened in the absence of the changed federal laws. From this taxonomy, we then construct a mapping of state changes in these policies coincident with the move from NCLB to ESSA. The matrix of specific teacher evaluation and accountability policies builds on the extensive longitudinal database of the National Center for Teacher Quality (NCTQ). The resulting extract of the multiple components of teacher policies provides a rich picture of state reactions to the relaxation of federal restrictions on their behavior.

---

authorizations of the Elementary and Secondary Education Act of 1965 (ESEA). ESEA signified in 1965 the start of substantial federal involvement in both funding and policy for public schools (Cross (2014)).

[2] Highly variable grants were made between 2010-2013. Grants were given to 18 states in three different phases. See https://en.wikipedia.org/wiki/Race_to_the_Top [accessed August 16, 2023]

[3] https://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html

We match this information on output-based and input-based teacher policies with changes in the growth of state achievement using the National Assessment of Educational Progress (NAEP). This panel covers achievement changes from just before and just after the federal legislative change, allowing us to assess whether the teacher policy responses that we observe are also important in determining learning outputs.

We find that the movement from NCLB and Race-to-the-top to ESSA led overall to large, systematic, and swift movement away from components of teacher policy that emphasize student outputs. On the other hand, components that emphasized background and certification of teachers but not performance in the classroom did not show such uniform changes. Some states increased these latter requirements for teachers while others decreased them. In other words, given the change in decision making from federal to state government, the clearest pattern is that states tended to pull away from the NCLB output-based policies.

When we link these teacher policy changes to state growth in student achievement between fourth and eighth grade, we find that strong output-based teacher policies are associated with greater student achievement gains in both math and reading. On the other hand, the set of input-based policies are associated with lower state achievement gains.

While the focus on differential achievement growth rules out the most significant confounders, it remains possible that unmeasured but coincidental other factors enter into the estimated impact parameters. The near-universal finding in the literature is, however, that other measurable characteristics of schools do not systematically characterize the effectiveness of schools.[4] This suggests that our estimates provide a plausible path for establishing a lower-bound on the impact of the changed decision making structure.

Combining changes in state teacher policy with estimates of their impact on achievement, we find that the shift in federalism contained in ESSA was associated with a small but significant fall in student achievement growth. We find that the altered locus of decision making on net hurt student achievement growth. In other words, the policy responses of the states when given more latitude in state teacher policies led to actions that were harmful to student achievement.

[4] Hanushek and Rivkin (2012). See also the debate about the impacts of funding: Handel and Hanushek (2023, 2024), Jackson and Mackevicius (2024).

Note that the focus of this analysis is not the impact of specific forms of teacher evaluation or personnel policies but the response of the educational system to the new regime implemented through a change in the locus of decision making.[5] The "treatment" that we are interested in is the provision of the new institutional structure within which states make varying decisions. As such, it represents a unique analysis of how changes in decision making authority within the federal system affect the performance of government.

Impacts are likely to grow over time. In order for the state policies to have an impact on students, local districts must change their policies and procedures, and the teacher force must also adjust because of the district actions. This implies inertia to the change in federalism with ESSA necessarily limits the ability to judge the full impact within our four-year observation period following the enactment of ESSA. Because of the incomplete adjustments of states and districts that we observe, we interpret our estimates as a lower bound on the modified federalism associated with the lessened federal presence.

The Move from NCLB to ESSA

Education is not mentioned in the U.S. Constitution, leaving primary responsibility for both policy and funding of schools up to each state. This perspective was, for example, reinforced by prior rulings of the U.S. Supreme Court in school finance matters.[6] At the same time, the federal government in a variety of other policy regimes seeks to achieve minimum national standards and to promote equity by providing incentives to states for their adoption. NCLB, adopted in 2001, follows this latter line by conditioning receipt of federal compensatory education funds (Title 1) on meeting the requirements of the law.

---

[5] Previous analyses have looked at specific policy initiatives either within or across states. These policies are components of the changed locus of decision making that we consider here. Taylor (2023) provides a detailed analysis of the many components of teacher evaluation and personnel policies that have be implemented in different states. Kraft, Brunner, Dougherty, and Schwegman (2020) investigate the impact of new teacher evaluation policies on teacher supply and also consider broader teacher personnel policies. Bleiberg et al. (2023), which we return to below, provide an analysis of the impact of introducing new teacher evaluation systems across the states on student achievement.

[6] The lawsuit of *San Antonio Independent School District v. Rodriguez*, 411 U.S. 1 (1973) concerned uneven school funding in Texas under the property tax and was brought under the 14th amendment to the U.S. Constitution. The U.S. Supreme Court, citing the absence of any discussion of education in the Constitution, ruled that education was not a fundamental right under the U.S. Constitution. Thus, the Texas funding formula could stand because it had a rational basis (see Hanushek and Lindseth (2009)).

NCLB was a very complicated law that introduced several components that had little precedent in federal education policy. It required each state to develop a test-based accountability system, it set a target of all students' meeting proficiency targets by the 2013-14 school year, and it defined a set of remedial actions for any school not on a path to meet the student achievement goals. It was supposed to be re-authorized in 2007, at which time the most problematic features could presumably be remedied, but Congress never was able to re-authorize it.

Ultimately, the unrealistic goal of having all students reach proficiency led to the broad consensus that NCLB had to be replaced. The Congress, which had not been able to re-authorize NBLB, sought compromise legislation that could lead to re-authorizing the Elementary and Secondary School Act, the basic authorization that not only housed the federal accountability rules but also the fundamental parts of all federal policy toward K-12 education. The Every Student Succeeds Act was voted into law in 2015. ESSA retained the idea of test-based accountability but turned decision making on major design and institutional features back to the states.

For our purposes, it is useful to put NCLB and the move to ESSA into the framework of national-state federalism. In many ways the overall structure of NCLB violated basic federalism principles. It charged each state with developing its own educational standards, testing, and achievement goals. Yet, because many students end up working in other states, the level of human capital produced in one state has economic implications for other states (Hanushek, Ruhose, and Woessmann (2017a, 2017b)) – implying that educational requirements for students might better be determined by the demands of the national labor market. On the other hand, the federal specification of how to change the schools when local schools failed to meet Adequate Yearly Progress (AYP)[7] directly conflicts with the principle of subsidiarity. The local school almost certainly has a better idea than the federal government of the demands and capacities of the local school.

While keeping student testing, ESSA loosened the use and reporting requirements for accountability. Thus, it did not address the externalities of state-defined education goals. But,

[7] Adequate Yearly Progress established intermediate performance targets that would lead each school to reach 100 percent proficiency by the 2013-14 school year. These targets applied to the whole school and to subgroups defined by racial/ethnic groups, poverty, and more. AYP was used to judge performance of schools, and failure to meet those requirements led to different required actions of each school.

ESSA also eliminated the anachronistic element of having the federal government specify how to remediate failing schools, returning the operational decisions to the states. Thus, ESSA partially but not completely fell more in line with basic principles of federalism.

A major component of our investigation of the changed decision making revolves around teacher policies. Consideration of the roles and evaluation of teachers has been a continuous part of the discussion about federal involvement in education policy. From early on, school personnel were rightfully concerned that the test results developed under NCLB might be used to assess the performance of individual teachers. Because the accountability systems focused on status measures, or the level of performance, the available scores necessarily conflated family and neighborhood factors with the impacts of schools and teachers.[8] Thus, student performance as measured would partially reflect educational inputs that were beyond the control of the teacher and that therefore arguably should not be central to the evaluation of the teacher.

The idea of employing the existing testing regimes for teacher evaluations was, however, elevated in policy and legislative circles with the development of the "Race-to-the-top" program (RTTT) in 2009 under President Obama (Duncan (2018)). As an extension of federal involvement in school accountability, the Race-to-the-top program was a competitive grant program at the state level, where states were requested to enter a competition for funds. The guidelines included a variety of elements for the state grants, but the two most important were adoption of the Common Core curriculum and the use of student achievement growth measures for teacher evaluations. By moving to achievement growth, the largest problems of inappropriate attribution of performance to the teacher were ameliorated if not eliminated, but the overall controversy over approaches and use teacher evaluations remained.

Importantly for our assessment of the changes in federalism, ESSA retained ideas of student testing and test-based accountability but quite consistently gave development of teacher evaluations and teacher policies back to the states. The behavior of the different states with this new flexibility and the impact of these choices on student outcomes are central to our evaluation.

[8] The U.S. Department of Education did offer a number of waivers for parts of NCLB including allowing the use of "growth models"; see, for example, Polikoff, McEachin, Wrabel, and Duque (2014), Derthick and Rotherham (2013).

State Responses to ESSA

The initial task is the identification and cataloging of changes in teacher policies that states made with the expanded decision-making authority under ESSA. Teacher policy is of course complex, involving a range of specific components, and states emphasize different rules, regulations, and procedures. We begin by tracing the use of a range of important teacher policies documented in the taxonomy of policies developed by the National Council on Teacher Quality (NCTQ).[9] We then confirm our interpretation of these measures as reflecting important elements of federalism by developing the crosswalk with NCLB, RTT, and ESSA.

We distinguish between teacher policies that are directly linked to effectiveness in the classroom and those that emphasize background and activities that do not call for inclusion of information about student outcomes. Conceptually, the test-based evaluation and accountability of NCLB are best seen as a principal-agent problem (Figlio and Loeb (2011)). When interested parties – from parents to policy makers – have difficulty in monitoring what teachers and schools are doing, output-oriented accountability can provide a mechanism to get the schools to work toward their desired results instead of the interests of the teachers and schools. Input-based approaches on the other hand generally presume a direct linkage of specific characteristics to classroom performance. As such, they require a detailed understanding of the mechanisms related to school effectiveness, something that appears largely beyond our current understanding and capacity in many educational areas.

But the simple appeal of output-oriented accountability is also misleading. Given the complexity of both desirable goals of interested parties and the operations of schools, this solution to the principal-agent problem can itself be complicated and prone to undesirable results.[10] Thus, the subsequent step in the evaluation is validating the relationship between these conceptual arguments and school outcomes.

The NCTQ teacher policy data for each year were assigned to one set or the other of these policy sets: outcome-based policies or input-based policies. This information was extracted and

[9] See, for example, https://www.nctq.org/publications/2017-State-Teacher-Policy-Yearbook and various other years.

[10] The most obvious concern is partial observability of relevant outcomes (Holmstrom and Milgrom (1991)), and this was central to many arguments against such test-based accountability. As a simple example, the NCLB focus on reading and math achievement in grades 3-8 leaves out all other subjects and grades along with outcomes other than achievement. Issues of noncognitive skills also enter.

compiled for the years 2011, 2013, 2015, 2017, and 2019 and then put into a state-by-year database. Considerable effort is, however, required to harmonize the questions and coding of teacher requirements over time, because both varied with new data collections by NCTQ. The resultant policy matrix allows us to understand the trends in these policies for each state.[11]

The key to the measurement of output-based policies (Table 1) is that there is explicit mention of teacher effectiveness as measured by student assessments. As can be seen, effectiveness ratings can be included at a variety of points from regular evaluation and tenure to dismissal. Importantly, while these policies generally require teacher evaluations based on student achievement, they go significantly beyond just having such an evaluation system. These policies go into detail on a broad range on uses of any evaluations and into the basis for more general personnel policies.

The input-based policies (Table 2) represent a more varied set of policies that generally involve actions that bear no clear relationship to observed and measured effectiveness. While pursuing policies that are generally plausible on historical or prima facie grounds, they lack explicit attachment to and emphasis on measured student outcomes. Such policies are generally motivated by "quality considerations" but are not keyed to what effectiveness means. Since these policies tend neither to preclude the use of information on student performance nor to mandate it in decision making, there is potential ambiguity in their classification and the actual application in different states. The classification of performance pay underscores such ambiguity (and is the subject of subsequent sensitivity analysis), but the long-standing inconsistency of such policies leads to the initial classification as an input-based policy.[12]

Policies in NCLB and ESSA

These various teacher policies are generally related directly to provisions of NCLB and ESSA, and those linkages support our analysis of how changes in federalism play out in terms of student results. The actual impact of federal statutes on schools and students of course depends first on the reactions of states, because the federal government cannot directly impose policies on states. The federal statutes provide guidelines, and the key to any effect of these comes through subsequent actions that states took after ESSA in comparison to those during NCLB.

[11] The resultant teacher evaluation matrix will be made publicly available.
[12] See the early assessment of merit pay by Cohen and Murnane (1985, 1986) for skepticism about the basis for merit pay and the evaluation of the relationship to impact on student achievement in Dee and Keys (2004).

NCLB included specific provisions for the development and implementation of output-based teacher policies, as documented in Appendix A. It called for more objective teacher and principal evaluation systems that included measurable student progress, as well as effectiveness-based tenure systems.[13] NCLB also included incentives for the collection and reporting of teacher effectiveness data including federal funding termination if, upon federal evaluation, there was no significant progress towards attaining student achievement goals. In contrast, ESSA leaves many of those measurements and policy decisions to the states' discretion. Although it includes provisions for the state to ensure that disadvantaged students are not disproportionately served by ineffective teachers, it is also explicit in noting that this shall not be construed as a requirement for the states to develop or implement teacher evaluation systems. Nor is there a requirement for the states to collect or report any data that the states are not already reporting as of the day of the enactment of ESSA.

Both NCLB and ESSA consider some input-related policies including mentions of differential pay and incentives for recruitment and retention of teachers in high-need subjects or schools, but ESSA generally does so without reference to consideration of actual student performance. Other input-related policies are largely absent from both pieces of legislation and implicitly left to the states.

Pre- and post- ESSA trends in state policies

After the enactment of ESSA, much of the momentum behind adopting and implementing more rigorous educator evaluation systems ground to a halt. Incentives were not in place for state departments of education or state school boards to address the issue of teacher accountability. The ability to sustain output-based evaluation and accountability policies is clearly affected by the political balance between state authorities and various interest groups including unions.[14]

[13] NCLB did have requirements for having "highly qualified teachers" (HQT), a requirement generally interpreted as common certification requirements such as having a bachelor's degree, passing exams if required, and so forth. This requirement, which was not based on student outcomes, was quite different from the main sections of NCLB. Race-to-the-Top and the flexibility that was introduced in NCLB over time added the perspective of the central role of student performance and value-added. Further, HQT was more an issue for school district teacher assignment policies and did not change the overall set of laws and regulations of the states. There is little evidence that it ever had much impact on district or state behavior.

[14] Jha, Banerjee, and Moller (2020)

Table 3 provides data on the prevalence of the individual input-based and output-based policies. These data by number of states with each provision are shown graphically in Figures 2 and 3. It is not surprising that in light of the withdrawing of federal mandates in favor of more local control, output-based policies have seen a retreat across the country. Uniformly the output-based policies rise over the NCLB/RTTT period, but then they consistently fall after the replacement with ESSA (Table 3, panel A, and Figure 2). The fall shows the withdrawal of significant numbers of states. The dramatic rise in requirements for including a significant component of student achievement growth is rather quickly reversed as ESSA comes into effect.

Each of the specific components shown in Figure 2 follow a similar inverted-V pattern of state usage over time. Across the output-based specific policies, the most prevalent is the requirement to use student growth evaluations in various ways, a component seen explicitly in NCLB/RTTT requirements.

Interestingly, when we turn to input-based teacher policies we do not see the same pattern (Table 3, panel B, and Figure 3). Except for offering more pay for advanced degrees, the largest number of states with the input-based requirements is consistently most prevalent in 2019, although the general movements remain rather small.

These data show the reactions of states to the change in federalism embodied by the move from NCLB to ESSA. They do not, however, indicate what impact on achievement these changes might have had. That is the subject of the next section.

Teacher Policies and Achievement Growth

In order to assess the overall impact of the altered federal-state balance and the subsequent state policy changes, we first relate the various teacher policies to growth in student achievement across states. This exercise allows us to estimate the impact of particular components that vary following the changed locus of decision making.

Achievement Growth

While the environment and structure of teaching is determined by many complex factors, overall teacher policies are largely determined at the state level. We presume states have an overriding objective of improving student outcomes, and they pursue alternative regulations and

policies that they believe will improve these outcomes. Interestingly, states make quite different decisions on the package of teacher policies, making it difficult to isolate individual elements.

We focus on the specific policies described above, but in order to understand the impact of the change in federal decision making policies, we begin with a generic representation of the education production process.

The achievement of a student i in grade g in state s and year t (A^{s}_{igt}) is:

$$A^{s}_{igt} = \rho_{s} + X^{s}_{igt}\Gamma + \varepsilon^{s}_{igt} \tag{1}$$

where $\mathbf{X}^{s}_{igt}$ is a vector of cumulative family inputs and school inputs (which might include peers, neighborhood factors, etc.) and ε^{s}_{igt} is an error term. ρ_{s} is a state-specific intercept that is the aggregation of fixed state differences in preferences, policies, and other factors entering into education in state s.

If we average across the students in each state and consider how the student performs in grade g conditional on prior performance in grade g*, we can substitute into Eq. 1 and write the average growth in achievement from g* to g as:

$$\underset{g\rightarrow g^{*}}{\Delta\overline{A}^{s}_{t}} = \underset{g\rightarrow g^{*}}{\Delta X^{s}_{t}\Gamma} + \underset{g\rightarrow g^{*}}{\Delta\overline{\varepsilon}^{s}_{t}} \tag{2}$$

The term $\Delta\overline{\mathbf{X}}^{s}_{t}$ is simply the flow of average school and family inputs over grades g* to g, measured at time t. In this formulation, the state-specific institutional and historic factors (ρ_{s}) drop out.

Our objective is to estimate how the change in federalism contained in moving from NCLB to ESSA affects the outcomes. In order to do that, we focus on the major change in policy of going from the significant emphasis on state policies that employed output-based measures of students (T^{s}_{t}) in the evaluation and management of the teacher corps under NCLB and Race-to-the-Top to allowing states freedom to design the teacher personnel policies. Note, however, that the federal policies did not emphasize different teacher input-based policies ($\tilde{T}^{s}_{t}$) built around background and experiences as opposed to student outputs.

We decompose the major components of $\mathbf{X}^{s}_{igt}$ in order to isolate the role of federal policy changes. We diverge slightly from past work on educational production functions. Although we highlight teachers as distinct from other inputs, we consider directly the two elements of teacher

policy (I_t^s and $\tilde{T}_t^s$). This separation also reflects the important aspect of federal policies derived from its treatment of teacher inputs.

$$\Delta\overline{X}_t^s = \gamma_1 T_t^s + \gamma_2 \tilde{T}_t^s + \gamma_3 S_t^s + \gamma_4 F_t^s \quad (3)$$

where S includes school inputs except for that coming through teacher policies, and F is the input of families.

Substituting Eq. 3 into Eq. 2 yields our estimation equation. Once we have the underlying parameters relating teacher policies to achievement growth (γ_1 and γ_2), we can apply them to the changes in policies related to the changed federal statutes previously identified and can estimate the impact of altered federalism on student outcomes.

$$\underset{g\to g^*}{\Delta\overline{A}_t^s} = \gamma_1 T_t^s + \gamma_2 \tilde{T}_t^s + \gamma_3 S_t^s + \gamma_4 F_t^s + \Delta\overline{\varepsilon}_t^s \quad (4)$$

Panel data construction

Data from the National Assessment of Educational Progress (NAEP) provide the basic measures of achievement that allow us to compare states. NAEP is often referred to as "the nation's report card." It is designed to provide consistent achievement measures that can be compared across states and across time.[15] Students are sampled and tested at two-year intervals, and state participation is mandatory since enactment of NCLB. Sample sizes are sufficient to allow state-level reporting of results, although not all demographic groups can be reported for each state because of the underlying distributions of students in each state.[16]

We specifically consider the assessments in reading and math for grades 4 and 8. While NAEP does not provide longitudinal data for individual students, it provides representative data for the student populations of each state at different times from 1990 to 2019.[17] We construct our

[15] For details on the NAEP sample design, see https://nces.ed.gov/nationsreportcard/tdw/sample_design/.
[16] For reporting requirements, see https://nces.ed.gov/nationsreportcard/tdw/analysis/summary_rules_minimum.aspx.
[17] NAEP comes in different versions. We use the Main NAEP data that is designed to assess state performance. It is grade-based and focused on math and reading in grades 4 and 8. The Long Term Trend (LTT) NAEP starts in the 1970s and is designed to produce comparable scores over time by keeping the same assessment framework. LTT NAEP provides national data but not state data and is age-based. These data were shown in Figure 1.

measure of average achievement growth in each state (i.e., $\underset{g*}{\Delta}\bar{A}_t^s$) by comparing grade 4 scores in math or reading to grade 8 scores four years later, i.e., by following the same cohort in each state.

Our analysis focuses on two cohorts – students in each state in grade 8 in 2015 and in 2019. This provides achievement growth for the cohort in school at the end of NCLB and for the cohort in school at the beginning of ESSA. Both the 8th grade and the 4th grade test scores are normalized by the national mean and standard for each test as of 2015. Instead of constraining the coefficient on fourth grade achievement to be -1 as implied in Equation 2, the subsequent statistical analysis puts prior achievement on the right hand side of the equation, which allows for depreciation or growth of the earlier achievement.

The estimation approach eliminates all inputs that are constant for each state over this period. This design clearly deals with the main family inputs which just slowly move over time for states. We do, however, include a measure of parental education (either average years of schooling or percent with a college degree) and of state expenditure per student in the estimation in order to capture any potentially important dynamic differences in *S* or *F*. Expenditure is measured on a per-pupil basis averaged for 2011-2015 and 2015-2019.(U.S. Department of Education (2022).

To capture the flow inputs that go into the achievement growth, we aggregate the data on teacher policies to the prevalence of each teacher policy during the last four years of NCLB and the first four years of ESSA. Table 4 displays the descriptive statistics on state policies, and it readily shows how the states changed policies that were relevant to the two cohorts.

The state averages, however, mask the underlying changes across states. Figure 4 (output-based) and Figure 5 (Input-based) show how states changed the emphasis on the two teacher policies over the study period. Moving from the NCLB period to the ESSA period, significant numbers of states backed away from output-based policies with just Texas moving more toward such policies. On the other hand, with input-based policies there is no simple pattern with some states moving toward these policies and some moving away from them.

Econometric analysis of state level achievement growth

Table 5 provides a clear picture of the relationship between teacher policies and the growth in achievement. We include estimates with pooled math and reading growth and with the subjects separately.[18]

The top rows of the table show the impact of the control variables: an indicator for ESSA, a subject indicator, and the measures of family (*F*) and school (*S*). Parental education attainment of the state has a strong and significant impact on achievement growth, but spending differences have a small negative and mostly insignificant effect on growth. Lagged achievement unsurprisingly has a strong effect on eighth grade performance with a coefficient that is very consistent with that from microdata estimates (Hanushek and Rivkin (2012)).

The central focus of these estimates is the impact of teacher policies, and there is a consistent pattern. The set of output-based policies systematically are associated with higher growth in achievement and are statistically significant. On the other hand, input-based policies are invariably harmful to growth, although the estimates are not always statistically significant. The findings for input-based policies are consistent with prior estimates of educational production functions (Hanushek (2003)), while the findings for output-based policies are relatively new.[19]

The separate math and reading models are very similar with one surprising exception. Outcome based teacher policies have a surprisingly stronger impact on reading growth than on math growth. A consistent finding of educational production functions and of estimates of teacher value-added has been the schools and teachers have a stronger effect on math outcomes (Hanushek and Rivkin (2012), Bacher-Hicks and Koedel (2023)).

We will return below to the issue of the magnitude of effects in terms of federalism, but we benchmark the estimates by considering the impact of adding one additional output-based policy. Going from no such policies to having all of the identified policies implies an increase in achievement growth of 0.2 s.d. (math) to 0.3 s.d. (reading). With the seven identified policies, adding one more explicit policy would then raise achievement growth by 0.03-0.04 s.d.

---

[18] All estimates are GLS assuming a random-effects model across states.

[19] Taylor (2023) reviews and analyzes the range of impact studies of teacher evaluation and personnel policies that have been conducted over time. Bleiberg et al. (2023) provide estimates of the impact of differing teacher evaluation policies across states and time in an event-study format, which we discuss below.

An alternative to the aggregate policy estimates is to disaggregate the separate elements of teacher policies. Unfortunately, with the limited cross-state variation and the correlations across specific teacher policies, we are unable to disentangle the separate impacts. The estimates, found in Appendix Tables B1 and B2, show insignificant impacts of the separate policies although they are jointly significant.

In sum, the estimates of growth in state achievement are entirely consistent with more micro studies of educational production functions. Policies that provide incentives related to teacher effectiveness lead to better student performance, while policies that emphasize background characteristics of teachers do not.

Investigations of racial/ethnic differences

It is possible to expand the analysis by recognizing that we can follow the growth of different racial and ethnic groups across states. The NAEP data are available for Asians, Blacks, Hispanics, and Whites. We follow the same strategy as before except that we rely on the sample of achievement growth by group and we now allow for different growth rates by adding an indicator for each group.[20] (White students are the comparison group).

When we duplicate the prior analyses for the enlarged racial/ethnic sample, we find equivalent results for the impact of the different teacher policies (Table 6). The output-based teacher policies are systematically related to improved student performance while the input-based policies point to lower achievement.

We also find that for comparable state policies, Asians outperform whites who outperform Hispanics and Blacks. The relative deficits of black students are particularly striking.

It is also possible to go further by stratifying the sample into separate racial/ethnic groups, although we now confront the more limited sampling of states based on the distribution of the subgroups across the country. The stratified results show some instability of impacts across the separate subgroups (Table 7). The black and white students follow very similar patterns that are close to the aggregate estimates with positive growth effects for outcome-based teacher policies and negative (albeit not as precise) estimates for input-based policies. Asians show stronger impacts from both types of policies, although this might well reflect the more limited sampling in

[20] Because of different population size of racial/ethnic groups in varying states, NAEP data are not available for each group in each state.

just half of the states. Finally, the impacts for Hispanics are quite anomalous with a reversal signs for the teacher policies. Differences in spending per pupil are also significantly negative in their impact on student achievement growth.

Overall, the estimates by racial and ethnic subgroups provide general support for the differential impact of the teacher policies, but their increased imprecision leads to some uncertainty.

Sensitivity Analysis

When originally classifying teacher policies, we coded the generic performance-based pay policy as an input-based policy because it did not explicitly reference objective measures of effectiveness. This corresponded to the historic ambiguity in the design of such programs. Of the various policy assignments, this classification has the largest possibility for misclassification. We pursue two approaches for assessing the potential influence of an incorrect assignment. In the simplest, we investigate putting this in the output-based aggregate instead of the input-based. When we do this, we find that the estimated impact of output-based policies is smaller but still statistically significant when estimated with either the sample of all students or the separate race/ethnicity sample (Appendix Tables B3 and B4). The negative impact of input-based policies is simultaneously smaller (i.e., closer to zero) and statistically insignificant although still negative.

A clearer perspective on generic performance pay policies comes from considering its separate impact on achievement growth while removing it from both outcome-based and input-based aggregates (Table 8).[21] Performance pay by itself has a statistically insignificant albeit a negative effect on achievement growth. The estimated impacts of outcome-based and input-based remain essentially unaffected.

The estimation constrained the introduction and the withdrawal of teacher policies to have symmetric impacts on achievement. We can relax this constraint and estimate differential impacts of introduction and withdrawal. This variant (not shown) produces insignificant differences, but the number of states following each change becomes small – making it difficult to interpret these results.

[21] The estimated separate impact of performance pay on achievement growth in the context of the race/ethnicity sample yields similar results; Appendix Table B5.

Discussion

By focusing on student growth, our estimation of the impact of teacher policies across states is designed to guard against any state differences that are constant across the observation period. Thus, such things as the structure of the schools and on-going regulations and policies, as the curriculum and testing structure, and as the cultural background and preferences of the families do not influence our estimates as long as they are not changing.

The main concern in our analysis is potential time-varying factors that both affect achievement and are correlated with the teacher policies. There obviously was considerable flux in school policies over the 21st Century, although the portion of that flux that is not captured by the identified teacher policies and by spending changes is more limited. ESSA does retain a requirement for employing test-based accountability, although the states were given new authority to set reactions to different degrees of educational performance. This new authority would be problematic if, for example, states reduced school performance incentives at the same time that they reduced output-based teacher policies. However, we do not have any measures of the "quality" of state accountability programs and how that might have changed.

The closest other analysis to ours focuses specifically on teacher evaluation policies. Bleiberg et al. (2023) investigate how achievement (and attainment) was affected by the substantial movement of states to alter their teacher evaluation policies from 2009-2017. They conclude that the altered teacher evaluation policies had no national impact on student outcomes. While there are methodological differences in their approach,[22] two key factors are likely related to the difference between our small but significant impacts and their insignificant ones. First, they focus specifically on teacher evaluation policies measured generally by the introduction of some new state policy but also by a quality measure of the new policy. They do not consider whether or how such evaluations enter into the educational process. We consider a much broader set of teacher policies (Tables 1 and 2). While the output-based policies generally involve evaluation at some stage, the policies themselves are directed at a range of usage domains. The input-based policies are additionally quite separate from teacher evaluation. Second, their methodology is focused on

[22] Bleiberg et al. (2023) employ event-study and difference-in-differences methods keyed to the introduction of new teacher evaluations. Perhaps more importantly, their samples include individual districts and grades (or individual schools in the analysis of NAEP scores) even though the central policy is one that is constant for all grades and schools within each state. They measure achievement by SEDA scores combine state tests by way of NAEP performance differences across states. Their standard approach includes districts fixed effects, leaving score variations that elevate the importance of random noise.

the results after the introduction of new teacher evaluations, but they do not include any consideration of the removal of policies. (Implicitly they do have nonimplementation in the sense that they note that there is limited differentiation among teachers even with new evaluations). Thus, while there is an obvious relationship between Bleiberg et al. (2023) and our work, they are clearly focused on different policies.

Specifically, the analysis of Bleiberg et al. (2023) indicates that enhanced teacher evaluations by themselves have little overall impact on achievement. Our analysis indicated that a range of teacher policies incorporating test-based evaluations within a larger policy framework are related to better achievement.

Effect of Changed Federalism

Measured across states, output-based policies decreased by 5 percent between 2015 and 2019. The impact of this change on student achievement growth, the crux of this analysis, is presented in Table 9. These estimates combine the changed prevalence with the impact parameters in Table 5. The net impact of the reduction in output-based policies following the introduction of ESSA was 0.01-0.015 s.d. lower growth in NAEP. The change in input-based policies amounted to another reduction of 0.01 s.d. Thus, the combined impact of changes in teacher policies is 0.02-0.025 s.d.

These immediate effects are small by the standards of conventional, small-scale interventions, but, as Kraft (2020) has pointed out, judgments about small and large effects must take the context into account. In the case of our analysis of the move from NCLB to ESSA, the context is crucial.

First, this represents the average change for the entire nation and not the observed impact of a small program that has never scaled up. It combines the impacts of the states that made changes during the four years of the observed ESSA regime with the majority of states that did not change over that period. Our calculations consider the impact of a national policy on the millions of public school students.

Second, the analysis rests on the immediate policy changes made by the first set of states that reacted to the new federalism. We have aggregated policies over the 2015-2019 period, but many of these policy changes necessarily occurred closer to 2019 because of the lags inherent in legislative actions. We also do not know how many other states might subsequently alter their

teacher policies. The pandemic that began in early 2020 led both to dramatic alterations in student testing policies along with a general reluctance to make major school policy changes in the face of the overall disruptions of COVID.[23] Thus, while we do not have reliable evidence on the continuing dynamics of policy changes, we are most likely observing just a portion of the total policy effect that will evolve from the changed federalism. Unfortunately, because of the disruptive effects of the pandemic, the observed state changes after March 2020 will not provide a good picture of the complete legislative dynamics surrounding the changed policies of ESSA.

Third, and most important, the change in state policy does not in itself change the education that students see. Impacts on teaching in the classrooms that are attributed to the state policy changes involve the added inertia of the reactions of individual school districts to the changes in state law and regulation.[24] Even if districts act immediately to the policy changes, there is considerable inertia in the teaching force since the bulk of teachers would not change quickly. The overall teacher corps changes with the entrants and exits from teaching, and the effectiveness of the teacher corps thus depends on the relative effectiveness of the new versus exiting teachers and on any change of effectiveness of the remaining teachers (see Kraft, Brunner, Dougherty, and Schwegman (2020). Therefore, the impact of teacher policies that work through changing the composition of the teacher corps are only partially observed by 2019, even if no further states alter their policies.

In sum, it is very likely that the observed changes represent just a portion of the total impact of changed federalism on the teacher corps and then on students.

Conclusions

The normal economic argument for the division of decision-making authority within a federal system is that decisions should be made at the lowest level that is capable of effectively doing them. This subsidiarity principle is built on the idea that local decision makers have a better understanding of the needs and capacities of their citizenry and governments, thus leading to more

[23] Interestingly, by our estimates, a portion of learning loss (e.g., Hanushek (2023), Lewis and Kuhfeld (2023) may have been related to the inability of apply the output-based policies that proved to be related to student outcomes.

[24] Indeed, Bleiberg et al. (2023) argue that one possible explanation for their null findings for teacher evaluation is a failure of districts to adopt meaningful evaluation systems, i.e., to incorporate state policies more fully in local district operations.

efficient results. Of course, such gains are potentially offset by limited decision-making capacity, political bargaining, economies of scale, and externalities.

The changing role of the federal government in education policy offers a chance to investigate the trade-offs in allocation of decision making across levels of government. With NCLB, the federal government assumed a much larger role in education decision making than it had before 2002. It pursued a national accountability policy but one that lost public and legislative support over time. It was replaced in 2016 by ESSA, a statute that turned much of educational decision making back to the states.

This paper addresses the net effect on student outcomes of these major changes in educational federalism. While it is not possible to cover all of the elements included in the federal statutes, it is possible to trace the impact of the law change on teacher policies. This focus is arguably the most important because of past research showing that teacher quality is the most important element of a high-quality school. By analyzing teacher policies generally covered by NCLB as related to an output-based focus, we see that the movement to the states with ESSA led to a lowering of mandates in this area and the subsequent alterations in state policies. On the other hand, input-based policies are not central to either NCLB or ESSA, and their pattern of change is less consistent across states.

By looking at student achievement growth across states, we show that output-based teacher policies are significantly related to achievement growth while input-based policies are negatively related (although less likely to be statistically significant). These results hold for both math and reading, and the models explain a majority of the state differences in achievement growth.

When we combine the policy choices followed by the states after ESSA came into effect with the estimated impacts from the production function estimates, we find a small but significant negative impact of the state policy choices. The impacts are small by the standards of small-scale individual programs – 0.02-0.025 standard deviations – but those standards are not appropriate in this case. We are looking at a program of national scale, but one that has not been fully completed. The estimates here reflect the immediate reactions of first-responder states and partial reactions of affected local school districts. The full impact is likely to grow larger as more states move from output-based to input-based teacher policies and as the policies work through individual district adjustments in their teacher forces.

References

Bacher-Hicks, Andrew, and Cory Koedel. 2023. "Estimation and interpretation of teacher value added in research applications." In *Handbook of the Economics of Education, Vol 6*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann: Elsevier: 93-134.

Baker, Eva L., Paul E. Barton, Linda Darling-Hammond, Edward Haertel, Helen F. Ladd, Robert L. Linn, Diane Ravitch, Richard Rothstein, Richard J. Shavelson, and Lorrie A. Shepard. 2010. *Problems with the use of student test scores to evaluate teachers*. Washington, DC: Economic Policy Institute (August 27).

Bleiberg, Joshua, Eric Brunner, Erica Harbatkin, Matthew A. Kraft, and Matthew G. Springer. 2023. "Taking Teacher Evaluation to Scale: The Effect of State Reforms on Achievement and Attainment." NBER Working Paper Series No. 30995. Cambridge, MA: National Bureau of Economic Research (March).

Carnoy, Martin, and Susanna Loeb. 2002. "Does external accountability affect student outcomes? A cross-state analysis." *Educational Evaluation and Policy Analysis* 24, no. 4 (Winter): 305-331.

Chetty, Raj, John N. Friedman, and Jonah Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review* 104, no. 9 (September): 2633-2679.

Cohen, David K., and Richard J. Murnane. 1985. "The merits of merit pay." *Public Interest* 80(Summer): 3-30.

Cohen, David K., and Richard J. Murnane. 1986. "Merit pay and the evaluation problem: Understanding why most merit pay plans fail and a few survive." *Harvard Educational Review* 56, no. 1 (February): 1-17.

Cross, Christopher T. 2014. *Political education: Setting the course for state and federal policy*. Second ed. New York, NY: Teachers College, Columbia University.

Dee, Thomas S., and Brian A. Jacob. 2011. "The impact of No Child Left Behind on student achievement." *Journal of Policy Analysis and Management* 30, no. 3: 418-446.

Dee, Thomas S., and Benjamin J. Keys. 2004. "Does merit pay reward good teachers? Evidence from a randomized experiment." *Journal of Policy Analysis and Management* 23, no. 3 (Summer): 471-488.

Derthick, Martha, and Andrew J. Rotherham. 2013. "Obama's NCLB Waivers: Are they necessary or illegal?" *Education Next* 12, no. 2 (Spring): 56-61.

Duncan, Arne. 2018. *How schools work: An inside account of failure and success from on the the nation's longest-serving Secretaries of Education*. New York: Simon and Schuster.

Figlio, David, and Susanna Loeb. 2011. "School Accountability." In *Handbook of the Economics of Education*, edited by Stephen Machin Eric A. Hanushek and Ludger Woessmann: Elsevier: 383-421.

Handel, Danielle V., and Eric A. Hanushek. 2023. "U.S. School Finance: Resources and Outcomes." In *Handbook of the Economics of Education. Volume 7*, edited by Eric A Hanushek, Stephen Machin, and Ludger Woessmann. Amsterdam: North Holland: 143-226.

Handel, Danielle V., and Eric A. Hanushek. 2024. "Contexts of Convenience: Generalizing from Published Evaluations of School Finance Policies." *Evaluation Review* 48, no. 3 (June): 461-494.

Hanushek, Eric A. 2023. *The Economic Cost of the Pandemic: State by State*. Stanford, CA: Hoover Education Success Initiative, Hoover Institution.

Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *Economic Journal* 113, no. 485 (February): F64-F98.

Hanushek, Eric A., and Alfred A. Lindseth. 2009. *Schoolhouses, courthouses, and statehouses: Solving the funding-achievement puzzle in America's public schools*. Princeton, NJ: Princeton University Press.
Hanushek, Eric A., and Margaret E. Raymond. 2005. "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24, no. 2: 297-327.
Hanushek, Eric A., and Steven G. Rivkin. 2010. "Generalizations about using value-added measures of teacher quality." *American Economic Review* 100, no. 2 (May): 267-271.
Hanushek, Eric A., and Steven G. Rivkin. 2012. "The distribution of teacher quality and implications for policy." *Annual Review of Economics* 4: 131-157.
Hanushek, Eric A., Jens Ruhose, and Ludger Woessmann. 2017a. "Economic gains from educational reform by US States." *Journal of Human Capital* 11, no. 4 (Winter): 447-486.
Hanushek, Eric A., Jens Ruhose, and Ludger Woessmann. 2017b. "Knowledge capital and aggregate income differences: Development accounting for U.S. states." *American Economic Journal: Macroeconomics* 9, no. 4 (October): 184-224.
Holmstrom, Bengt, and Paul Milgrom. 1991. "Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design." *Journal of Law, Economics and Organization* 7: 24-52.
Jackson, C. Kirabo, and Claire L. Mackevicius. 2024. "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States." *American Economic Journal: Applied Economics* 16, no. 1: 412-46.
Jha, Nandan K., Neena Banerjee, and Stephanie Moller. 2020. "Assessing the Role of Teachers' Unions in the Adoption of Accountability Policies in Public Education." *The Urban Review* 52, no. 2 (2020/06/01): 299-330.
Koedel, Cory, Kata Mihaly, and Jonah E. Rockoff. 2015. "Value-added modeling: A review." *Economics of Education Review* 47: 180-195.
Koretz, Daniel. 2009. "Moving Past No Child Left Behind." *Science* 326, no. 5954 (November 6, 2009): 803-804.
Kraft, Matthew A. 2020. "Interpreting Effect Sizes of Education Interventions." *Educational Researcher* 49, no. 4 (May): 241-253.
Kraft, Matthew A., Eric J. Brunner, Shaun M. Dougherty, and David J. Schwegman. 2020. "Teacher accountability reforms and the supply and quality of new teachers." *Journal of Public Economics* 188(August): 104212.
Lewis, Karyn, and Megan Kuhfeld. 2023. *Education's long COVID: 2022–23 achievement data reveal stalled progress toward pandemic recovery*. Center for school and student progress: NWEA (July).
McGuinn, Patrick. 2016. "From No Child Left behind to the Every Student Succeeds Act: Federalism and the Education Legacy of the Obama Administration." *Publius: The Journal of Federalism* 46, no. 3: 392-415.
Oates, Wallace E. 1972. *Fiscal Federalism*. United States: Harcourt Brace Jovanovich, Inc.
Oates, Wallace E. 1999. "An essay on fiscal federalism." *Journal of Economic Literature* 37, no. 3 (September): 1120-1149.
Polikoff, Morgan S., Andrew McEachin, Stephani L. Wrabel, and Matthew Duque. 2014. *Grading the No Child Left Behind Waivers*. Education Outlook. Washington, DC: American Enterprise Institute (February).
Taylor, Eric S. 2023. "Teacher evaluation and training." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann: Elsevier: 61-141.
U.S. Department of Education. 2022. *Digest of Education Statistics 2022*. Washington, DC: National Center for Education Statistics.

Figure 1. NAEP Performance in Math and Reading

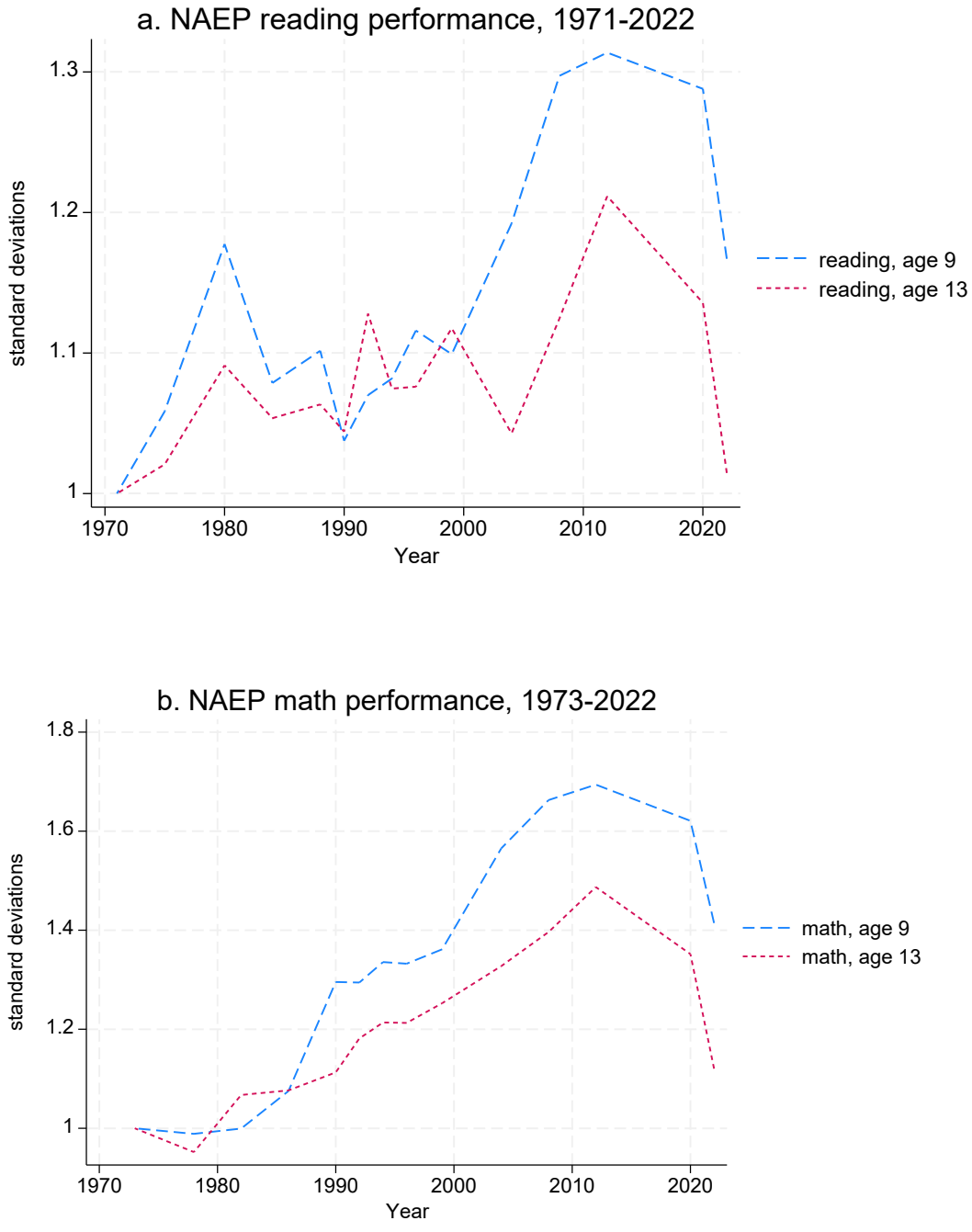

Figure 2. Pattern of Individual Output-based Teacher Policies

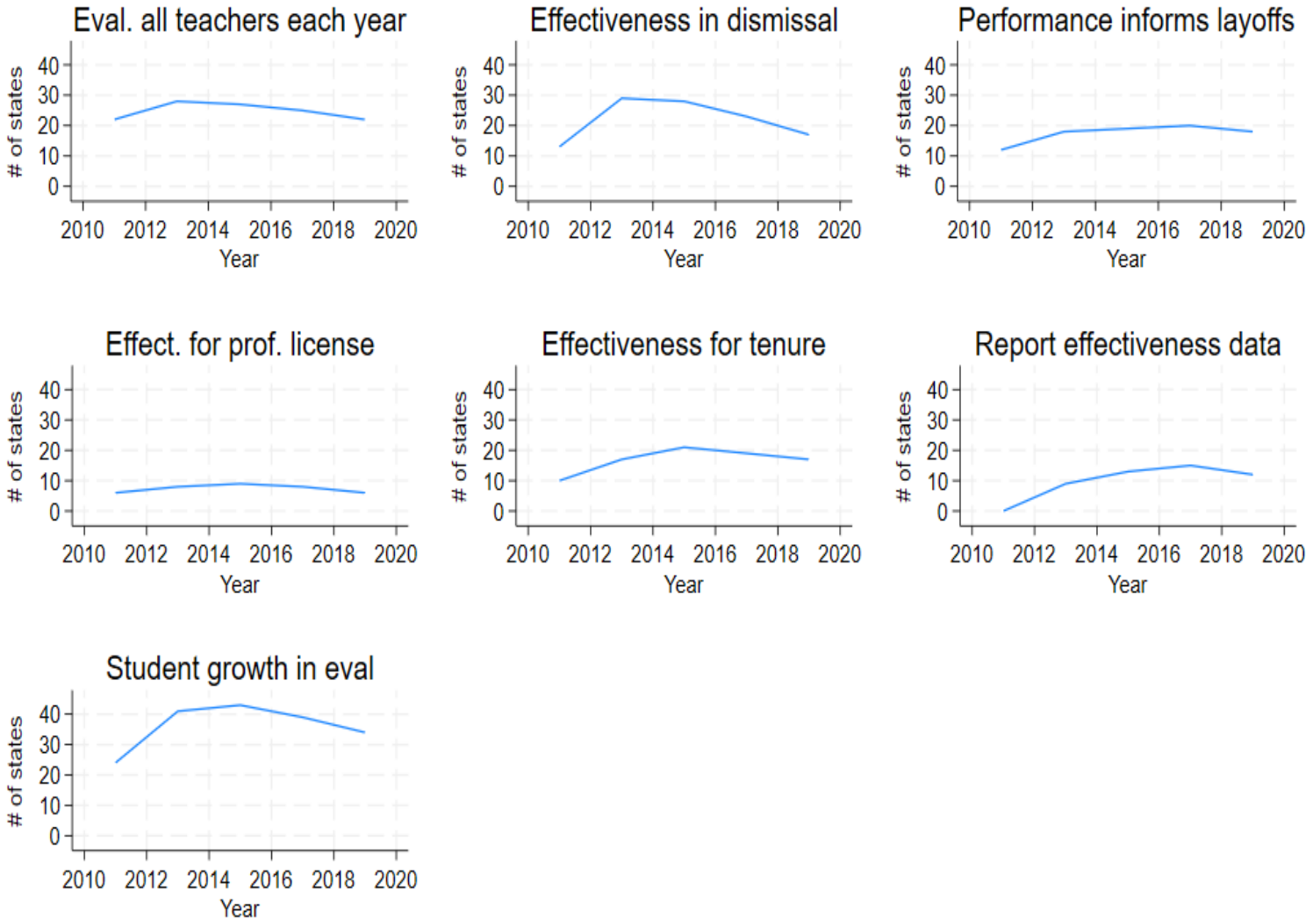

Figure 3. Pattern of Individual Input-based Teacher Policies

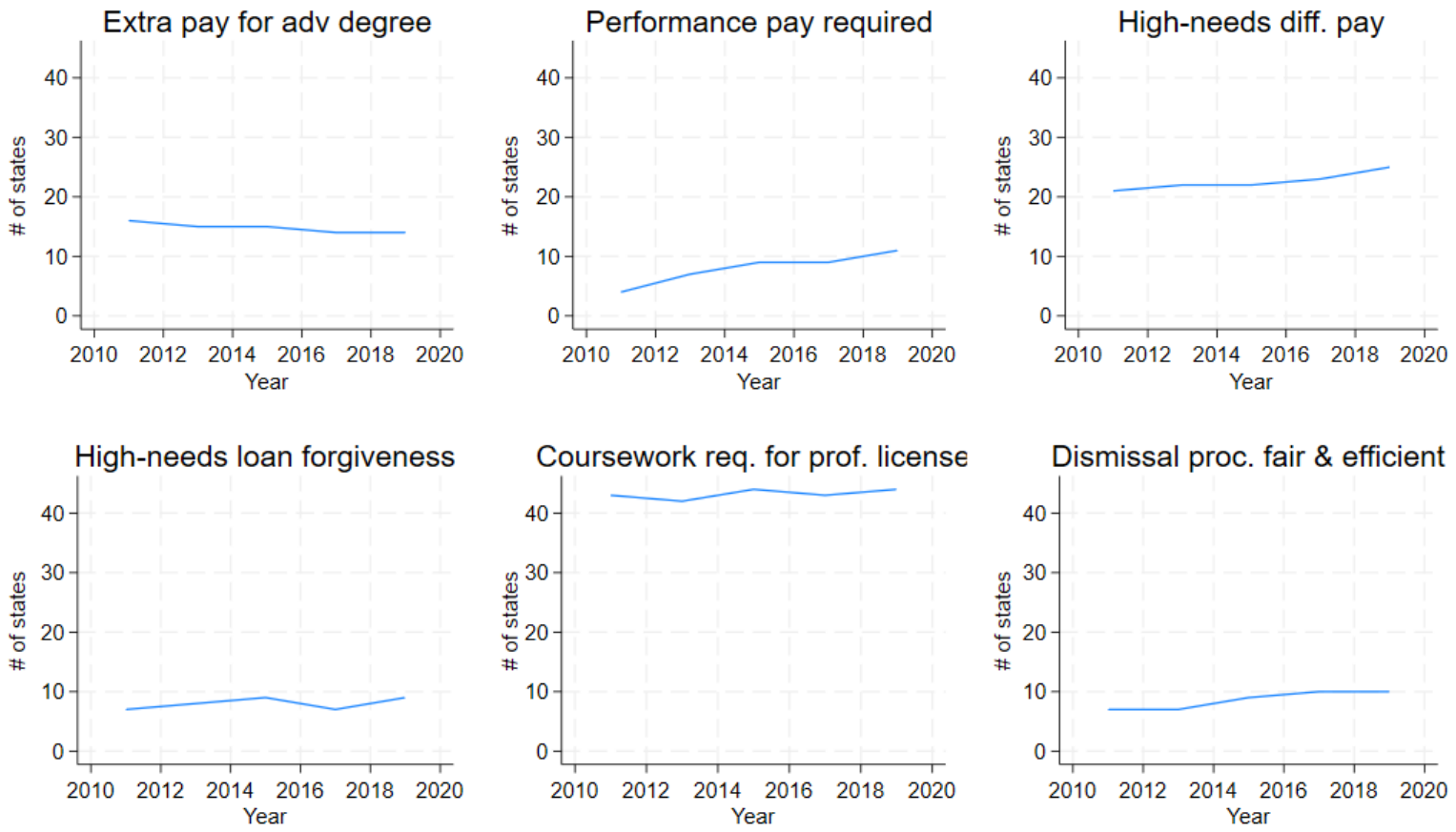

Figure 4. Three-Year Change of Adoption of Output-based Policies: 2015 & 2019

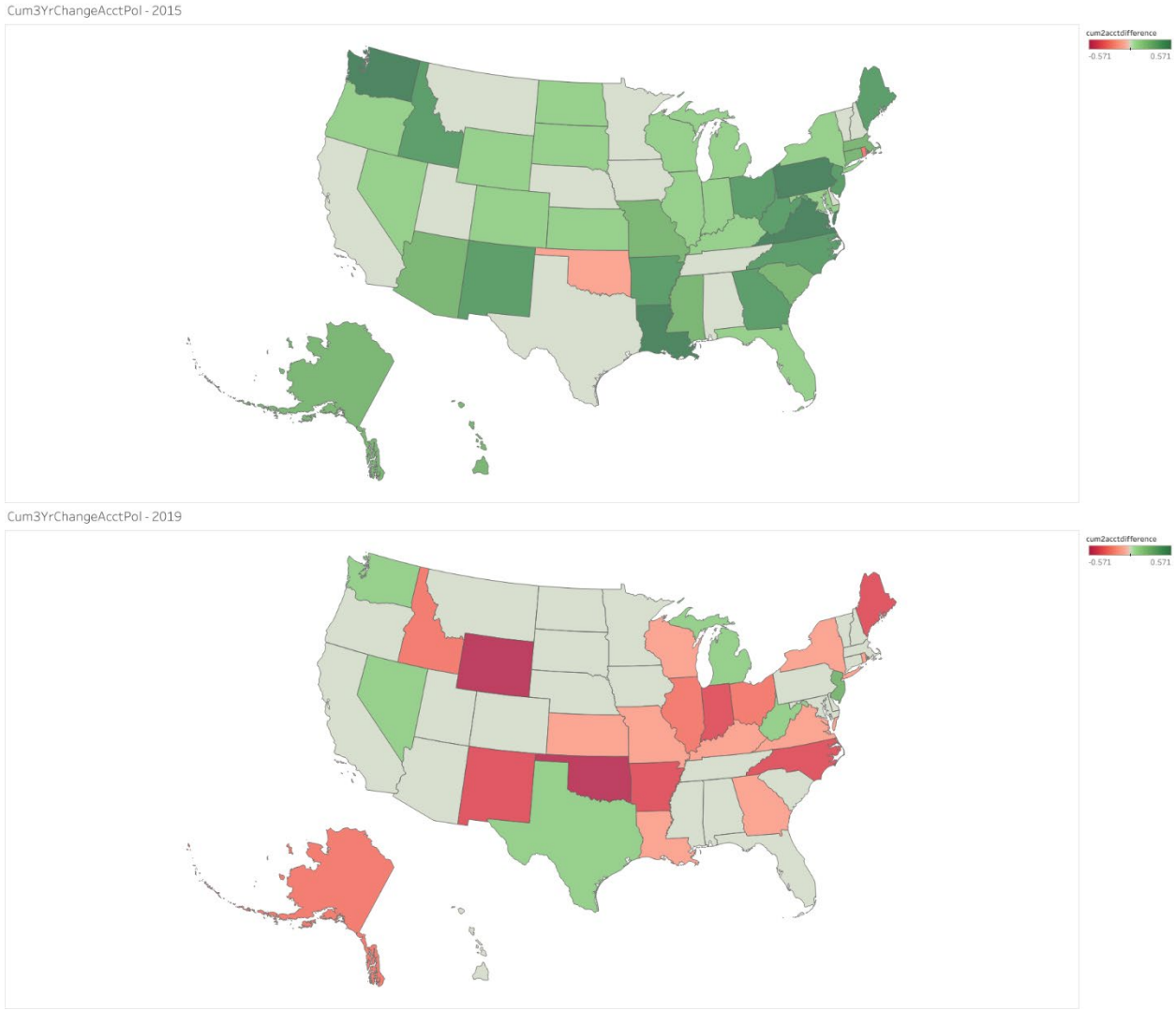

Figure 5. Three-Year Change of Adoption of Input-based Policies: 2015 & 2019

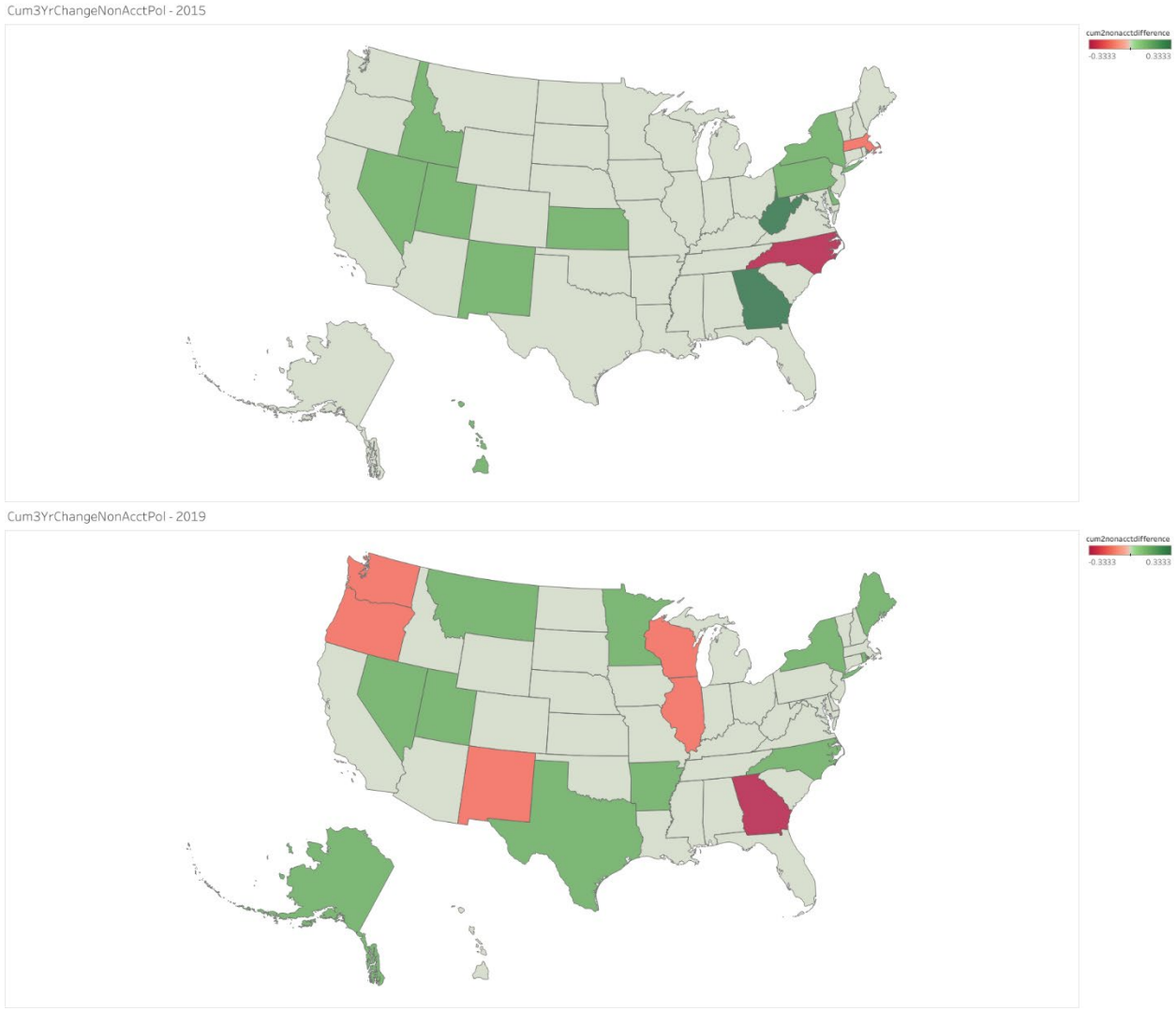

Table 1. Identification and Coding of Output-based Teacher Policies

Policy component	Coding	Definition
Student growth in teacher evaluation	1	State legislation requires that objective measures of student learning are included in their teacher evaluation instruments
	0	Policy is not required, or that measures of student performance are not objective, or measures not learning related (e.g. attendance), or school level data can be used in place of classroom or student level data
Evaluate all teachers each year	1	State legislation explicitly requires evaluation of all teachers each year
	0	A portion of teachers are not evaluated annually; e.g., by experience, tenure status, or previous effectiveness ratings.
Dismissal for instructional ineffectiveness	1	State legislation articulates instructional ineffectiveness is adequate grounds for dismissal with effectiveness directly related to objective measures of student performance and there is a direct link between assessment of effectiveness and teacher evaluation scores.
	0	Otherwise
Performance in layoff decisions	1	State legislation explicitly requires teacher effectiveness (measured by student outcomes) be considered when making layoff decisions, regardless of whether teacher effectiveness is the main or sole criterion or one of many criteria
	0	Otherwise
Performance to qualify for professional license	1	State legislation explicitly requires effectiveness (measured by student outcomes) as a condition for advancement to a professional license, regardless of when during the probationary period that the requirement is met
	0	Otherwise
Performance in tenure process	1	State legislation explicitly requires evidence of teacher effectiveness (measured by student outcomes) be considered in the tenure process, regardless of when during the probationary period the effectiveness requirement is met
	0	Otherwise, including when tenure is not available
Public reporting of teacher effectiveness data	1	State reports publicly the number or percentage of effective or ineffective teachers, aggregated consistently with applicable privacy constraints; observed from web search
	0	No report could be found in public sites or was reported just at the district level

Table 2. Identification and Coding of Input-based Teacher Policies

Policy component	Coding	Definition
Compensation for advanced degrees	1	State legislation as to whether it explicitly requires compensation for advanced degrees
	0	State discourages or prohibits compensation for advanced degrees, requires other measures such as performance to count more than advanced degrees in the compensation calculation, or leaves decision to the discretion of each district
Performance pay	1	While requiring consideration of effectiveness in awarding pay, does not explicitly address effectiveness measurement
	0	Otherwise
High-needs pay	1	States incentivize teaching in high-needs schools with differential pay to teachers who choose to teach in specially designated schools without regard to effectiveness
	0	Otherwise
Loan forgiveness	1	States relieve the loan burden of teachers in specially designated schools
	0	Otherwise
Coursework for professional licenses	1	State legislative requirement for additional coursework before conferring or renewing professional licenses.
	0	Otherwise
Fair and efficient dismissal process	1	States received a 1 if there is a clear distinction between due process based on ineffective classroom performance and other grounds for dismissal, if number of appeals is limited, and if requirements are in place for the timing of the appeals process
	0	Otherwise

Table 3. Prevalence across States of Alternative Teacher Policies, 2011-2019

	Prevalence of policy in 51 states (in % of all states)				
	2011	**2013**	**2015**	**2017**	**2019**
Output-based policies	***0.24***	***0.42***	***0.45***	***0.42***	***0.35***
Measures of student growth required in evaluations	0.46	0.80	0.84	0.76	0.67
Evaluate all teachers each year	0.42	0.55	0.53	0.49	0.43
Dismissal for instructional effectiveness	0.26	0.57	0.55	0.45	0.33
Performance informs layoffs	0.24	0.35	0.37	0.39	0.35
Effectiveness for professional license	0.12	0.16	0.18	0.16	0.12
Effectiveness for tenure	0.20	0.33	0.41	0.37	0.33
Public reporting of effectiveness data	0.00	0.18	0.25	0.29	0.24
Input-based policies	***0.32***	***0.33***	***0.35***	***0.35***	***0.37***
Extra pay for advanced degrees	0.30	0.29	0.29	0.27	0.27
Performance pay	0.08	0.14	0.18	0.18	0.22
High needs differential pay	0.40	0.43	0.43	0.45	0.49
High needs loan forgiveness	0.14	0.16	0.18	0.14	0.18
Course requirements for professional license	0.86	0.82	0.86	0.84	0.86
Dismissal process is fair and efficient	0.14	0.14	0.18	0.20	0.20

Note: Shaded cells represent peak inclusion of each provision across states.

Source: NCTQ data

Table 4. Aggregated Teacher Policies

	Obs	2015 Average prevalence	2015 Std. dev.	2019 Average prevalence	2019 Std. dev.
Output-based policies	***51***	***0.43***	***0.27***	***0.39***	***0.29***
Measures of student growth required in evaluations	51	0.82	0.37	0.72	0.43
Evaluate all teachers each year	51	0.54	0.47	0.46	0.49
Dismissal for instructional effectiveness	51	0.56	0.50	0.39	0.46
Performance informs layoffs	51	0.36	0.48	0.37	0.47
Effectiveness for professional license	51	0.17	0.36	0.14	0.33
Effectiveness for tenure	51	0.37	0.47	0.35	0.47
Public reporting of effectiveness data	51	0.22	0.39	0.26	0.42
Input-based policies	***51***	***0.34***	***0.17***	***0.36***	***0.18***
Extra pay for advanced degrees	51	0.29	0.46	0.27	0.45
Performance pay	51	0.16	0.35	0.20	0.39
High needs differential pay	51	0.43	0.49	0.47	0.48
High needs loan forgiveness	51	0.17	0.37	0.16	0.35
Course requirements for professional license	51	0.84	0.35	0.85	0.34
Dismissal process is fair and efficient	51	0.16	0.35	0.20	0.40

Table 5. Baseline Estimates of State Teacher Policies

8th grade scores - all students

Independent variables	(1) Pooled (Reading+Math)	(2) Reading	(3) Math
ESSA (=1)	-0.212***	-0.335***	-0.079***
	(0.03)	0.03	0.03
4th grade scores	0.538***	0.606***	0.629***
	(0.05)	(0.07)	(0.06)
Subject (Math=1)	0.117***		
	(0.02)		
Parental educational attainment	0.326***	0.281***	0.268***
	(0.08)	(0.10)	(0.08)
Per-pupil state expenditure	-0.210*	-0.259*	-0.129
	(0.11)	(0.13)	(0.11)
Prevalence of outcome-based policies	***0.290****	***0.313****	***0.211****
	(0.10)	(0.13)	(0.10)
Prevalence of input-based policies	***-0.456*****	***-0.437****	***-0.432****
	(0.17)	(0.21)	(0.17)
Observations	204	102	102

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table 6. Estimates of State Teacher Policies with Race/Ethnicity Disaggregation

8th grade scores - students by race/ethnicity

Independent variables	(1) Pooled (Reading+Math)	(2) Reading	(3) Math
ESSA (=1)	-0.184***	-0.328***	-0.035
	(0.02)	(0.04)	(0.03)
Race/Ethnicity - Asian	0.623***	0.518***	0.684***
	(0.04)	(0.06)	(0.06)
Race/Ethnicity - Black	-0.884***	-1.029***	-0.669***
	(0.06)	(0.08)	(0.08)
Race/Ethnicity - Hispanic	-0.196***	-0.278***	-0.086
	(0.07)	(0.10)	(0.08)
Corresponding 4th grade scores	0.577***	0.562***	0.632***
	(0.03)	(0.04)	(0.04)
Subject (Math=1)	0.150***		
	(0.02)		
Parental educational attainment	0.138***	0.127***	0.139***
	(0.02)	(0.03)	(0.03)
Per-pupil state expenditure	-0.074	-0.101	-0.026
	(0.08)	(0.09)	(0.09)
Prevalence of outcome-based policies	***0.249******	***0.245****	***0.203****
	(0.08)	(0.01)	(0.10)
Prevalence of input-based policies	***-0.302****	***-0.270****	***-0.241***
	(0.14)	(0.16)	(0.16)
Observations	660	331	329

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table 7. Estimates of State Teacher Policies Stratified by Race/Ethnicity

Independent variables	(1) Asian	(2) Black	(3) Hispanic	(4) White
ESSA (=1)	-0.0458	-0.280***	-0.199***	-0.190***
	-0.0742	-0.0449	-0.0361	-0.0265
4th grade scores	0.457***	0.385***	0.322***	0.375***
	-0.0852	-0.0554	-0.0561	-0.0608
Subject (Math=1)	0.293***	0.209***	0.104***	0.0261
	-0.0717	-0.0476	-0.0352	-0.0256
Parental educational attainment	0.296***	0.0049	0.0925***	0.402***
	-0.0826	-0.058	-0.0332	-0.0572
Per-pupil state expenditure	-0.196	0.101	-0.216**	0.0175
	-0.26	-0.114	-0.106	-0.0891
Prevalence of outcome-based policies	***0.920******	***0.364******	***-0.131***	***0.284******
	-0.281	-0.121	-0.112	-0.0911
Prevalence of input-based policies	***-1.089****	***-0.294***	***0.372****	***-0.271****
	-0.453	-0.2	-0.196	-0.153
Observations	101	164	191	191

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table 8. Removing Performance Pay from Both Input-based and Outcome-based Policies

Independent variables	(1) Pooled (Reading+Math)	(2) Reading	(3) Math
ESSA (=1)	-0.213***	-0.337***	-0.0776***
	-0.0258	-0.0342	-0.0287
4th grade scores	0.535***	0.598***	0.630***
	-0.0502	-0.075	-0.0579
Subject (Math=1)	0.117***		
	-0.0224		
Parental educational attainment	0.326***	0.281***	0.269***
	-0.0818	-0.106	-0.0831
Per-pupil state expenditure	-0.206*	-0.244*	-0.136
	-0.111	-0.138	-0.111
Prevalence of outcome-based policies without Performance Pay	***0.288******	***0.306****	***0.219****
	-0.103	-0.131	-0.104
Prevalence of input-based policies without Performance Pay	***-0.385******	***-0.381****	***-0.354****
	-0.139	-0.176	-0.141
Performance Pay	***-0.0589***	***-0.0209***	***-0.0983***
	-0.0727	-0.0923	-0.0734
Observations	204	102	102

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01

Table 9. Impact of the Move from NCLB to ESSA through Teacher Policies (s.d.)

	Aggregate	Reading	Math
Change in output-based	-0.015	-0.016	-0.011
Change in input-based	-0.009	-0.009	-0.009

Note: Estimates based on teacher policy changes from Table 4 and estimated impacts from Table 5

Appendix A: References to Specific Teacher Policies in Federal Legislation

Sources:
NCLB: **The Elementary and Secondary Education Act (The No Child Left Behind Act of 2001)**
ESEA Flexibility Policy Document
Race-to-the-top grant competition
ESSA: **The Every Student Succeeds Act**

Citations:
Every Student Succeeds Act, 20 U.S.C. §6301 (2015). https://congress.gov/114/plaws/publ95/PLAW-114publ95.pdf

No Child Left Behind Act, 20 U.S.C. §6301 (2001). https://www.govinfo.gov/content/pkg/PLAW-107publ110/pdf/PLAW-107publ110.pdf

U.S. Department of Education. (2012). ESEA Flexibility Policy Document. https://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html

U.S. Department of Education (2009). Race-to-the-top Program Executive Summary. https://files.eric.ed.gov/fulltext/ED557422.pdf

Output-based Policies

Use of Objective Measures of Student Growth in Teacher Evaluations

NCLB:
Title I, Part A – Improving Basic Programs Operated by Local Educational Agencies, Sec. 1119, 115 STAT. 1505
(2) STATE PLAN.—As part of the plan described in section 1111, each State educational agency receiving assistance under this part shall develop a plan to ensure that all teachers teaching in core academic subjects within the State are highly qualified not later than the end of the 2005–2006 school year. Such plan shall establish annual measurable objectives for each local educational agency and school that, at a minimum—
(A) shall include an annual increase in the percentage of highly qualified teachers at each local educational agency and school, to ensure that all teachers teaching in core academic subjects in each public elementary school and secondary school are highly qualified not later than the end of the 2005–2006 school year;
(B) shall include an annual increase in the percentage of teachers who are receiving high-quality professional development to enable such teachers to become highly qualified and successful classroom teachers; and
(C) may include such other measures as the State educational agency determines to be appropriate to increase teacher qualifications.

Title I, Part A – Improving Basic Programs Operated by Local Educational Agencies, Sec. 1119, 115 STAT. 1959
(23) HIGHLY QUALIFIED.—The term 'highly qualified'—

(A) when used with respect to any public elementary school or secondary school teacher teaching in a State, means that—
(i) the teacher has obtained full State certification as a teacher (including certification obtained through alternative routes to certification) or passed the State teacher licensing examination, and holds a license to teach in such State, except that when used with respect to any teacher teaching in a public charter school, the term means that the teacher meets the requirements set forth in the State's public charter school law; and
(ii) the teacher has not had certification or licensure requirements waived on an emergency, temporary, or provisional basis;
(B) when used with respect to—
(i) an elementary school teacher who is new to the profession, means that the teacher—
(I) holds at least a bachelor's degree; and
(II) has demonstrated, by passing a rigorous State test, subject knowledge and teaching skills in reading, writing, mathematics, and other areas of the basic elementary school curriculum (which may consist of passing a State-required certification or licensing test or tests in reading, writing, mathematics, and other areas of the basic elementary school curriculum); or
(ii) a middle or secondary school teacher who is new to the profession, means that the teacher holds at least a bachelor's degree and has demonstrated a high level of competency in each of the academic subjects in which the teacher teaches by—
(I) passing a rigorous State academic subject test in each of the academic subjects in which the teacher teaches (which may consist of a passing level of performance on a State-required certification or licensing test or tests in each of the academic subjects in which the teacher teaches); or
(II) successful completion, in each of the academic subjects in which the teacher teaches, of an academic major, a graduate degree, coursework equivalent to an undergraduate academic major, or advanced certification or credentialing; and
(C) when used with respect to an elementary, middle, or secondary school teacher who is not new to the profession, means that the teacher holds at least a bachelor's degree and—
(i) has met the applicable standard in clause (i) or (ii) of subparagraph (B), which includes an option for a test; or
(ii) demonstrates competence in all the academic subjects in which the teacher teaches based on a high objective uniform State standard of evaluation that—
(I) is set by the State for both grade appropriate academic subject matter knowledge and teaching skills; (II) is aligned with challenging State academic content and student academic achievement standards and developed in consultation with core content specialists, teachers, principals, and school administrators; (III) provides objective, coherent information about the teacher's attainment of core content knowledge in the academic subjects in which a teacher teaches;
(IV) is applied uniformly to all teachers in the same academic subject and the same grade level throughout the State;
(V) takes into consideration, but not be based primarily on, the time the teacher has been teaching in the academic subject;
(VI) is made available to the public upon request; and
(VII) may involve multiple, objective measures of teacher competency.

ESEA Flexibility Document:
"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through

multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:
Reform Plan Criteria
(D)(2) Improving teacher and principal effectiveness based on performance (58 points) The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—
(ii) Design and implement rigorous, transparent, and fair evaluation systems for teachers and principals that (a) differentiate effectiveness using multiple rating categories that take into account data on student growth (as defined in this notice) as a significant factor, and (b) are designed and developed with teacher and principal involvement; (15 points)

ESSA:
Title II, Part A – Supporting Effective Instruction, sec. 2101 S.1177–119
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:

(ii) Developing, improving, or providing assistance to local educational agencies to support the design and implementation of teacher, principal, or other school leader evaluation and support systems that are based in part on evidence of student academic achievement, which may include student growth, and shall include multiple measures of educator performance and provide clear, timely, and useful feedback to teachers, principals, or other school leaders, such as by—

(I) developing and disseminating high-quality evaluation tools, such as classroom observation rubrics, and methods, including training and auditing, for ensuring inter-rater reliability of evaluation results;

(II) developing and providing training to principals, other school leaders, coaches, mentors, and evaluators on how to accurately differentiate performance, provide useful and timely feedback, and use evaluation results to inform decisionmaking about professional development, improvement strategies, and personnel decisions; and

(III) developing a system for auditing the quality of evaluation and support systems.

Title I, Part A – Improving Basic Programs Operated by Local Educational Agencies, Sec. 1112, S.1177–52,53
(b) PLAN PROVISIONS.—To ensure that all children receive a high-quality education, and to close the achievement gap between children meeting the challenging State academic standards and those children who are not meeting such standards, each local educational agency plan shall describe—

(2) how the local educational agency will identify and address, as required under State plans as described in section 1111(g)(1)(B), any disparities that result in low-income students and minority students being taught at higher rates than other students by ineffective, inexperienced, or out-of-field teachers;

Annual evaluations for all teachers

NCLB:
Not addressed.

ESEA Flexibility Document:
"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that:

(1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:
Reform Plan Criteria
(D)(2) Improving teacher and principal effectiveness based on performance (58 points) The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—
(iii) Conduct annual evaluations of teachers and principals that include timely and constructive feedback; as part of such evaluations, provide teachers and principals with data on student growth for their students, classes, and schools; and (10 points)

ESSA:
Not addressed.

Instructional ineffectiveness as grounds for dismissal

NCLB:
Not addressed.

ESEA Flexibility Document:
"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:
Reform Plan Criteria
(D)(2) Improving teacher and principal effectiveness based on performance (58 points) The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—
(iv) Use these evaluations, at a minimum, to inform decisions regarding— (28 points)
(d) Removing ineffective tenured and untenured teachers and principals after they have had ample opportunities to improve, and ensuring that such decisions are made using rigorous standards and streamlined, transparent, and fair procedures.

ESSA:

Title II, Part B – National Activities, Subpart 1 – Teacher and School Leader Incentive Program, Sec. 2211, S11-77-130

(a) PURPOSES.—The purposes of this subpart are—
(1) to assist States, local educational agencies, and nonprofit organizations to develop, implement, improve, or expand comprehensive performance-based compensation systems or human capital management systems for teachers, principals, or other school leaders (especially for teachers, principals, or other school leaders in high-need schools) who raise student academic achievement and close the achievement gap between high- and low-performing students; and
(2) to study and review performance-based compensation systems or human capital management systems for teachers, principals, or other school leaders to evaluate the effectiveness, fairness, quality, consistency, and reliability of the systems.
(3) HUMAN CAPITAL MANAGEMENT SYSTEM.—
The term 'human capital management system' means a system—
(A) by which a local educational agency makes and implements human capital decisions, such as decisions on preparation, recruitment, hiring, placement, retention, dismissal, compensation, professional development, tenure, and promotion; and
(B) that includes a performance-based compensation system.

Title II, Part A – Supporting Effective Instruction, sec. 2101 S.1177–119

(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(ii) Developing, improving, or providing assistance to local educational agencies to support the design and implementation of teacher, principal, or other school leader evaluation and support systems that are based in part on evidence of student academic achievement, which may include student growth, and shall include multiple measures of educator performance and provide clear, timely, and useful feedback to teachers, principals, or other school leaders, such as by—
(II) developing and providing training to principals, other school leaders, coaches, mentors, and evaluators on how to accurately differentiate performance, provide useful and timely feedback, and use evaluation results to inform decisionmaking about professional development, improvement strategies, and personnel decisions;

Effectiveness considered in layoff decisions

NCLB:

Not addressed.

ESEA Flexibility Document:

"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:

Not addressed.

ESSA:
Title II, Part A – Supporting Effective Instruction, sec. 2101 S.1177–119
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(ii) Developing, improving, or providing assistance to local educational agencies to support the design and implementation of teacher, principal, or other school leader evaluation and support systems that are based in part on evidence of student academic achievement, which may include student growth, and shall include multiple measures of educator performance and provide clear, timely, and useful feedback to teachers, principals, or other school leaders, such as by—
(II) developing and providing training to principals, other school leaders, coaches, mentors, and evaluators on how to accurately differentiate performance, provide useful and timely feedback, and use evaluation results to inform decisionmaking about professional development, improvement strategies, and personnel decisions;

Evidence of effectiveness to qualify for a professional license

NCLB:
Not addressed.

ESEA Flexibility Document:
"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:
Reform Plan Criteria
(D)(2) Improving teacher and principal effectiveness based on performance (58 points) The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—
(iv) Use these evaluations, at a minimum, to inform decisions regarding— (28 points)
(c) Whether to grant tenure and/or full certification (where applicable) to teachers and principals using rigorous standards and streamlined, transparent, and fair procedures;

ESSA:
Title II, Part A – Supporting Effective Instruction, sec. 2101 S.1177–119
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(i) Reforming teacher, principal, or other school leader certification, recertification, licensing, or tenure systems or preparation program standards and approval processes to ensure that—
(I) teachers have the necessary subject-matter knowledge and teaching skills, as demonstrated through measures determined by the State, which may include teacher performance assessments, in the academic subjects that the teachers teach to help students meet challenging State academic standards;

(II) principals or other school leaders have the instructional leadership skills to help teachers teach and to help students meet such challenging State academic standards; and
(III) teacher certification or licensing requirements are aligned with such challenging State academic standards.

Effectiveness is considered in the tenure process

NCLB:
No mention of tenure process.

ESEA Flexibility Document:
"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:
Reform Plan Criteria
(D)(2) Improving teacher and principal effectiveness based on performance (58 points) The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—
(iv) Use these evaluations, at a minimum, to inform decisions regarding— (28 points)
(c) Whether to grant tenure and/or full certification (where applicable) to teachers and principals using rigorous standards and streamlined, transparent, and fair procedures;

ESSA:
Title II, Part A – Supporting Effective Instruction, sec. 2101 S.1177–119
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(i) Reforming teacher, principal, or other school leader certification, recertification, licensing, or tenure systems or preparation program standards and approval processes to ensure that—
(I) teachers have the necessary subject-matter knowledge and teaching skills, as demonstrated through measures determined by the State, which may include teacher performance assessments, in the academic subjects that the teachers teach to help students meet challenging State academic standards;
(II) principals or other school leaders have the instructional leadership skills to help teachers teach and to help students meet such challenging State academic standards; and
(III) teacher certification or licensing requirements are aligned with such challenging State academic standards.

ESSA:
Title II, Part A – Supporting Effective Instruction, sec. 2101 S.1177–119
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(ii) Developing, improving, or providing assistance to local educational agencies to support the design and implementation of teacher, principal, or other school leader evaluation and support systems that are based

in part on evidence of student academic achievement, which may include student growth, and shall include multiple measures of educator performance and provide clear, timely, and useful feedback to teachers, principals, or other school leaders, such as by—
(II) developing and providing training to principals, other school leaders, coaches, mentors, and evaluators on how to accurately differentiate performance, provide useful and timely feedback, and use evaluation results to inform decisionmaking about professional development, improvement strategies, and personnel decisions;

Title II, Part B – National Activities, Subpart 1 – Teacher and School Leader Incentive Program, Sec. 2211, S11-77-130
(a) PURPOSES.—The purposes of this subpart are—
(1) to assist States, local educational agencies, and nonprofit organizations to develop, implement, improve, or expand comprehensive performance-based compensation systems or human capital management systems for teachers, principals, or other school leaders (especially for teachers, principals, or other school leaders in high-need schools) who raise student academic achievement and close the achievement gap between high- and low-performing students; and
(2) to study and review performance-based compensation systems or human capital management systems for teachers, principals, or other school leaders to evaluate the effectiveness, fairness, quality, consistency, and reliability of the systems.
(3) HUMAN CAPITAL MANAGEMENT SYSTEM.—
The term 'human capital management system' means a system—
(A) by which a local educational agency makes and implements human capital decisions, such as decisions on preparation, recruitment, hiring, placement, retention, dismissal, compensation, professional development, tenure, and promotion; and
(B) that includes a performance-based compensation system.

Reporting of teacher effectiveness data

NCLB:
Title I, Part A – Improving Basic Programs Operated by Local Educational Agencies, Sec. 1119, 115 STAT. 1506
(2) ANNUAL REPORTS BY THE SECRETARY.—Each year, beginning with the 2002–2003 school year, the Secretary shall publicly report the annual progress of State educational agencies, local educational agencies, and schools, in meeting the measurable objectives described in subsection (a)(2).

ESEA Flexibility Document:
Not addressed.

Race-to-the-top:
(D)(4) Improving the effectiveness of teacher and principal preparation programs (14 points)
The extent to which the State has a high-quality plan and ambitious yet achievable annual targets to—
(i) Link student achievement and student growth (both as defined in this notice) data to the students' teachers and principals, to link this information to the in-State programs where those teachers and principals were prepared for credentialing, and to publicly report the data for each credentialing program in the State;

ESSA:
Title II, Part A- Supporting Effective Instruction, sec. 2104, S. 1177-129

(a) STATE REPORT.—Each State educational agency receiving funds under this part shall annually submit to the Secretary a report that provides—
(3) for a State that implements a teacher, principal, or other school leader evaluation and support system, consistent with section 2101(c)(4)(B)(ii), using funds under this part, the evaluation results of teachers, principals, or other school leaders, except that such information shall not provide personally identifiable information on individual teachers, principals, or other school leaders;
(c) AVAILABILITY.—The reports and information provided under subsections (a) and (b) shall be made readily available to the public.

Input-based Policies:

Compensation for Advanced Degrees

NCLB:
Not addressed.

ESEA Flexibility Document:
Not addressed.

Race-to-the-top:
Not addressed.

ESSA:
Not addressed.

Performance Pay

NCLB:
Title II, Part A – Teacher and Principal Training and Recruitment Fund, sec. 2113, 115 STAT 1625
STATE ACTIVITIES.—The State educational agency for a State that receives a grant under section 2111 shall use the funds described in subsection (a)(3) to carry out one or more of the following activities, which may be carried out through a grant or contract with a for-profit or nonprofit entity:
(12) Developing, or assisting local educational agencies in developing, merit-based performance systems, and strategies that provide differential and bonus pay for teachers in high-need academic subjects such as reading, mathematics, and science and teachers in high-poverty schools and districts.

ESEA Flexibility Document:
"To receive this flexibility, an SEA and each LEA must commit to develop, adopt, pilot, and implement, with the involvement of teachers and principals, teacher and principal evaluation and support systems that: (1) will be used for continual improvement of instruction; (2) meaningfully differentiate performance using at least three performance levels; (3) use multiple valid measures in determining performance levels, including as a significant factor data on student growth for all students (including English Learners and students with disabilities), and other measures of professional practice (which may be gathered through multiple formats and sources, such as observations based on rigorous teacher performance standards, teacher portfolios, and student and parent surveys); (4) evaluate teachers and principals on a regular basis; (5) provide clear, timely, and useful feedback, including feedback that identifies needs and guides professional development; and (6) will be used to inform personnel decisions."

Race-to-the-top Criteria:
Reform Plan Criteria
(D)(2) Improving teacher and principal effectiveness based on performance (58 points) The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to ensure that participating LEAs (as defined in this notice)—
(iv) Use these evaluations, at a minimum, to inform decisions regarding— (28 points)
(b) Compensating, promoting, and retaining teachers and principals, including by providing opportunities for highly effective teachers and principals (both as defined in this notice) to obtain additional compensation and be given additional responsibilities;

ESSA:
Title II, Part A- Supporting Effective Instruction, sec. 2102, S1177-120
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(vii) Developing, or assisting local educational agencies in developing—
(II) strategies that provide differential pay, or other incentives, to recruit and retain teachers in high-need academic subjects and teachers, principals, or other school leaders, in low-income schools and school districts, which may include performance-based pay systems;

Title II, Part B – National Activities, Subpart 1 – Teacher and School Leader Incentive Program, Sec. 2211, S11-77-130
(a) PURPOSES.—The purposes of this subpart are—
(1) to assist States, local educational agencies, and nonprofit organizations to develop, implement, improve, or expand comprehensive performance-based compensation systems or human capital management systems for teachers, principals, or other school leaders (especially for teachers, principals, or other school leaders in high-need schools) who raise student academic achievement and close the achievement gap between high- and low-performing students; and
(2) to study and review performance-based compensation systems or human capital management systems for teachers, principals, or other school leaders to evaluate the effectiveness, fairness, quality, consistency, and reliability of the systems.
(3) HUMAN CAPITAL MANAGEMENT SYSTEM.—
The term 'human capital management system' means a system—
(A) by which a local educational agency makes and implements human capital decisions, such as decisions on preparation, recruitment, hiring, placement, retention, dismissal, compensation, professional development, tenure, and promotion; and
(B) that includes a performance-based compensation system.
(4) PERFORMANCE-BASED COMPENSATION SYSTEM.—The term 'performance-based compensation system' means a system of compensation for teachers, principals, or other school leaders—
(A) that differentiates levels of compensation based in part on measurable increases in student academic achievement; and
(B) which may include—
(i) differentiated levels of compensation, which may include bonus pay, on the basis of the employment responsibilities and success of effective teachers, principals, or other school leaders in hard-to-staff schools or high-need subject areas; and
(ii) recognition of the skills and knowledge of teachers, principals, or other school leaders as demonstrated through—
successful fulfillment of additional responsibilities or job functions, such as teacher leadership roles; and evidence of professional achievement and mastery of content knowledge and superior teaching and leadership skills.

High-Needs Differential Pay

NCLB:
Title II, Part A – Teacher and Principal Training and Recruitment Fund, sec. 2113, 115 STAT 1625
STATE ACTIVITIES.—The State educational agency for a State that receives a grant under section 2111 shall use the funds described in subsection (a)(3) to carry out one or more of the following activities, which may be carried out through a grant or contract with a for-profit or nonprofit entity:
(12) Developing, or assisting local educational agencies in developing, merit-based performance systems, and strategies that provide differential and bonus pay for teachers in high-need academic subjects such as reading, mathematics, and science and teachers in high-poverty schools and districts.

ESEA Flexibility Document:
Not addressed.

Race-to-the-top:
(D)(3) Ensuring equitable distribution of effective teachers and principals (25 points)
The extent to which the State, in collaboration with its participating LEAs (as defined in this notice), has a high-quality plan and ambitious yet achievable annual targets to—
(i) Ensure the equitable distribution of teachers and principals by developing a plan, informed by reviews of prior actions and data, to ensure that students in high-poverty and/or high-minority schools (both as defined in this notice) have equitable access to highly effective teachers and principals (both as defined in this notice) and are not served by ineffective teachers and principals at higher rates than other students; and (15 points)
(ii) Increase the number and percentage of effective teachers (as defined in this notice) teaching hard-to-staff subjects and specialty areas including mathematics, science, and special education; teaching in language instruction educational programs (as defined under Title III of the ESEA); and teaching in other areas as identified by the State or LEA. (10 points)
Plans for (i) and (ii) may include, but are not limited to, the implementation of incentives and strategies in such areas as recruitment, compensation, teaching and learning environments, professional development, and human resources practices and processes.

ESSA:
Title II, Part A- Supporting Effective Instruction, sec. 2102, S1177-120
(B) TYPES OF STATE ACTIVITIES.—The activities described in this subparagraph are the following:
(vii) Developing, or assisting local educational agencies in developing—
(II) strategies that provide differential pay, or other incentives, to recruit and retain teachers in high-need academic subjects and teachers, principals, or other school leaders, in low-income schools and school districts, which may include performance-based pay systems;

Incentivize teaching in high-need schools through loan forgiveness

NCLB:
Not addressed.

ESEA Flexibility Document:
Not addressed.

Race-to-the-top:

Not addressed.

ESSA:
Not addressed.

Coursework for professional license

NCLB:

Not addressed.

ESEA Flexibility Document:
Not addressed.

Race-to-the-top:
Not addressed.

ESSA:
Not addressed.

Fair and efficient dismissal processes

NCLB:
Not addressed.

ESEA Flexibility Document:
Not addressed.

Race-to-the-top:
Not addressed.

ESSA:
Not addressed.

Appendix Table B1: Estimates with disaggregated teacher policies, all students

Independent Variables	(1) Pooled, single policies	(2) Pooled, grouped input based	(3) Reading, single policies	(4) Reading, grouped input based	(5) Math, single policies	(6) Math, grouped input based
Year 2019	-0.222***	-0.222***	-0.354***	-0.351***	-0.0749**	-0.0777**
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Corresponding 4th grade scores	0.528***	0.531***	0.532***	0.578***	0.639***	0.631***
	(0.05)	(0.05)	(0.09)	(0.08)	(0.06)	(0.06)
Subject (Math=1)	0.116***	0.117***				
	(0.02)	(0.02)				
Evaluate all teachers each year	-0.0501	-0.0534	-0.0289	-0.0602	-0.0479	-0.053
	(0.06)	(0.05)	(0.08)	(0.07)	(0.06)	(0.06)
Dismissal for instruct. ineffectiveness	0.0928	0.0864	0.0506	0.0607	0.107	0.102
	(0.08)	(0.08)	(0.11)	(0.10)	(0.08)	(0.08)
Performance informs layoffs	0.0102	0.0122	0.00692	0.00916	0.00873	0.00124
	(0.07)	(0.06)	(0.09)	(0.08)	(0.07)	(0.06)
Effectiveness for professional license	-0.0413	-0.0302	-0.043	-0.0399	-0.0506	-0.0202
	(0.08)	(0.07)	(0.10)	(0.09)	(0.08)	(0.07)
Effectiveness for tenure	0.104*	0.103*	0.0986	0.0978	0.114*	0.112*
	(0.06)	(0.06)	(0.09)	(0.08)	(0.06)	(0.06)
Public reporting of effectiveness data	0.103	0.112*	0.204**	0.176**	0.00268	0.0132
	(0.07)	(0.07)	(0.09)	(0.09)	(0.07)	(0.07)
Measures of student growth req. in eval.	-0.0322	0.0335	-0.00635	0.0798	-0.0427	0.00654
	(0.07)	(0.07)	(0.09)	(0.10)	(0.07)	(0.08)
Extra pay for advanced degrees	-0.124**		-0.1		-0.122**	
	(0.06)		(0.09)		(0.06)	
Performance pay	-0.0873		-0.0263		-0.126	
	(0.08)		(0.11)		(0.08)	
High needs differential pay	-0.0747		-0.0987		-0.0777	
	(0.06)		(0.08)		(0.06)	
High needs loan forgiveness	-0.013		-0.0469		0.0236	
	(0.07)		(0.09)		(0.07)	
Course requirements for prof. license	-0.1		-0.0687		-0.131	
	(0.08)		(0.12)		(0.08)	
Parental educational attainment	0.337***	0.337***	0.301**	0.295***	0.275***	0.277***
	(0.09)	(0.08)	(0.13)	(0.11)	(0.09)	(0.08)
Per-pupil state expenditure	-0.277**	-0.261**	-0.266	-0.301**	-0.217*	-0.193
	(0.13)	(0.12)	(0.18)	(0.15)	(0.12)	(0.12)
Prevalence of input-based policies		-0.480***		-0.487**		-0.469***
		(0.18)		(0.23)		(0.18)
Observations	204	204	102	102	102	102

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix Table B2: Estimates with disaggregated teacher policies and race/ethnicity of students

Independent Variables	(1) Pooled, single policies	(2) Pooled, grouped input based	(3) Reading, single policies	(4) Reading, grouped input based	(5) Math, single policies	(6) Math, grouped input based
Year 2019	-0.187***	-0.188***	-0.344***	-0.341***	-0.0254	-0.032
	(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)
Race/Ethnicity - Asian	0.625***	0.623***	0.521***	0.515***	0.683***	0.685***
	(0.04)	(0.04)	(0.06)	(0.06)	(0.06)	(0.06)
Race/Ethnicity - Black	-0.870***	-0.886***	-1.001***	-1.024***	-0.644***	-0.670***
	(0.06)	(0.06)	(0.09)	(0.09)	(0.08)	(0.08)
Race/Ethnicity - Hispanic	-0.183***	-0.196***	-0.254***	-0.269***	-0.0714	-0.0875
	(0.07)	(0.07)	(0.10)	(0.10)	(0.08)	(0.08)
Corresponding 4th grade scores	0.584***	0.576***	0.576***	0.564***	0.645***	0.631***
	(0.03)	(0.03)	(0.04)	(0.04)	(0.04)	(0.04)
Subject (Math=1)	0.150***	0.150***				
	(0.02)	(0.02)				
Evaluate all teachers each year	0.0485	0.0308	0.0308	0.00905	0.0457	0.0294
	(0.05)	(0.05)	(0.06)	(0.06)	(0.05)	(0.06)
Dismissal for instruct. ineffectiv.	0.00249	0.00258	-0.056	-0.0567	0.0438	0.0427
	(0.07)	(0.07)	(0.08)	(0.08)	(0.08)	(0.08)
Performance informs layoffs	0.0191	0.000528	0.0171	0.0056	0.0168	-0.0158
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Effectiveness for professional license	-0.0206	-0.012	0.0264	0.0209	-0.0824	-0.0456
	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
Effectiveness for tenure	0.0877*	0.0893*	0.0957	0.0957	0.0832	0.0843
	(0.05)	(0.05)	(0.06)	(0.06)	(0.06)	(0.06)
Public reporting of effectiveness data	0.0457	0.0486	0.0636	0.0779	0.0102	0.0111
	(0.06)	(0.06)	(0.07)	(0.07)	(0.07)	(0.07)
Measures of student growth req. in eval.	0.0352	0.0849	0.0598	0.129	0.0582	0.0682
	(0.06)	(0.07)	(0.08)	(0.08)	(0.07)	(0.08)
Extra pay for advanced degrees	-0.0188		-0.0129		-0.0362	
	(0.05)		(0.06)		(0.05)	
Performance pay	-0.0674		0.00383		-0.154**	
	(0.06)		(0.07)		(0.07)	
High needs differential pay	-0.145***		-0.147***		-0.126***	
	(0.04)		(0.05)		(0.05)	
High needs loan forgiveness	-0.0125		-0.0229		0.052	

	(0.06)		(0.07)		(0.06)	
Course requirements for prof. license	-0.0189		0.0219		-0.0712	
	(0.06)		(0.07)		(0.07)	
Parental educational attainment	0.139***	0.139***	0.127***	0.130***	0.138***	0.139***
	(0.02)	(0.02)	(0.03)	(0.03)	(0.03)	(0.03)
Per-pupil state expenditure	-0.0925	-0.0868	-0.105	-0.109	-0.0853	-0.0549
	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)
Prevalence of input-based policies		-0.337**		-0.298*		-0.304*
		(0.15)		(0.17)		(0.18)
Observations	660	660	331	331	329	329

Standard errors in parentheses. * p<0.10, ** p<0.05, *** p<0.01

Appendix Table B3: Effect of Categorizing Performance Pay as Outcome-based Policy

Independent variables	(1) Pooled (Reading+Math)	(2) Reading	(3) Math
ESSA (=1)	-0.218***	-0.341***	-0.0842***
	-0.0256	-0.0338	-0.0286
4th grade scores	0.538***	0.603***	0.629***
	-0.0503	-0.0744	-0.0586
Subject (Math=1)	0.117***		
	-0.0225		
Parental educational attainment	0.321***	0.275***	0.265***
	-0.0819	-0.106	-0.0844
Per-pupil state expenditure	-0.167	-0.219*	-0.0867
	-0.107	-0.132	-0.108
Prevalence of outcome-based policies plus Perf. Pay	***0.251****	***0.297****	***0.149***
	-0.0997	-0.125	-0.102
Prevalence of input-based policies minus Perf. Pay	***-0.325****	***-0.341****	***-0.276****
	-0.131	-0.165	-0.134
Observations	204	102	102

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01

Appendix Table B4: Effect of Categorizing Performance Pay as Outcome-based Policy with Differential Racial/Ethnic Growth

Independent variables	(1) Pooled (Reading+Math)	(2) Reading	(3) Math
ESSA (=1)	-0.189***	-0.333***	-0.0384
	-0.0244	-0.0369	-0.0288
Race/Ethnicity - Asian	0.621***	0.512***	0.683***
	-0.0433	-0.0617	-0.0556
Race/Ethnicity - Black	-0.881***	-1.020***	-0.666***
	-0.0619	-0.0846	-0.0818
Race/Ethnicity - Hispanic	-0.193***	-0.275***	-0.0798
	-0.0662	-0.0951	-0.0784
Corresponding 4th grade scores	0.579***	0.568***	0.633***
	-0.0305	-0.0427	-0.04
Subject (Math=1)	0.150***		
	-0.0238		
Parental educational attainment	0.138***	0.124***	0.141***
	-0.0228	-0.0323	-0.0282
Per-pupil state expenditure	-0.0367	-0.0717	0.0152
	-0.0778	-0.0844	-0.089
Prevalence of outcome-based policies plus Perf. Pay	***0.204****	***0.223****	***0.129***
	-0.0846	-0.0938	-0.097
Prevalence of input-based policies minus Perf. Pay	***-0.174***	***-0.192***	***-0.0661***
	-0.113	-0.125	-0.13
Observations	660	331	329

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01

Appendix Table B5. The Independent Impact of Performance Pay in the Racial/Ethnic Samples

Independent variables	(1) Pooled (Reading+Math)	(2) Reading	(3) Math
ESSA (=1)	-0.182***	-0.330***	-0.0286
	-0.0247	-0.0372	-0.029
Race/Ethnicity - Asian	0.624***	0.516***	0.688***
	-0.0432	-0.0619	-0.0555
Race/Ethnicity - Black	-0.885***	-1.025***	-0.670***
	-0.0618	-0.085	-0.081
Race/Ethnicity - Hispanic	-0.195***	-0.278***	-0.0834
	-0.0661	-0.0953	-0.0779
Corresponding 4th grade scores	0.577***	0.565***	0.631***
	-0.0304	-0.043	-0.0395
Subject (Math=1)	0.150***		
	-0.0238		
Parental educational attainment	0.138***	0.125***	0.140***
	-0.0228	-0.0324	-0.028
Per-pupil state expenditure	-0.0815	-0.0953	-0.0542
	-0.0809	-0.0904	-0.0899
Prevalence of outcome-based policies minus Perf. Pay	***0.259******	***0.238****	***0.239****
	-0.0872	-0.0999	-0.0974
Prevalence of input-based policies minus Perf. Pay	***-0.247****	***-0.232****	***-0.181***
	-0.119	-0.136	-0.134
Performance Pay	***-0.0789***	***-0.0229***	***-0.144****
	-0.0611	-0.0688	-0.0682
Observations	660	331	329

Standard errors in parentheses * p<0.10, ** p<0.05, *** p<0.01