Disciplined Devolution and the New Education Federalism

Martin A. Kurzweil*

In the face of congressional gridlock and failed bureaucratic efforts to address evolving challenges, policy makers have increasingly turned to alternative governance frameworks. Three such frameworks are executive waiver of legislative schemes, enlistment of states to make policy under federal programs, and experimentalist collaboration between policy makers and citizens focused on continuous learning.

Analysis of each framework in isolation has overlooked the fact that, in a growing number of significant federal programs, they interact to create a distinct and promising governance model, which this Article calls “disciplined devolution.” In these programs, a federal agency grants states waivers to design policies that deviate from a pre-existing legislative scheme, subject to a set of broad goals. The agency evaluates the initial plans and provides monitoring and feedback on their implementation. The agency also promotes collaboration with local stakeholders and benchmarks each state’s success against that of other states.
The Obama Administration’s education policies present a particularly vivid illustration of disciplined devolution. The Administration has used statutory waiver authority to allow forty-three states to replace the requirements of the No Child Left Behind Act with comprehensive education reform plans of their own devising, subject to ongoing monitoring for progress toward broad federal policy goals and efforts to incorporate stakeholders in decision making.

Analysis of these education initiatives shows that disciplined devolution can promote innovation, foster locally tailored policy, and encourage continuous institutional learning. Beyond that, disciplined devolution also answers some of the most trenchant critiques of waiver, cooperative federalism, and experimentalism, by offering greater democratic legitimacy as well as greater accountability for process and outcomes. At the same time, the disciplined devolution model represents a fragile balance of centralizing and decentralizing forces. This Article identifies and recommends steps to avoid the risk that internal tensions or external shocks will destabilize the new framework and tip it into rigid bureaucracy or unaccountable devolution.
Governance through traditional legislation and regulation is on the decline. Complex and rapidly changing economic and social circumstances have put pressure on the federal legislative process and the expansive statutory schemes that have emerged from it, raising questions about whether they are agile enough to meet modern needs.\(^1\) Congressional gridlock\(^2\) has created additional obstacles and kept obsolete statutes on the books without updates.\(^3\) At a more basic level, the *ex ante*, rule-based regulation that Congress has traditionally adopted or instructed agencies to undertake has proved inadequate in a number of rapidly changing policy domains.\(^4\)

As a result of these challenges, the federal government and policy advocates have increasingly turned to alternative governance frameworks to achieve policy ends. Three strategies in particular have received significant scholarly attention. First, many recent statutes have granted an executive officer authority to waive significant portions of the legislative scheme (what leading scholars of the subject call “big waiver”), giving the executive branch greater flexibility to address diverse and changing circumstances.\(^5\) Second, federal statutes and initiatives have shifted the site of policy making away from Congress and the federal bureaucracy by enlisting state and local governments to develop, rather than just implement, policy in the national interest.\(^6\) And

---


4. See generally Dorf & Sabel, supra note 1.

5. See, e.g., Barron & Rakoff, supra note 1, at 279–83 (discussing No Child Left Behind and the Affordable Care Act).

6. See, e.g., Gillian E. Metzger, *Federalism Under Obama*, 53 WM. & MARY L. REV. 567 (2011) (exploring a trend of providing states a significant policy-making role in federal legislation and regulation under the Obama Administration). Cooperative federalism is, of course, a long-standing phenomenon. See, e.g., Daniel J. Elazar, *American Federalism: A View from the States* (2d ed. 1972); Mortimer Z. Good, *The American System: A New View of Government in the United States* (Daniel J. Elazar ed., 1966). What I describe here is a specific form of cooperative federalism in which a federal regulatory scheme assigns states responsibilities with broad discretion to shape the federal policy within their own jurisdictions and to influence the national policy. See Metzger, supra, at 569–70. This is in contrast to other programs that have been labeled cooperative federalism, in which the states are expected to carry out federally specified policy (whether or not they actually do). See Jessica Bulman-Pozen & Heather K. Gerken, *Uncooperative Federalism*, 118 YALE
third, a variety of private and public organizations have turned away from traditional, command-and-control management to a more open-ended and collaborative problem-solving approach.\textsuperscript{7}

Big waiver, cooperative federalism, and experimentalism respond in different ways to a federal legislative process and bureaucratic structure that commentators have criticized as both ineffective and inflexible. Big waiver vests administrative agencies with the ability to alter legislative schemes to make them more responsive and differentiated.\textsuperscript{8} Cooperative federalism locates policy making closer to the ground and takes advantage of fewer legal, structural, and in some cases political constraints on state and local legislative and administrative agencies.\textsuperscript{9} And experimentalism replaces the bureaucratic structure of command-and-control rules with a set of processes designed to empower those closest to a problem, generate information, and foster continuous organizational learning about both means and ends.\textsuperscript{10}

Yet each of these models introduces or exacerbates a set of problems as well. Executive waiver authority raises concerns about democratic legitimacy and transparency, as agency officials are empowered to selectively unmake law.\textsuperscript{11} Cooperative federalism risks shortchanging national policy goals, as states may lack motivation or capacity to carry them out.\textsuperscript{12} Experimentalism may never take hold in the first place, or fade away if it does, because of disagreements among putative participants or a failure of effort.\textsuperscript{13}

Commentators and scholars have analyzed big waiver, cooperative federalism, and experimentalism in isolation. But the same scholars have largely overlooked the fact that, in a growing number of significant federal programs, the strategies intersect to create a governance model that is greater than the sum of its parts.\textsuperscript{14} In these programs, a federal agency uses waiver

---


8. See infra notes 57–66 and accompanying text.

9. See infra notes 79–86 and accompanying text.

10. See infra notes 100–07 and accompanying text.

11. See infra note 67 and accompanying text.

12. See infra notes 87–88 and accompanying text.

13. See infra notes 108–13 and accompanying text.

14. Scholars focused on the individual theoretical models have recognized that each has a relationship to the others, but have not identified the programs in which they converge as distinct. See, e.g., Barron & Rakoff, supra note 1, at 295–97 (discussing state implementation of federal spending programs and growing appreciation for experimentalism as among the factors contributing to the rise of big waiver); Dorf & Sabel, supra note 1, at 469–70 (noting affinity between American federalist
authority to permit states to design policies that deviate from a default legislative scheme, subject to a set of broad, agency-fashioned goals. The federal agency evaluates the initial plan and provides ongoing monitoring and feedback focused on states’ success in meeting the agency’s broad goals. The agency also promotes collaboration with local stakeholders in planning and implementation, and, in the most fully developed regimes, benchmarks each state’s experience against other states’ experiences.

This governance model, which I call “disciplined devolution,” has great theoretical appeal, yielding the advantages of the individual strategies with which it overlaps while mitigating some of their deficiencies. Like the more familiar models, disciplined devolution promotes policy innovation and experimentation, tailoring of policy to local conditions and capacities, and continuous institutional learning. Beyond that, disciplined devolution combines features of big waiver, cooperative federalism, and experimentalism that offset the most trenchant critiques of each.

Specifically, vesting policy making in politically accountable state and local governments while involving affected stakeholders in the policy-making process mitigates concerns about big waiver’s democratic legitimacy and transparency.15 The federal power to deny or revoke a state’s waiver, ongoing monitoring and comparative benchmarking, and the reliance built up by interstate collaboration respond to contentions that vesting states with responsibility for federal policy risks shortchanging federal goals.16 And the federal government’s ability to focus and motivate state activity through control of the waiver and the enlistment of state and local political processes addresses worries that participation in experimentalist regimes is difficult to initiate and maintain.17

While nascent versions of the disciplined devolution framework are evident in recent federal health care18 and welfare19 initiatives, a particularly vivid example is found in K–12 education. This Article uses major recent changes in education policy to explore disciplined devolution in action and to evaluate its strengths and weaknesses as a distinct form of governance. In doing so, this Article not only clarifies the attributes of disciplined devolution, it also disrupts the debate within education circles about the nature of these changes and identifies advantages as well as concerns that remain hidden under more conventional analyses.

structure and democratic experimentalism); Metzger, supra note 6, at 600–01 (recognizing that fostering experimentalism is one rationale for the federal government to empower states in the context of a federal program).

15. See infra notes 119–22 and accompanying text.
16. See infra notes 123–24 and accompanying text.
17. See infra notes 125–30 and accompanying text.
18. See infra notes 114–15 and accompanying text.
19. See infra notes 116–17 and accompanying text.
From 1965 through 2011, education federalism featured a traditional bureaucratic structure. The major federal education law of this period, the Elementary and Secondary Education Act of 196520 (ESEA)—reauthorized in 2002 as the No Child Left Behind Act of 200121 (NCLB)—was characterized by a set of specific federal rules covering an expanding set of policies and funding streams, and a bureaucratic structure at the federal, state, and local levels to monitor and demonstrate compliance. The scope of compliance monitoring excluded many important policy domains and implementation details. While states and local school districts demonstrated formal compliance with the federal rules, they would often ignore or undermine the underlying goals of the federal framework in their areas of autonomy. In some jurisdictions, promising local initiatives competed for resources or otherwise conflicted with bureaucratic demands, and as a result, were not fully realized.22 This governance approach left the United States with an education system that is far more expensive but less effective and significantly less equitable than those of other developed countries.23

The Obama Administration has upended that long-lived regime. Its 2009 Race to the Top grant program24 funded incipient state and local reform initiatives focused on school accountability and design, teacher quality, technology, and information systems.25 The program also funded the pre-existing efforts of forty-five states and the District of Columbia to develop common academic standards and assessments.26 And starting in 2011, the ESEA Flexibility program27 went a step further, allowing forty-three states and the District of Columbia to replace many of the requirements of NCLB with comprehensive education governance plans of their own devising, subject to federal review.28

The “new education federalism” that emerged from Race to the Top and ESEA Flexibility incorporates big waiver, cooperative federalism, and

---

22. See infra Part II.A.
25. See id.
experimentalism to reshape education policy making at the federal and state levels. The new education federalism establishes broad federal goals for developing academic standards, improving teacher quality, and intervening in struggling schools through plans that states are invited to develop. The state plans are meant to be comprehensive approaches to the management of K–12 public education. Further, states are required to demonstrate that they have collaborated with local stakeholders in designing the plans. The federal government evaluates and recommends revisions to the plans and monitors implementation, with a focus on the states’ progress toward the original, broad goals.29

Some critics have asserted that the Obama Administration’s initiatives represent a federal usurpation of traditionally state and local functions.30 Others contend that the effect is nearly the opposite: an abdication of federal responsibility to hold states and districts accountable.31 Viewed through the lens of disciplined devolution, however, the Obama Administration’s initiatives instead create new roles for actors at each level of government, as well as non-governmental stakeholders.

The early stages of implementing the new education programs suggest that many of the theoretical expectations of combining waiver, cooperative federalism, and experimentalism in disciplined devolution are borne out. States were motivated to participate in the new regime and undertook their responsibilities sincerely, overcoming the initiation and consensus concerns that may attach to experimentalism. Policy making by state officials and intensive stakeholder participation in the initial planning enhanced the legitimacy of a regime based on deviation from existing federal law. Many states generated policies tailored to their needs and capacities, innovated, evaluated their innovations, and adapted policy to experience. Finally, federal oversight and interstate collaboration have largely held states accountable to their commitments to the broad goals, mitigating the risk of shirking that comes with other forms of devolved policy making.32

Yet implementation of the new education federalism has also fallen short of the theoretical expectations for disciplined devolution in important ways. In particular, although individual states appear to be learning from their own experiences, best practices have mostly not spread across states. Additionally,

29. See infra Part II.C.
30. See infra note 256 and accompanying text.
31. See infra note 257 and accompanying text. These views are primarily representative of debates within the education policy community, though some legal scholars have made similar claims. Of course, legal commentators have also described the changes as part of the waiver, progressive federalism, or experimentalism trends. See, e.g., Barron & Rakoff, supra note 1, at 277–78 (2013) (describing ESEA Flexibility as an example of big waiver); Metzger, supra note 6, at 590 (describing Race to the Top as an example of progressive federalism); Sabel & Simon, supra note 7, at 56, 85 (describing Race to the Top as an example of experimentalism).
32. See infra Part III.A.
the collaboration with local stakeholders that was evident as the state plans were being drafted has faded in some states, creating a risk that local school districts and other stakeholders will come to see the policies as just another set of state mandates.\footnote{See infra Part III.B.}

Viewed more broadly, each of these shortcomings represents an unbalancing of disciplined devolution. Siloed policy making is a case of excess decentralization, while the decline in stakeholder participation means decision making is too centralized (at the state, rather than the federal level). These problems might be remedied, however, by corrections at the federal level, illustrating the lynchpin role of the federal agency in the disciplined devolution framework.

In addition to these internal tensions in the disciplined devolution model, the education case study reveals systemic risks for the governance structure. These risks include the shifting winds of politics, a hasty effort to consolidate the substantive policies of the Obama initiatives in legislation, a decrease in the credibility of NCLB as a default scheme, or a simple failure to recognize the nature and benefits of the new structure. Each could tip education federalism back to the rule-driven, compliance-oriented governance structure of the past or leave policy making to states and local school districts without the discipline of federal or interstate oversight.

In generalized form, these risks—political change, premature legislative lock-in, obsolescence of the default regime, and category error—may threaten any program premised on disciplined devolution and reveal the inherent fragility of the model. Yet these risks, although a serious concern for disciplined devolution, are not unique to it; waiver, cooperative federalism, and experimentalism bear them in their own ways, as well. Indeed, a disciplined devolution structure, while not immune to these problems, may be better suited to respond to them than alternative forms of governance.

The Article proceeds as follows. Part I elaborates the concept of disciplined devolution and its relationship to each of the three more familiar models of governance. While each model addresses the inflexible and bureaucratic nature of traditional legislation and regulation, disciplined devolution offers a balance among local variation and institutional learning, rigorous federal oversight, and democratic legitimacy that the other models lack. Part II turns to the case study of disciplined devolution in K–12 education, explaining the past fifty years of education federalism as a transition from various iterations of compliance-focused bureaucracy under ESEA and NCLB to the new framework of disciplined devolution. Part III evaluates disciplined devolution in practice by reviewing the early implementation of the new education federalism and finds many ways in which the theoretical advantages of disciplined devolution are borne out, and several ways in which
they are not. Finally, Part IV raises broader concerns about the sustainability of the disciplined devolution model and offers some tentative responses to those concerns.

I. DISCIPLINED DEVOLUTION

A. The Problems of Traditional Bureaucratic Governance

The traditional federal approach to regulation and the provision of social services is linear and hierarchical. Congress passes legislation setting out requirements, authorizes an agency to do so, or directs some combination of the two. The legislative or regulatory requirements are in the form of uniform rules that federal officials, regulated entities, or, in the case of conditional spending programs, state officials must follow. The rules are enforced through a bureaucratic structure of compliance monitoring, which is itself governed by a hierarchy and set of rules.

While there are circumstances in which this traditional approach to federal governance has been and could continue to be successful, at least since the 1970s, the traditional model has been subject to increasing criticism. The uniform and inflexible rules created through this process are a poor fit for differentiated and changing circumstances. The problems of regulatory mismatch and obsolescence have become acute as the American population has diversified and technological and economic change (including mutually reinforcing trends such as the integration of new markets, the shift from manufacturing to service, employee mobility, and complex financing) has accelerated.

Bureaucratic enforcement compounds the rigidity of the traditional regime while simultaneously facilitating unaccountable and inconsistent divergences from it. Front-line workers, or “street-level bureaucrats,” frequently have the discretion to interpret and apply the general rules with limited observation and review by superiors. Concerned that workers will apply this discretion on the basis of their sympathies rather than legislative or regulatory goals, the organization develops ever-more-detailed rules designed to limit the street-

34. Detailed legislation and bureaucratic enforcement tend to work well when the problem to be addressed is well understood, adequate and routine solutions are known, and the primary objective is efficiency in implementation. See RUDI VOLTI, AN INTRODUCTION TO THE SOCIOLOGY OF WORK AND OCCUPATIONS 85 (2008). Indeed, for much of the twentieth century, the American bureaucratic structure worked well enough to evade serious criticism. See Dorf & Sabel, supra note 1, at 277–78.
36. See generally MICHAEL LIPSKY, STREET-LEVEL BUREAUCRACY: DILEMMAS OF THE INDIVIDUAL IN PUBLIC SERVICES (30th anniversary expanded ed. 2010).
37. See id. at 13–16.
level bureaucrats’ discretion. Yet these efforts to rein in discretion often backfire, as the organization becomes less flexible and street-level bureaucrats either manipulate conflicting rules to continue to exercise discretion or resist by working abjectly. The resulting administration can be sclerotic, ineffective, and perceived as arbitrary.

National politics have exacerbated these problems. Thirty-five years of conservative attacks on government intervention have undermined public faith in government’s effectiveness, efficiency, and good intentions. Underfunded and otherwise handcuffed agencies confirm these low expectations. More recently, procedural obstruction and partisan differences have hamstrung Congress, lowering its legislative production as well as its popular standing, and generating repeated crises of governance in standoffs with the President.

One set of solutions to these problems of traditional bureaucracy seeks to further diminish the role of government. In different ways, privatization, market-based regulation, and new public management seek to fix what ails bureaucracy by giving bureaucrats precious little to do. In doing so, however, these approaches may leave important public goals and values unmet or simply replicate the bureaucratic model around a new focal point of outcome targets rather than inputs.

More recently, scholars have put forward strategies to revitalize, rather than diminish, government’s role in policy making. The following sections

---

38. See id.; see also CHARLES SABEL ET AL., INDIVIDUALIZED SERVICE PROVISION IN THE NEW WELFARE STATE: LESSONS FROM SPECIAL EDUCATION IN FINLAND 14 (Sitra Studies No. 62, 2011); JAMES Q. WILSON, BUREAUCRACY: WHAT GOVERNMENT AGENCIES DO AND WHY THEY DO IT (1989).


42. See Terkel, supra note 2.


44. See RICHARD EPSTEIN, SIMPLE RULES FOR A COMPLEX WORLD (1995).

45. See MARKET-BASED GOVERNANCE: SUPPLY SIDE, DEMAND SIDE, UPSIDE, AND DOWNSIDE (John D. Donahue & Joseph S. Nye, Jr., eds. 2002).


47. See, e.g., Sabel & Simon, supra note 7, at 71–74. But see Jody Freeman, Extending Public Law Norms Through Privatization, 116 HARV. L. REV. 1285 (2003) (arguing that privatization may enable the extension of public goals to the private sphere).

48. As I explain below, my view is that the NCLB regime was a new public management scheme that functioned as a bureaucracy. See infra text accompanying note 153.
discuss three of these governance strategies: big waiver of statutory provisions, cooperative federalism, and experimentalism.

B. Big Waiver

One strategy for overcoming congressional nonresponsiveness and rigid statutory rules is to empower members of the executive branch to deviate from and even replace the legislative scheme. Several recent scholarly works have identified and examined the use of what David Barron and Todd Rakoff call “big waiver”—a delegation to the executive branch of “broad, discretionary power to determine whether the rule or rules that Congress has established should be dispensed with.”

Not all waivers are “big” waivers: many statutes provide agencies the authority to modify a statutory requirement in a limited way to deal with exceptional cases, a sort of equitable power to relieve regulated parties of extraordinary and unintended burdens. In contrast to these “little” waivers, big waiver authority enables an agency to cancel requirements at the very heart of the statutory scheme and, in some cases, to create its own alternative requirements for the regulated parties.

While there is historical precedent for big waiver, Congress’s inclusion of broad waiver authority in recent statutes and the executive branch’s willingness to exercise that authority appear to have increased in recent years. The waiver authority provided for in ESEA and the Obama Administration’s Flexibility program created pursuant to that authority is one example that is explored at length below. Other examples include the Secretary of Homeland Security’s

49. Barron & Rakoff, supra note 1, at 267; see also Kate R. Bowers, Saying What the Law Isn’t: Legislative Delegations of Waiver Authority in Environmental Laws, 34 HARV. ENVT’L. L. REV. 257, 257–59 (2010); R. Craig Kitchen, Negative Lawmaking Delegations: Discretionary Executive Authority to Amend, Waive, and Cancel Statutory Text, 40 HASTINGS CONST. L.Q. 525 (2013); infra notes 54, 56.

50. Barron & Rakoff, supra note 1, at 277.

51. See id. at 277–78. Big waiver is also distinct from mere nonenforcement. For the duration of the waiver, it makes lawful conduct that would otherwise be unlawful under the statute. By contrast, a decision not to enforce has no effect on the legality of conduct; it is merely a case of the enforcer “look[ing] the other way.” Id. at 274.

52. See id. at 279–90; Samuel R. Bagenstos, Federalism by Waiver After the Health Care Case 1–4 (U. Mich. Pub. Law Research Paper No. 294, 2012), available at http://papers.ssrn.com/abstract=2161599. Barron and Rakoff attribute the proliferation of big waiver to a number of mutually reinforcing trends. Since the Great Society, federal spending programs that distribute benefits through state and local governments have proliferated, and Congress has established rigid and highly detailed sets of requirements to govern these programs. See Barron & Rakoff, supra note 1, at 295–96. The rigidity of these mandates on spending has in turn generated a reaction demanding greater flexibility and a set of new theories for how to offer that flexibility. See id. at 296–97, 299–301. Furthermore, the sheer number of federal statutes has resulted in a pervasive background of potentially conflicting legislation for every new piece of legislation, and Congress does not have the wherewithal to determine in advance when and where those conflicts would arise, or how they should be resolved. See id. at 301–04. Divided government populated by highly polarized partisans has made congressional resolution of conflicts and adaptation to experience and new facts even less likely. See id. at 306–09.

53. Barron & Rakoff, supra note 1, at 279–81.
authority to waive any law that inhibits construction of a border fence and a provision in the Patient Protection and Affordable Care Act (ACA) permitting states to request a waiver in order to pursue an alternative health care framework.

From a governance perspective, big waiver offers three primary advantages over traditional bureaucratic regulation. First, waiver permits policy to be tailored to the context of specific cases as they arise, while traditional regulation through rule promulgation must anticipate the need for tailoring ex ante. At the same time, a waiver regime, because it includes a baseline scheme, provides greater predictability than a regulatory regime that exclusively regulates in response to cases as they arise—for example, a regime based on adjudication.

Second, waiver facilitates policy change in response to evolving needs. For various reasons of structure and political economy, Congress is frequently slow to legislate to address new pressing needs, and previously enacted legislation is notoriously difficult to amend. Similarly, while there is some debate over the extent of “ossification,” it is clear that notice-and-comment


57. There are other advantages, from the perspective of separation of powers and administrative law. For example, big waiver allows Congress to take control of the “first draft” of a regulatory scheme, something that traditional delegation often leaves to an administrative agency. Barron & Rakoff, supra note 1, at 270. Further, when waiver authority is premised on the condition that the scheme under the waiver is at least as effective along some specified dimensions as the original legislative scheme, it creates a narrower range of discretion for the administrative agency than a delegation to do something effective, without the congressionally designed baseline. Id. Additionally, a waiver provision may save a conditional spending program from being overturned as coercive in the wake of NFIB v. Sebelius, 132 S. Ct. 2566 (2012), and states may have more leverage to take advantage of administrative waiver in the wake of that decision. See Metzger, supra note 56, at 115; Bagenstos, supra note 52, at 6–8.


59. See Barron & Rakoff, supra note 1, at 270–71.

rulemaking,\textsuperscript{61} as well as ex ante White House\textsuperscript{62} and ex post judicial review,\textsuperscript{63} make policy adjustment through traditional processes difficult and slow.\textsuperscript{64} Waiver, by comparison, offers a streamlined process for adjusting policy in the face of changing circumstances.\textsuperscript{65}

Third, some waiver regimes enable the federal government to promote innovation while maintaining a point of control. Many waiver provisions allow a statutory requirement to be waived only if alternative actions undertaken pursuant to the waiver are more effective than the original requirement in achieving the statutory goals.\textsuperscript{66} This permits experimentation on the part of the agency or a regulated entity while providing a benchmark against which to measure the success of the alternative schemes. And when a waiver ends, the default regime goes back into place, thus limiting the downside risk of innovation. If the default is burdensome, it also provides ongoing motivation to the recipient of the waiver to take its responsibilities seriously.

Unconstrained use of waiver authority, however, raises concerns about democratic accountability and legitimacy. Big waiver empowers the executive

\textsuperscript{61} See Administrative Procedure Act, 5 U.S.C. § 553 (2012); McGarity, \textit{supra} note 60.


\textsuperscript{65} In this way, it is similar to informal modes of regulation that have proliferated as a response to the difficulty of the notice-and-comment rulemaking process. See, e.g., Connor N. Rasor, \textit{Note, Strategic or Sincere: Agency Use of Guidance Documents}, 119 YALE L.J. 782 (2010); O’Connell, \textit{supra} note 64, at 936. See generally Robert A. Anthony, \textit{Interpretive Rules, Policy Statements, Manuals and the Like—Should Federal Agencies Use Them to Bind the Public?}, 41 DUKE L.J. 1311, 1319 (1992).

\textsuperscript{66} Barron & Rakoff, \textit{supra} note 1, at 270. While a number of the most high-profile big waivers are based on authority including some kind of “more-effective-than” provision, many are not. Waiver authority in both the ACA and NCLB is subject to a “more-effective-than” constraint. See 42 U.S.C. § 18052(b)(1) (2012); 20 U.S.C. § 7861(b) (2012). Other waiver provisions, such as the authority given to the Secretary of Homeland Security to waive statutory impediments to the construction of a border fence, contain no justification requirement whatsoever. See Pub. L. No. 109–13, § 102(c)(1), 119 Stat. 321, 306 (2005) (current version at 8 U.S.C. § 1103 (2012)). Barron and Rakoff recommend reading such a requirement into waivers that are silent on the matter, as a principle of administrative law. Barron & Rakoff, \textit{supra} note 1, at 332.
branch to determine at its discretion that duly enacted legislation applies to some actors in some circumstances but not to others. There is a risk that powerful interests or merely those aligned with the President’s politics will be favored and others disfavored. There is also a risk that the government may use the promise of a waiver to extract concessions that are unrelated to the purpose of the original legislation. For these reasons, libertarians such as Richard Epstein have warned that big waiver authority risks undercutting the rule of law.67

Furthermore, although big waiver provides more notice and transparency than a regime based entirely on post hoc adjudication,68 it is less predictable and transparent than a regime based on statutory rules. Under a waiver regime, regulated entities will have a hard time knowing in advance whether the law on the books will apply to them. The predictability of legal application will depend on factors such as how much discretion the relevant executive branch official has, who participates in the decision, and the level of publicity around the whole process. In the most extreme case—an official with discretion to waive law at will, without publicity—there will be a high degree of uncertainty among parties affected by the legislation’s reach.

C. Cooperative Federalism

A second alternative to the traditional bureaucratic model shifts the locus of policy making—including decisions with respect to federal policy—to the states. With some exceptions,69 political liberals spent much of the twentieth century advocating the expansion of federal power and the limitation of state power.70 The cooperative federalism programs of the Great Society reflected this preference: even as they directed funding to states for services for the poor and other disadvantaged groups, they created a set of detailed, highly specific implementation rules that suggested a lack of trust in the states’ motives.71 By contrast, devolution of decision-making authority to the states had a more conservative flavor, linked to the legacy of antebellum- and Jim Crow-era racial oppression masquerading as an interest in states’ rights.

Recently, progressive attitudes toward devolution have shifted and a new “blue” federalism is in ascendance.72 With an all-Republican federal government during portions of the George W. Bush presidency, states became

68. See supra note 58 and accompanying text.
70. See Robert A. Schapiro, Not Old or Borrowed: The Truly New Blue Federalism, 3 HARV. L. & POL’Y REV. 33, 33 (2009).
72. See David J. Barron, Foreword: Blue State Federalism at the Crossroads, 3 HARV. L. & POL’Y REV. 1 (2009); Schapiro, supra note 70.
a focus for progressive policy making. Notwithstanding periods of Democratic control of the presidency and both houses of Congress, the Obama Administration continued and expanded this approach with major legislative achievements as well as administrative decisions emphasizing independent or coordinated state policy making. This new, progressive trust in state policy making is evident in three of the most significant legislative achievements for President Obama and the Democratic Congress of his first two years in office: the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank), which limits preemption of state financial regulation, permits states to enforce federal consumer financial protection provisions against state-chartered banks, and includes state officials on federal regulatory boards; the ACA, which gives states substantial control over the expansion of Medicaid to individuals earning up to 133 percent of the federal poverty line and the creation of individual insurance exchanges; and the American Recovery and Reinvestment Act (ARRA), which allocates over a third of its funding ($282 billion out of $787 billion) to or through state and local governments, much of it with minimal restrictions on use.

Beyond purely political calculations, the incorporation of state policy-making authority in these programs recognizes that many of the most innovative government interventions of recent years have occurred at the state and local level. Health care reforms in Massachusetts and Oregon, marijuana

---

73. See Barron, supra note 72, at 1–2; Schapiro, supra note 70, at 33–34, 40–51; Ernest A. Young, Welcome to the Dark Side: Liberals Rediscover Federalism in the Wake of the War on Terror, 69 BROOK. L. REV. 1277, 1277–78, 1302–08 (2004).
74. See Metzger, supra note 6.
76. See Metzger, supra note 6, at 581–85. The legislation also requires the Consumer Financial Protection Bureau to act when a majority of states pass resolutions requesting modifications to its regulations. See id. at 585.
77. See Patient Protection and Affordable Care Act §§ 1311–13, 2001(a)(1)(C) (codified as amended in scattered sections of the Internal Revenue Code and 42 U.S.C.); Metzger, supra note 6, at 574; see also Abbe R. Gluck, Intrastatutory Federalism and Statutory Interpretation: State Implementation of Federal Law in Health Reform and Beyond, 121 YALE L.J. 534 (2011). As noted above, the ACA also permits states to deviate from the legislative scheme with their own innovative alternatives, subject to approval by the Secretary of Health and Human Services. See supra notes 55–56 and accompanying text.
78. See U.S. GOV’T ACCOUNTABILITY OFFICE, GAO-10-999, RECOVERY ACT: OPPORTUNITIES TO IMPROVE MANAGEMENT AND STRENGTHEN ACCOUNTABILITY OVER STATES’ AND LOCALITIES’ USES OF FUNDS 4 (2010). Much of the funding went to fill state budgetary gaps caused by declining revenues, with only limited maintenance of effort requirements and therefore a great deal of flexibility. Metzger, supra note 6, at 588–89. In addition, some funds were set aside for competitive grants, such as the Race to the Top program, to support state-designed policy innovations. Id. at 590.
legalization in Colorado and Washington, and the comprehensive use of data analysis in managing government services in Maryland offer some prominent examples of creative state policy making. Cities such as New York, Chicago, and Denver have also become proving grounds for proactive government efforts to promote social and economic well-being. Relatively flexible state implementation of federal programs harnesses this creativity.

Cooperative federalism permits the state government to distribute benefits or regulate in ways that are tailored to local conditions, including further devolution of authority to the local level. Local decision making yields policy that is not only a closer fit to local needs, but also more closely resembles self-government. State and local political representatives and administrative officials are more familiar (sometimes personally so) with the people affected by their policies. With less separation between government and citizenry, there also may be greater opportunity for those affected to participate in or influence the policy-making process.

Decentralizing policy formulation may also help evade national political gridlock. Inviting state participation in federal policy formulation may go further and disrupt that gridlock. States that favor the federal government’s policy are empowered to implement it and even go beyond what the federal government would do; states that generally oppose the federal government’s policy have greater leverage to change or resist it.

80. See Abby Rapoport, Washington, Colorado, and the Headaches of a Legal High, AM. PROSPECT (May 24, 2013), http://prospect.org/article/washington-colorado-and-headaches-legal-high (“When Colorado and Washington State passed ballot measures legalizing marijuana last November, they weren’t just the first states in the country to do so—they were the first governments in the world to do so.”).


83. See Metzger, supra note 6, at 600; see also Roderick M. Hills, Jr., The Political Economy of Cooperative Federalism: Why State Autonomy Makes Sense and “Dual Sovereignty” Doesn’t, 96 Mich. L. Rev. 813 (1998) (making a functional argument for state autonomy based on analogy to contracting for services).

84. But see Richard Briffault, Our Localism: Part II—Localism and Legal Theory, 90 COLUM. L. REV. 346, 446–47 (1990) (“Local autonomy, thus, should be seen as normatively ambiguous. Although it may provide opportunities for political participation and responsive government, local autonomy also contributes to the preservation of the political, economic and social status quo and to the privatization of American politics.”).

85. See Metzger, supra note 6, at 601. State participation could also serve to reinforce Congress’s goals against an executive branch seeking to depart from them. See Jessica Bulman-Pozen, Federalism as a Safeguard of the Separation of Powers, 112 COLUM. L. REV. 459 (2012).

86. See Bulman-Pozen & Gerken, supra note 6; see also Heather K. Gerken, Dissenting by Deciding, 57 STAN. L. REV. 1745 (2005).
The proliferation of these devolutionary approaches to progressive policy making also bears risks, however. Most notably, devolving policy-making authority to states without sufficient accountability for their adherence to federal goals may result in inadequate implementation of the federal policy. That risk is particularly salient when the federal government seeks to provide a benefit or protect vulnerable populations that lack political power in the states. For instance, resistance by Republican-led states to the ACA’s Medicaid expansion came at the expense of providing medical coverage to uninsured individuals.87 Further, without strong federal involvement, states may end up regulating in a way that is in their narrow self-interest but contrary to the broader national interest. For example, devolution raises the classic concern of a “race to the bottom” in the provision of social services or regulation of corporate or environmental practices.88

Moreover, although the aforementioned examples of creative policy making serve as an important counterpoint, many state and local government agencies are themselves organized as hidebound bureaucracies.89 Transferring policy-making responsibility from the federal bureaucracy to a state or local bureaucracy may result in the same rule-based, compliance-focused governance. Although disaggregating policy making may result in more differentiation by jurisdiction, the same problems of uniformity, inflexibility, and street-level arbitrariness may persist within each jurisdiction. Relatedly, national political disputes are increasingly replicated at the state or even local levels, further rendering the devolutionary strategy simply a microcosm of the traditional federal structure’s dysfunction.90

Furthermore, state and local officials may lack the resources, perspective, or expertise to address particularly challenging problems. Resource constraints may result from limited appropriations, an overly broad or burdensome portfolio of responsibilities, or both. Moreover, state and local officials are likely to be focused on the design and effects of policy in their jurisdictions, but what makes sense locally may not make sense nationally, because of either

89. The classic studies of American bureaucracy, such as LIPSKY, supra note 36, and WILSON, supra note 38, focus on state and local government.
externalities or aggregation problems. In addition to concerns about conflicting local and national incentives, state and local officials may have little awareness of successes or failures outside their jurisdiction and therefore be unable to learn from those experiences. And while policy expertise can be built through experience, the distribution of talent is such that there are undoubtedly state and local offices struggling with policy problems for which they lack expertise.

D. Experimentalism

Another alternative to traditional bureaucracy seeks to reform the bureaucratic process itself. Experimentalism is a governance structure in which “central institutions give autonomy to local ones to pursue generally declared goals[,] . . . monitor[,] local performance, pool[,] information in disciplined comparisons, and create[,] pressures and opportunities for continuous improvement at all levels.”

While experimentalist systems have evolved in both private and public organizations, in the public context, they are most closely associated with the democratic theory of John Dewey. For Dewey, democracy was “a process of collaborative discovery,” and public policies were “experimental in the sense that they will be entertained subject to constant and well-equipped observation of the consequences they entail when acted upon, and subject to ready and flexible revision in the light of observed consequences.” Dewey and modern

91. Sabel & Simon, supra note 7, at 55; see also Dorf & Sabel, supra note 1. Experimentalism shares much in common with governance structures described as “responsive regulation,” “new governance,” and “collaborative governance.” See Ian Ayres & John Braithwaite, Responsive Regulation: Transcending the Deregulation Debate 4–7 (1992) (discussing responsive regulation); Gráinne de Búrca & Joanne Scott, Introduction: New Governance, Law and Constitutionalism, in Law and New Governance in the EU and the US 1, 2–3 (Gráinne de Búrca & Joanne Scott eds., 2006) (defining the concept of new governance); Jody Freeman, Collaborative Governance in the Administrative State, 45 UCLA L. REV. 1 (1997) (discussing collaborative governance); see also Lobel, supra note 7 (discussing the similarities among new governance, democratic experimentalism, and responsive regulation).

92. Dorf & Sabel, supra note 1, at 292–315. The manufacturing process developed by Toyota in the 1970s is often held up as the prototypical illustration of experimentalism in the corporate context. See id. at 297–300. In the Toyota process, teams of employees are responsible for accomplishing broad manufacturing or progress-based goals and for providing information back to supervisors about their process. As the teams work within the framework of the best current practice, they reflect on both the process and the original benchmarks and revise them in light of experience. Errors trigger a halt in the process and are closely analyzed by both the teams and supervisors, with a goal of understanding the root source of the error and developing a solution to that source problem. See id. at 298–300; William H. Simon, Toyota Jurisprudence: Legal Theory and Rolling Rule Regimes, in Law and New Governance in the EU and the US, supra note 91, at 37, 44–55. The development and implementation of safety standards at Alcoa in the 1990s is another frequently cited example. See, e.g., Steven J. Spear, The High Velocity Edge: How Market Leaders Leverage Operational Excellence to Beat the Competition 87–108 (2d ed. 2010).


proponents of experimentalism envision policy making as a transparent and deliberative process of problem solving involving a comprehensive and diverse group of stakeholders. That collaborative “public” uses information, which emerges through close observation of implementation and outcomes, to continuously evaluate and revise both means and ends.95

Experimentalism has increasingly become a feature of federal administration, where it takes several different forms.96 One form, which Cary Coglianese and David Lazer refer to as “management-based,” involves a regulator requiring a regulated entity to develop and continuously update plans to address specific harms, which are monitored, compared, and used as a basis for intervention by the regulator.97 Another type of experimentalist regulation involves agreements either between regulators and regulated entities or among regulated entities as validated by the regulator to achieve certain observable regulatory goals.98 Finally, there is “multi-level governance,” in which regulatory authority is devolved to a lower level of government, conditioned on meeting certain standards.99

In contrast to bureaucracy’s uniform and inflexible ex ante rules, experimentalism ensures that policies are context-specific and provisional. Subject to broad outlines, local actors design policy at the local level in a collaborative effort between those with the responsibility for carrying out the policy and those who will benefit or be constrained by it.100 No policy is imposed or imported; it is either created or adapted. There is also no such thing as a final policy. Implementation is continuously monitored at the local and central level, and the information generated by that close observation is used to

95. See id. at 123–28, 148–71; Liebman & Sabel, supra note 7, at 189 n.23, 279; Sabel & Simon, supra note 7, at 55; Simon, supra note 93, at 734–35. See generally Dorf & Sabel, supra note 1.
96. See Sabel & Simon, supra note 7, at 55–56.
97. See Cary Coglianese & David Lazer, Management-Based Regulation: Prescribing Private Management to Achieve Public Goals, 37 L. & SOC’Y REV. 691 (2003); see also Sabel & Simon, supra note 7, at 83. A recent example of management-based regulation is the Food Safety Modernization Act of 2010, FDA Food Safety Modernization Act, Pub. L. No. 111-353, 124 Stat. 3885 (2011) (codified in scattered sections of 21 U.S.C.), under which the Food and Drug Administration (FDA) requires all food processing facilities above a certain size to formulate, validate, and continuously update Hazard Analysis and Preventive Controls plans. The FDA then allocates inspections according to riskiness, with risk factors including the adequacy of a facility’s plan and its implementation. See FDA Food Safety Modernization Act §§ 103, 201, 307; Sabel & Simon, supra note 7, at 85.
98. Sabel & Simon, supra note 7, at 83. The Clinton-era Project XL, under which the Environmental Protection Agency issued flexible emissions permits to corporations that promised superior environmental performance and agreed to a performance-based, ongoing evaluation, is an example of this type of negotiated experimentalism (though perhaps not a fully realized one). See Dorf & Sabel, supra note 1, at 375; Freeman, supra note 91, at 55.
99. Sabel & Simon, supra note 7, at 83. Sabel and Simon point to Race to the Top, discussed in detail below, as an example of multilevel experimentalist governance. Id. at 85. Other examples include elements of the Clean Air Act, the Clean Water Act, and the Occupational Safety and Health Act that contemplate state regulation in the relevant domain subject to minimum national standards. Id.
100. See id. at 82.
revise not only the means of achieving a policy goal but the ends, as well. Thus experimentalism may be a particularly useful organizing principle where problems are complex and uncertain.

Additionally, experimentalism places a high value on making local behaviors and decisions explicit, rather than tacit. The information-generating function of experimentalism permits local units and central authorities to evaluate not only the outcomes of various local units, but the reasons for those outcomes, enabling continuous improvement and diffusion of policy learning. The local unit itself can use these highly detailed observations to self-correct. And just as important, the central authority and other local units also have access to this information. Other local units can apply the lessons learned in another context to their own. The central authority can compare the implementation and success of the various local units, enabling it to set ambitious yet realistic goals for policy development and outcomes in the local units, target its interventions, and disseminate best practices.

Furthermore, the collaborative nature and local focus of experimentalism lend it a democratic legitimacy that is different from that of traditional bureaucratic governance, and perhaps more vigorous. The traditional model of governance is built on a conception of democracy in which the public expresses its will through elected representatives who in turn control the bureaucracy. Yet electoral decisions only reflect the preferences of majorities within electoral boundaries and only ambiguously express the majority’s will with respect to each of the issues and candidate positions that are bundled together in an election. The assumption that elected officials or even their appointees control the bureaucracy through rules of behavior also breaks down at the door of the street-level bureaucrat.

In contrast to this traditional electoral concept of legitimacy, experimentalism offers a concept of legitimacy grounded in stakeholder participation and transparency. First, where experimentalism succeeds in fostering broad stakeholder participation in the development and implementation of policy, those who are affected by government action have a direct role in setting policy. Experimentalism acts as a form of self-governance rather than representative government. Second, experimentalism makes explicit

101. See Freeman, supra note 91, at 28–29; Sabel & Simon, supra note 7, at 79.
103. See id. at 734–35.
104. See, e.g., Sabel & Simon, supra note 7, at 88–89.
107. See supra notes 36–39 and accompanying text.
the close observation and course correction at the local level, thereby facilitating more rigorous oversight by elected representatives.

Conversely, critiques of experimentalism focus on the difficulty of initiating and maintaining an experimentalist regime, as well as on the potential breakdowns that threaten experimentalism’s democratic bona fide. Bureaucracies tend to build entrenched interests, both among those who work within the bureaucracy (the street-level bureaucrats) and those who are regulated by the bureaucracy. Those with a stake in the current regime will be resistant to changing it. And on the other side of the coin, those more inclined to be collaborative may give away too much, effectively entrenching existing policy. A compounding concern is that experimentalism assumes a consensus on policy goals that may not exist in practice. This lack of consensus may lead to lengthy debate regarding ends, thereby delaying experimentation with regard to means.

Even when a diverse public does undertake experimentalist collaborative policy making, maintaining the system can be difficult. Because experimentalism requires close observation, critical self-evaluation, and constant revision, significant effort is expected of participants. It is a far more active and mentally taxing form of governance than bureaucracy. Experimentalism also requires this active engagement from more people than traditional bureaucracy, including in some cases from ordinary citizens who receive no pay for their public work.

Relatedly, experimentalism’s expectation of broad participation may be difficult to achieve in practice, threatening the model’s democratic legitimacy. Some stakeholders may not have the resources to participate (e.g., the citizen stakeholder with a full-time job), or may not have the knowledge to participate on equal terms with other stakeholders. The result may be a collaboration of experts and powerful interests, with other, less-savvy stakeholders participating as mere tokens or excluded altogether. Such an outcome could yield a


110. Cf. Lisa T. Alexander, Stakeholder Participation in New Governance: Lessons from Chicago’s Public Housing Reform Experiment, 16 GEO. J. ON POVERTY L. & POL’Y 117, 134 (2009) (noting that “[n]ew governance studies have given little attention to how race, class, gender and other social fissures impact the effective representation of marginalized constituents’ interests in such reform collaborations.”); Simon, supra note 93, at 732 (noting critique that experimentalism “sometimes treats issues as uncertain—and hence a subject for deliberation—when in fact they ought to be treated as certain, and hence a constraint on deliberation”).

111. See Minow, supra note 109, at 336.


policy-making structure as exclusive as that of traditional bureaucracy, but without the formal limits on discretion.

E. The Disciplined Devolution Model

Big waiver, cooperative federalism, and experimentalism each offer a useful corrective to bureaucratic governance but come with some costs. It is important to recognize, however, that each of these strategies has been used to describe a large set of programs. Some of these programs do not suffer from the model’s shortcomings to the same extent as others.

In particular, there is a set of programs in which all three strategies converge in a way that mitigates many of the shortcomings of each model standing alone. This convergence, which I call “disciplined devolution,” strikes a balance between the strong central oversight of big waiver and the decentralizing tendencies of cooperative federalism and experimentalism to yield a form of organization that is localized in its particulars and centralized in its overarching goals and accountability.

Three features characterize disciplined devolution:

1) A federal agency uses its authority to deviate from an existing statutory scheme. The agency will generally employ authority to waive requirements granted by the target statute itself to achieve this deviation but may also use regulatory, spending, or waiver power from a different statute.

2) The federal agency invites state or local governments to develop their own policies in the relevant domain, oriented toward broad goals identified by the agency. Each state or locality’s decision to participate is voluntary; it may instead choose to continue abiding by the original legislative scheme.

3) The federal agency evaluates and has approval authority over the state or locality’s initial plan and provides feedback to the state as part of that process. After approving a state’s plan, the agency undertakes ongoing monitoring of implementation, with a focus on whether there is progress toward the original goals. The federal agency benchmarks each state’s successes against that of other states. The federal agency also may encourage collaboration with intrastate stakeholders or other states and facilitate the sharing of information across states.

This combination of big waiver, cooperative federalism, and experimentalism is emerging in several significant and recent federal programs. The Obama Administration’s efforts to allow states to create new education reform schemes that deviate from the NCLB Act provide a particularly striking example and are the focus of much of the remainder of this Article. New programs in health care and welfare also have the potential to develop into disciplined devolution. In health care, the ACA authorizes the Secretary of Health and Human Services to waive major provisions of the law upon
application of a state that seeks to implement an alternative health care framework. The Secretary may only issue the waiver, however, if the state’s alternative framework provides health care coverage “at least as comprehensive [and] . . . at least as affordable . . . to at least a comparable number of its residents” as that required by the ACA. A recent welfare rule creates a similar structure. Pursuant to the Personal Responsibility and Work Opportunity Reconciliation Act of 1996, the Secretary of Health and Human Services has recently issued guidance that invites states to apply for waivers from the law’s welfare-to-work requirement in order to implement program structures that are more likely to promote a transition to the workforce.

Disciplined devolution addresses the problems of bureaucracy in many of the same ways as the strategies with which it overlaps. Like big waiver, disciplined devolution accommodates diversity and flexibility in policy making, while still using Congress’s preferred scheme as a baseline. Like cooperative federalism, it encourages state and local governments to take ownership of federal goals, tailor policies to local conditions, and sidestep national political disputes. And like experimentalism, disciplined devolution establishes a process of self-evaluation and adaptation in light of experience that also produces information about policy success and failure to be used by the federal government or shared across states.

Yet disciplined devolution is not simply coterminous with the other governance models. There are programs qualifying as one or more of those governance approaches that do not qualify as disciplined devolution. For example, while the Secretary of Homeland Security’s authority to waive any statutory provision that inhibits the construction of a border fence is a prime example of big waiver, it does not involve state and local policy making and in no way incorporates experimentalist features. Likewise, while most of the state aid provisions of the American Recovery and Reinvestment Act vest significant policy-making authority in the states, they do not give states that authority as a deviation from an existing legislative scheme and do not use continuing federal oversight to promote experimentation and institutional learning. The private sphere applications of experimentalism certainly do not fit the mold of disciplined devolution, and neither do the public sphere experimentalist

114. See 42 U.S.C. § 18052(a)(1)–(2), (b)(1) (2012); see also Barron & Rakoff, supra note 1, at 281–84.
115. 42 U.S.C. § 18052(b)(1). The alternative framework also must “not increase the Federal deficit.” Id.
118. See supra note 78 and accompanying text. Race to the Top, which was created from ARRA funds, is a notable exception described in more detail below.
programs employing direct regulation of or contracting with regulated entities, because they do not involve the states. Experimentalist, multi-level governance schemes tick two of the three boxes for disciplined devolution, but not all of those programs are triggered by a deviation from a default legislative scheme.

What makes disciplined devolution distinctive and worthy of attention is that it combines offsetting features of each of the individual models. It is conceivable that the combination of waiver, cooperative federalism, and experimentalism could cancel out the positive aspects of one or more of the other strategies. But as I explain below and explore through the example of K–12 education, there is reason to believe that disciplined devolution achieves a complementary balance, in that features of each of the converging governance strategies address potential weaknesses of the others while preserving their benefits.

**Disciplined Devolution v. Waiver**—As discussed above, one of the core concerns about big waiver is that it allows executive agencies to pick and choose what law will apply for different citizens or entities. This raises concerns about the fairness and democratic legitimacy of the practice as well as the transparency and predictability of the legal regime. In disciplined devolution, however, state or local officials choose whether to pursue a waiver and assume the burden of policy design. Compared to other waiver regimes, this transfers a significant degree of discretion from the federal agency to actors who are both more informed about and more accountable to the beneficiaries of the regulation. Experimentalist procedural expectations of extensive stakeholder engagement and the generation and sharing of information about policy implementation further reinforce democratic norms by involving the beneficiaries themselves in the deliberative process.

In light of this, there is a strong argument that disciplined devolution is more democratically legitimate than agency delegation typical of traditional bureaucratic governance. Congress’s delegation to agencies to fill in the interstitial details of legislation depends on Congress’s initial imprimatur and presidential oversight for its legitimacy. As discussed above, the chain of accountability from people’s initial electoral choices, through congressional or presidential action (or inaction), and finally to the day-to-day decisions of agency staff is a long one and may easily break. By comparison, those affected by government action are more directly involved and more influential in the disciplined devolution framework.

119. See supra notes 67–68 and accompanying text.
121. See supra note 105 and accompanying text.
122. See supra notes 105–07 and accompanying text.
Because the waiver applicants are state or local governments, there is also likely to be greater transparency about the waiver process than under other circumstances. For example, states may have to pass legislation regarding their plans or applicants may have to comply with local and open government requirements. State or local applicants that have to account to local constituents and interest groups may also put pressure on the federal agency to make its decision-making process transparent. Experimentalist deliberation and information sharing also enhance visibility into the process.

**Disciplined Devolution v. Cooperative Federalism**—The disciplined devolution framework also mitigates concerns that state and local governments in a devolved policy-making structure will fail to vindicate national goals through misaligned interests, lack of information, or shortfalls in capacity._first_ approval of initial plans and ongoing monitoring provide the federal agency with opportunities to correct states that stray from the federal goals. The ability to deny or cancel the waiver gives the agency the leverage to reinforce those goals._second_ The experimentalist process bolsters the alignment of interests by encouraging states to repeatedly self-evaluate their policies in light of the federal goals. This process also generates information for the federal agency about whether the states are on target. The experiences of different states or localities may also lead the federal agency to reexamine its goals—reorienting toward weaknesses or underlying problems that come to light, or raising expectations for all states based on some states’ successes.

Second, the federal agency in a disciplined devolution program can share information across states and localities, enabling the spread of best practices that may be missing from other devolved governance structures. If the program has an interstate collaboration element, this too can facilitate the sharing of information across jurisdictions. And the state and local governments are better able to synthesize the policy knowledge coming from the federal government and other jurisdictions as a result of their ongoing process of policy experimentation.

Third, disciplined devolution can build problem-solving capacity and expertise among state and local actors through collaborative learning. The process encourages discussion of the effects of different policy details among local participants, facilitating the rapid diffusion of policy knowledge. Repeated self-evaluation highlights the most effective practices and weeds out those unrelated to the program’s goals. It also channels information to the federal agency in a form that allows it to identify areas in which capacity is lacking. This better prepares the agency to intervene in order to vindicate overarching regulatory objectives.

---

123. See supra notes 87–90 and accompanying text.

124. I discuss the logic behind the motivational power of the waiver regime below. See infra note 129 and accompanying text.
Disciplined Devolution v. Experimentalism—Critiques of experimentalism highlight the difficulty of initiating and maintaining such a regime.\textsuperscript{125} Disciplined devolution addresses the need for focus, motivation, and participation that underlie these critiques. First, the federal agency in a disciplined devolution program constrains the choices of states by setting the initial goals and, depending on the circumstances, the initial policy options. While it is important for the agency to consider relaxing these constraints as state and local experience with policy implementation reveals dead ends or new possibilities, the initial limitation takes some topics off the table for deliberation, reducing the likelihood of a failure of consensus and focusing state and local actors on a narrower set of decisions.\textsuperscript{126}

Importantly, in contrast to multilevel experimentalist programs that do not operate against a default legislative scheme, the waiver component of disciplined devolution may make it easier for the federal agency itself to set those initial goals and policy constraints. The agency’s goals will be tied to explicit or implicit goals in the statute. The statute may limit the scope, duration, or purpose of waiver.\textsuperscript{127} And at the very least, the statutory scheme will set the baseline against which any state plan will be compared.\textsuperscript{128}

Second, disciplined devolution motivates state and local actors to maintain the effort of the experimentalist process with the possibility of reversion to the default regime. The possibility of reversion is likely to be a more credible motivator than the risk of losing or the promise of gaining federal funding, which is the usual incentive for states to participate in cooperative federal programs. Nondiscretionary spending and regular appropriations develop constituencies that make it politically difficult to cut off funding. Temporary grants lose their power to motivate as the funding runs out. In both cases, inertia goes against the incentive for state and local participation—doing nothing means that the incentive (more temporary funding or cutting off appropriations) disappears. But in disciplined devolution, reversion to the default regime is the automatic result when a waiver expires. Doing nothing means that the motivating threat is carried out. This arrangement gives the federal agency in a disciplined devolution regime significant leverage vis-à-vis the states, as well as natural points at which to exercise its power as it reviews waiver applications, milestones, and possible renewal.\textsuperscript{129}

\begin{itemize}
\item \textsuperscript{125} See supra notes 108–13 and accompanying text.
\item \textsuperscript{126} See supra note 110.
\item \textsuperscript{127} For example, the waiver provision in NCLB limits waivers to four years (plus extensions), and makes certain provisions unwaivable. No Child Left Behind Act § 9401(c), (d), Pub. L. No. 107-110, 115 Stat. 1425 (2002) (codified as amended in scattered sections of 20 U.S.C.).
\item \textsuperscript{128} The ACA waiver, for example, explicitly requires waiver plans to be on par with or better than the default scheme along several dimensions. See supra note 115 and accompanying text.
\item \textsuperscript{129} This logic assumes that the waiver granted to states is either temporary or dependent on the continuation of some condition, both reasonable assumptions in the context of disciplined devolution regimes as I have described them. It also assumes that the agency is motivated to have the states continue the disciplined devolution regime, that the default statutory scheme can be
Third, disciplined devolution’s reliance on state and local governments as well as the strength of the federal agency’s position can help encourage the broad stakeholder participation necessary to experimentalism’s democratic bona fides. The state and local officials who participate in policy planning may themselves be elected or report to elected representatives; they may accordingly be sensitive to constituent interests or likely to seek inclusion of constituent voices in the process. Along the same lines, state and local governments have existing forums for stakeholder participation, which can be adapted to suit the deliberative requirements of disciplined devolution. Advocacy groups and interested citizens are already aware of these forums, familiar with state and local government processes, and primed to participate. The federal agency can make broad stakeholder participation a condition of the waiver, and use its points of control—initial evaluation and ongoing monitoring—to observe the process and reinforce the principle.

* * *

Disciplined devolution has a number of theoretical advantages over traditional bureaucratic governance as well as each of the three models with which it overlaps. But the model is not a panacea. Disciplined devolution may be poorly executed, or countervailing factors may undermine the theoretical benefits. The next Section takes up K–12 education as a case study to assess whether the theoretical advantages of disciplined devolution exist in practice or are replaced by other deficiencies.

II.
THE EVOLUTION OF EDUCATION FEDERALISM

American public education, after the Obama Administration initiated changes to the federal role, presents an instructive illustration of disciplined devolution in action. Education commentators and legal scholars have misinterpreted the changes in education federalism as examples of big waiver, cooperative federalism, or experimentalism. This Article clarifies that the new structure can and should be properly understood as disciplined devolution. Viewing the new education federalism as disciplined devolution will help both those working within the system and those outside it trying to hold government actors to account. As explained in further detail below, a misapprehension of the nature of governance model could lead the Administration or state actors to take actions that undermine its effectiveness; understanding the model is a necessary condition for refinement and improvement. Similarly, those outside the system should know what to expect from it and critique both the

implemented, and that Congress does not intervene to change the statutory scheme. These assumptions may not hold; I explore the consequences of their possible breakdown in the context of the Obama Administration’s education reforms in Part IV, below.

130. Examples of these existing forums include legislative hearings, state advisory boards, town hall meetings, local board meetings, advocacy group meetings, and local citizen councils.
mechanism and the results; neither can be done effectively without an accurate view of what the system is.

Understanding the significance of the new education federalism requires an overview of the system of bureaucratic governance from which it emerged. Thus, Part II.A briefly surveys education federalism from the passage of ESEA in 1965 through the implementation of NCLB during the presidency of George W. Bush. Part II.B explains the Race to the Top and ESEA Flexibility programs that the Obama Administration introduced to spur innovations that were constrained by the old regime and to grant relief from NCLB’s requirements. Finally, Part II.C explains how the Obama initiatives differ from the ESEA/NCLB regime and create a new framework of disciplined devolution.

A. From ESEA to NCLB

The Elementary and Secondary Education Act of 1965 directed funding to states to be redistributed to local school districts based on the number of low-income students they served. While the statute was initially vague about how the federal funds should be spent, after several years of widespread misappropriation by districts, congressional amendments and federal Office of Education regulations added numerous detailed requirements. Congress also increased the size of the federal education bureaucracy and funded a substantial increase in states’ education bureaucracies to monitor compliance with ESEA’s requirements. This regime, in which federal and state bureaucracies created and enforced compliance with thousands of detailed


135. See McGuinn, supra note 131, at 37.
administrative rules for school districts, would persist with little variation through the 1990s.

Consistent with traditional bureaucratic governance, ESEA’s requirements for how federal funds could be spent were input oriented and uniform.136 Rules established categories of expenditures and limits for each category, as well as reporting requirements designed to facilitate federal audits.137 Districts were required to spend funds on schools with the highest concentration of poverty, to supplement rather than supplant local funds, to provide the same comparable amounts of local funding to all schools, and to maintain fiscal effort from year to year.138 Schools were permitted to spend ESEA funds only on compensatory education for those students with learning deficiencies, which frequently meant pulling such students out of class to receive separate instruction for the sole reason that this practice made it easier to track the funding.139

Evidence began to emerge in the mid-1970s that ESEA programs and funding had largely failed to improve educational opportunities for disadvantaged students.140 Among other reasons, ESEA’s process rules diverted school officials’ and teachers’ attention and resources away from instruction and to the administrative tasks and reporting demanded by state and federal bureaucracies.141

---

136. ESEA’s focus on administrative requirements also reflected resistance to federal micromanagement of schools by local officials and national Republican politicians, as well as a lack of consensus on how to measure the effectiveness of reform efforts. See Mcguinn, supra note 131, at 36–37.

137. See chubb, supra note 133.


140. See, e.g., Stephen K. Bailey & Edith K. Mosher, ESEA: The Office of Education Administers a Law (1968); Joel S. Berke & Michael W. Kirst, Federal Aid to Education: Who Benefits? Who Governs? (1972); Milbrey Wallin McLaughlin, Evaluation and Reform: The Elementary and Secondary Education Act of 1965, Title I (1975); see also Mcguinn, supra note 131, at 39 & n.59 (collecting research). This is not to say that there was no improvement in educational outcomes of disadvantaged students during this period. In fact, researchers have traced significant improvement in the educational outcomes of Black students during this period to integration, particularly the integration of schools in the South. See James E. Ryan, Schools, Race, and Money, 109 Yale L.J. 249, 297–304 (1999) (describing research findings). To the extent ESEA reinforced integration, then, it can be said to have led to an improvement in outcomes for disadvantaged students. See also Gary Orfield, The Reconstruction of Southern Education: The Schools and the 1964 Civil Rights Act 102–50 (1969).

141. See Richard F. Elmore, Differential Treatment of States in Federal Education Policy, 60 Peabody J. Educ. 34, 41 (1982) (noting that “[t]ight regulations and guidelines appear only to produce pro forma ‘compliance behavior’”).
By the beginning of the 1980s, it was widely believed that the United States was lagging behind other countries in educational performance and that bureaucratic federal intrusion in education was a major reason for the shortcoming. President Reagan sought to harness this frustration by constraining the federal education bureaucracy. In contrast, the reaction in many states was to shift bureaucratic focus to student outcomes: raising academic standards, instituting curriculum requirements, and creating high-stakes assessments for students.

Limited federal efforts to deregulate education eventually gave way to efforts to appropriate the popularity of standards-based reform. In 1989, President George H.W. Bush, along with the National Governors Association, convened an education summit with reform-minded governors (including Arkansas’s Bill Clinton) to discuss the core elements and challenges of standards-based reform. While Bush was unsuccessful in moving legislation, President Clinton signed the Improving America’s Schools Act of 1994.

---


143. The Republican Party’s broader critique of federal intervention had particular resonance for the public in the area of K–12 education, in light of the deep unpopularity of federally (typically court-) mandated busing in support of integration. See McGuinn, supra note 131, at 39; James, supra note 134, at 190.

144. James, supra note 134, at 190. Between fiscal years 1981 and 1988, the budget for the Education Department was cut by 11 percent in real dollars. McGuinn, supra note 131, at 45–46.

145. McGuinn, supra note 131, at 22, 46; Fuhrman & Elmore, supra note 134, at 82. By 1984, thirty-five states had set new graduation requirements, twenty-two instituted curriculum reforms, and twenty-nine adopted new testing policies. Viteritti, supra note 142, at 2092 (citing U.S. Dep’t of Educ., *The Nation Responds* 144–46 (1984)). The standards-based reforms in the states had a mixed record. In many states, the academic requirements lacked enforcement mechanisms and local districts retained significant flexibility in how they implemented the state reforms. See Fuhrman & Elmore, supra note 134, at 86–87. Although this soft touch led to uneven implementation across states and regions, significant local initiatives to implement the reforms or parallel initiatives occurred in many states. See id. at 88.


147. See id. at 60–61.
which required states receiving ESEA funds to set benchmarks for schools based on their standards and assessments and to impose “corrective action” on schools that failed to meet those benchmarks. Yet by the end of 2001, only nineteen states were in compliance with IASA assessment requirements, and none of the noncomplying states lost funding.

By 2002, Republicans (led by President George W. Bush) and congressional Democrats (led by Senator Ted Kennedy) were able to reach a compromise to add teeth to federal standards and testing requirements attached to ESEA funding. The NCLB Act has been both extolled and excoriated for shifting focus to student outcomes as measured by standardized tests and away from provision of inputs under ESEA. This outcome-based approach had some important positive effects—including highlighting previously hidden educational disparities and facilitating some state and local policy innovations. But NCLB’s inflexibility in some areas, lack of oversight in other areas, and compliance-focused administration limited those positive impacts and caused other significant instructional problems. Although it superficially had the features of new public management, NCLB operated as an outcome-focused version of traditional bureaucratic governance, with all the limitations of such an approach.

At the core of NCLB is a set of requirements for states to receive funding under ESEA related to testing, school and district accountability, interventions in struggling schools, and teacher qualifications. First, states must adopt standards in literacy and math for all grade levels and create annual assessments in those subjects for grades three through eight and one year in high school. Scores on the state tests must be reported publicly each year at the school, district, and state level for all students, as well as for defined

---

149. See McGUINN, supra note 131, at 94–95.
150. See MARIS A. VINOVSKIS, FROM A NATION AT RISK TO NO CHILD LEFT BEHIND 172 (2009). While IASA nominally tied compliance with its mandates to receipt of federal education funds, the timelines for completion and penalties for failure to comply were weak and unclear. See McGUINN, supra note 131, at 97–98.
151. To attract Democratic support, the bill included a 20 percent increase in federal education funding, preserved programs for class size reduction and school construction, and eschewed private school vouchers. To quell conservative concerns about the federalization of instruction, it left the creation of standards and assessments to the states. See McGUINN, supra note 131, at 176. President Bush’s popularity and feelings of national unity in the wake of the attacks of September 11 also played a role in the bill’s passage. Id.
152. While many of reform elements of NCLB were similar to those in Clinton’s IASA, NCLB enforced its provisions with strict deadlines and the threat of funding penalties for noncompliance. See McGUINN, supra note 131, at 182.
153. See supra note 46.
demographic subgroups including each major racial and ethnic group, students in poverty, students with disabilities, and English language learners.\footnote{155}{Id. § 1111(b)(3)(C)(xiii), (h)(1)(C)(i).}

Second, NCLB requires states to establish annual targets for the percentage of students meeting grade-level proficiency standards, with the eventual goal of all students in the state meeting those proficiency standards within twelve years of the law’s passage (i.e., by 2014).\footnote{156}{Id. § 1111(b)(2)(F)–(H). A “safe harbor” provision means that a school, district, or state could be compliant while never reaching an annual target or the target of having all students achieve proficiency. Id. § 1111(b)(2)(i)(i).} For a school or district to make “adequate yearly progress” (AYP), it must meet the target percentage of students achieving proficiency for all students as well as each demographic subgroup of students.\footnote{157}{Id. § 1111(b)(2)(C), (I).}

Third, the law requires an escalating series of interventions for schools and districts that fail to make AYP.\footnote{158}{While the testing, reporting, and AYP provisions apply to every school in the state, consequences for failing to make AYP are reserved for schools receiving federal Title I funds. Id. § 1116(b)(1)(A), (5), (7), (8).} Schools that miss AYP for two consecutive years are designated for “school improvement” and must develop a plan for improving student achievement and set aside funds for professional development.\footnote{159}{Id. § 1116(a)(1), (b)(3), (b)(4).} Students in those schools are also entitled to transfer to another public school within the same district.\footnote{160}{Id. § 1116(b)(1)(A), (E).} After three consecutive years of failure to make AYP, schools must offer their students after-school tutoring services by contracted providers.\footnote{161}{Id. § 1116(b)(5), (e)(1).} After four consecutive years, schools must take one of several “corrective actions,” including replacing school staff or introducing new curriculum.\footnote{162}{Id. § 1116(b)(8).} And after five consecutive years of missing AYP, schools undergo “restructuring,” whereby either management of the school is taken over by a private company or the state, or the school is converted to charter status.\footnote{163}{Id. § 1111(b)(2)(i)(i).} One year of making AYP keeps a school in the same status (with the same interventions) and two consecutive years of making AYP returns a school to “good standing.”\footnote{164}{Id. § 1119(a)(1), (3). For schools not receiving Title I funds, all teachers of “core academic subjects” had to be “highly qualified” by 2005–06. Id. § 1119(a)(1), (3). For schools not receiving Title I funds, all teachers of “core academic subjects” had to be “highly qualified” by 2005–06. Id. § 1119(a)(2); 34 C.F.R. § 200.55 (2003).}

Finally, NCLB requires most teachers employed in public schools to be “highly qualified.” Teachers are considered highly qualified if they have a
state teaching license appropriate for the grade level and subject they teach and if they either majored in the subject in college, pass a state test in that subject, or otherwise demonstrate knowledge in the area.\footnote{166}{No Child Left Behind Act § 9101(23), Pub. L. No. 107-110, 115 Stat. 1425 (2002) (codified as amended in scattered sections of 20 U.S.C.); 34 C.F.R. § 200.56 (2015).}

NCLB had a substantial effect on the way states approached education policy, including some distinctly positive effects. The law’s testing and reporting requirements amplified the focus on student achievement that emerged in the 1980s and reorganized federal, state, and local bureaucracies around these requirements.\footnote{167}{See McGuinn, supra note 131, at 182; Michael Heise, The Political Economy of Education Federalism, 56 EMORY L.J. 125, 141–48 (2006).} The disaggregation of results by demographics revealed previously hidden but grave disparities in educational outcomes across and within schools.\footnote{168}{See PAUL MANNA, COLLISION COURSE: FEDERAL EDUCATION POLICY MEETS STATE AND LOCAL REALITIES 150 (2011).} NCLB also gave local policy entrepreneurs political cover to pursue bold reforms, some of which built on the requirements of NCLB and some of which differed significantly from them.\footnote{169}{See id. at 151–52. Some of the most significant local reforms that emerged during the NCLB period are described infra notes 180–83 and accompanying text.}

NCLB did not, however, change education federalism in a fundamental way, as some have suggested,\footnote{170}{See Patrick McGuinn, The National Schoolmarm: No Child Left Behind and the New Educational Federalism, 35 PUBLIUS 41, 68 (2005) (“Regardless of whether NCLB ultimately improves schools or student achievement, the law has created a new educational federalism in the United States.”).} nor did it end up leading to the kind of experimentalist continuous learning cycle that others predicted.\footnote{171}{See Liebman & Sabel, supra note 7.} NCLB’s approach to instruction and school management was simultaneously too directive and not directive enough.\footnote{172}{See James E. Ryan, The Perverse Incentives of the No Child Left Behind Act, 79 N.Y.U. L. REV. 932, 987 (2004) (“In an attempt to drive education policy without intruding too greatly upon state authority, the federal government has combined regulatory stringency regarding AYP with regulatory laxity regarding the quality of standards and assessments.”).} In the areas of school evaluation, consequences for poor performance, and teacher qualifications, the law was uniform and directive, with numerous, specific requirements that were enforced through bureaucratic monitoring. But in the area of standards and testing, the federal government was uninvolved, and every state had a different approach. Both practical challenges and a perverse incentive structure yielded lower standards and low-quality tests. The result was typical of bureaucratic governance: a great deal of activity in states and districts to show compliance with the specific requirements of the law, but little connection between that activity and the underlying goals of the legislation to improve student outcomes and reduce achievement gaps.

The law’s specific and uniform requirements were problematic for at least three reasons. First, NCLB’s requirements did not accommodate differences in
state capacity to meet them. Many states struggled to meet the law’s deadline for establishing annual tests by 2006. States also struggled with implementation of after-school tutoring and the transfer option, and take-up rates for eligible students were consistently low. No state achieved compliance with the highly qualified teacher requirements by the 2006 deadline, and, after several extensions, the Education Department abandoned enforcement. Paradoxically, states with the most developed accountability systems before NCLB often had the most difficulty reconciling their preexisting policies with the new federally imposed ones.

Second, the requirements were in many cases poorly aligned to actual circumstances in schools, districts, and states. For example, the transfer option has little purchase in rural areas with long distances between schools or in districts where there are too few seats in good schools to receive transferring students. Additionally, although states had some flexibility to front-load or back-load their annual targets, the law’s requirement that all students achieve proficiency by 2014 ignored differences in students’ and states’ starting points and circumstances, as well as differences in students’ learning trajectories.

Third, NCLB’s requirements limited policy innovation. For many jurisdictions, NCLB’s requirements took up time and resources that might have been devoted to pursuing alternatives and relieved state and local officials of the responsibility to think critically about standards, testing, and interventions. Other states and districts did innovate, but NCLB was a hindrance to fully realizing the benefits of these innovations. For example, several important developments in the use of student test results emerged in the decade following NCLB’s passage, including sophisticated methodologies for measuring student learning growth and the value added by particular

---

173. MANNA, supra note 168, at 44–45. Once the tests were in place, some state education agencies lacked the capacity to monitor their contractors’ scoring practices or provide results in a timely manner. Id. at 50–51.

174. Id. at 51–52.

175. Id. at 76–77.

176. Id. at 58.

177. Id. at 45–46. As a result, NCLB may have slowed progress in those states if for no other reason than state and local officials had to divert focus from improving the systems they had set up to complying with the new federal requirements.


179. See MANNA, supra note 168, at 154 (“[NCLB] caused concerns over procedural bureaucratic requirements to mushroom, despite promises from the law’s advocates that it would stress substantive outcomes, such as increasing student achievement and decreasing achievement gaps. As a result, state and local education officials spent much time showing how they had complied with those procedural expectations.”).

180. See Damien W. Betebenner, Toward a Normative Understanding of Student Growth, in THE FUTURE OF TEST-BASED EDUCATIONAL ACCOUNTABILITY 155 (Katherine E. Ryan & Lorrie A. Shepard eds., 2008).
Results using these methodologies conflicted with and raised doubts about NCLB’s proficiency-based AYP measures and highly qualified teacher criteria. With few exceptions, the jurisdictions experimenting with growth and value-added studies nevertheless had to comply with federal requirements. Another underutilized innovation was the so-called “portfolio” strategy, in which a district would promote parental choice among a diverse set of schooling options by allowing schools to operate free of district mandates. A portfolio district would also closely monitor school performance and intervene in or replace struggling schools with new schools, including charter schools. While NCLB is conceptually consistent with this strategy, its timeline and specific requirements are not, and many portfolio districts ended up developing performance measures and interventions that ran in parallel to NCLB, duplicating effort and causing confusion among educators and the public.

NCLB created as many problems by offering no specific guidance as by being too specific in its guidance. In contrast to the highly detailed rules governing testing process, reporting, accountability, and teacher qualifications, NCLB left the content of standards and tests entirely to the states. The statute requires states to establish “challenging” academic content standards in math, literacy, and science, and three achievement thresholds (basic, proficient, and advanced) aligned to those standards. There are no specifications with respect to either the content standards or the achievement levels and, in fact, the law provides explicitly that states “shall not be required to submit [their] standards to the Secretary.” The one, loose check on the quality of state standards is that NCLB requires participation in the National Assessment of Education Progress (NAEP)—a long-standing, federally designed and administered set of assessments periodically offered to a sample of students.


184. See, e.g., Stacy Childress et al., Managing for Results at the New York City Department of Education, in Education Reform in New York City: Ambitious Change in the Nation’s Most Complex School System 87, 95–97 (Jennifer O’Day et al. eds., 2011).


186. Id. § 1111(b)(1)(A).

187. Id. § 1111(c)(2).
which was voluntary before NCLB. By comparing states’ results on their own tests and NAEP, researchers and the public could determine whether state tests were more or less challenging than NAEP.

The structure of NCLB created a perverse incentive for states to set low standards and create poor quality tests. NCLB’s combination of federally defined, rapidly increasing performance targets and penalties, on the one hand, and state control of standards and testing, on the other hand, led many states to set low cutoff scores to more easily meet performance targets. As a result, for nearly every state, a lower percentage of students achieved proficiency on NAEP than on the state test. Furthermore, the rapid pace at which tests had to be designed, insufficient state appropriations to develop the tests, and the lack of capacity of testing contractors led to cheaper and lower quality assessments.

Despite the apparent “dumbing down” of states’ standards and tests, a design flaw in the rules governing AYP ensured that many schools were identified for school improvement and that the number would ratchet up significantly over time. In 2002–03, the first year of implementation, 28 percent of schools nationwide failed to make AYP. By the 2007–08 school year, 35 percent of schools failed to make AYP. Some states had as many as 76 percent of schools miss AYP for a given school year.

The combination of testing to low-quality standards and the idiosyncrasies of AYP had a negative effect on instruction. Schools began to emphasize the tested subjects of literacy and math to the exclusion of other subject areas. As the state tests typically covered a broad set of standards and skills with superficial multiple-choice questions, curricula and instruction tended toward broad and shallow coverage, rather than deeper knowledge or higher-order skills.

---

189. See Manna, supra note 168, at 48–49; Ryan, supra note 172.
190. See Manna, supra note 168, at 48. Although the NAEP requirement revealed the decline in standards, revelation apparently did little to prevent it.
191. See id. at 158.
192. For reasons such as measurement error, random variation, and the fact that the students in a school change each year, schools are far more likely to experience volatility than consistent improvement in the percentage of students achieving proficiency based on test scores. See id. at 73; Thomas J. Kane & Douglas O. Staiger, Volatility in School Test Scores: Implications for Test-Based Accountability Systems, in Brooking Papers on Education Policy 235 (Diane Ravitch ed., 2002). Further, because a school overall fails to make AYP if any of its disaggregated demographic subgroups fails to make AYP, schools with more demographic subgroups have more opportunities to fall short. See Manna, supra note 168, at 74.
194. Id.
195. Id.
196. Id. at 115–17.
197. Id. at 109–11. Furthermore, because the only measure that mattered for AYP was the percentage of students above the state’s proficiency cutoff score, many teachers, principals, and school
Viewed in this way, NCLB suffered from the same bureaucratic governance problems as ESEA. Although NCLB expanded federal rules into new instructional and managerial areas, like ESEA, the focus of the federal-state interaction was on administrative compliance with those rules. And like ESEA, NCLB did not subject core policy domains (standards and testing) to federal evaluation or oversight. And, again, like ESEA, NCLB shifted resources and attention to compliance monitoring in the areas covered by the law and permitted states and local districts to use the unmonitored areas to undermine the underlying goals of the legislation.

B. The New Education Federalism

Almost from its inception, NCLB generated resistance from states, school districts, and teachers’ unions. Some states and districts brought suit (ultimately unsuccessfully), contending that some of the law’s requirements violated a prohibition on unfunded mandates, or that the conditions and penalties attached to the federal funding were too vague and were not within Congress’s authority under the Spending Clause. By 2006, six states had passed laws or resolutions expressing opposition to NCLB, and another thirty-two states had such a bill making its way through the legislative process. Although NCLB was due for reauthorization in 2007, Congress failed to move any bill to a vote that year, leading to the first of several temporary extensions of the existing legislation.

After reauthorization failed, states more aggressively rebelled against NCLB, with some declining to reset proficiency targets for their schools to remain on track for the 2014 deadline. The state resistance mirrored negative public opinion regarding NCLB. For example, in an August 2009 Gallup poll, 45 percent of respondents who reported familiarity with the law believed NCLB had “not made much difference” in the education received by public districts adopted a strategy of focusing on “bubble students”: students that in previous years or on practice exams were just above or just below the proficiency threshold. Id. at 137–38.

198. Connecticut v. Duncan, 612 F.3d 107 (2d Cir. 2010). Section 9527(a) of NCLB prohibits any “officer or employee of the Federal Government to mandate, direct, or control a State, local educational agency, or school’s curriculum, program of instruction, or allocation of State or local resources, or mandate a State or any subdivision thereof to spend any funds or incur any costs not paid for under this Act.” No Child Left Behind Act § 9527(a), Pub. L. No. 107-110, 115 Stat. 1425 (2002) (codified as amended in scattered sections of 20 U.S.C.).

199. Sch. Dist. of Pontiac v. Sec. of the U.S. Dep’t of Educ., 584 F.3d 253 (6th Cir. 2009) (en banc).


201. Sam Dillon, For a Key Education Law, Reauthorization Stalls, N.Y. TIMES, Nov. 6, 2007, http://www.nytimes.com/2007/11/06/washington/06child.html. Failure to reauthorize meant that the law would continue in the same form; a provision of NCLB permits Congress to continue funding the law’s current programs in the absence of reauthorization or amendment. Id.

school students, 29 percent believed it had made education worse, and just 21 percent believed it had made education better.\textsuperscript{203}

At the same time, momentum was building behind some of the innovations that were inhibited by NCLB’s requirements. In addition to the testing methodologies and portfolio strategy described above, the development of rigorous national standards—a talking point since the first President Bush’s 1989 education summit—finally seemed plausible. By early 2009, the National Governors Association and the Council of Chief State School Officers had organized their state members around a project that they named the Common Core State Standards Initiative. The Initiative would develop standards for all grade levels in literacy and math aligned to eventual college- and career-readiness, as well as assessments to measure whether students were meeting those standards.\textsuperscript{204} All states other than Alaska and Texas signed memoranda of understanding committing to help develop the standards, though not necessarily to adopt them or the assessments tied to them.\textsuperscript{205}

As the head of the Chicago Public Schools, President Obama’s choice for Secretary of Education, Arne Duncan, was a key participant and advocate for those innovations. From the earliest days of the Obama presidency, he indicated a desire to overhaul NCLB, emphasizing the need to raise standards. In 2009, for example, Secretary Duncan argued that the “biggest problem” with the law is that “it doesn’t encourage high learning standards,” and “[i]n fact, it inadvertently encourages states to lower them.”\textsuperscript{206} With a polarized Congress still unable to agree on how to modify and reauthorize NCLB,\textsuperscript{207} the Administration began to use regulatory authority to reshape the federal role in education.

1. Race to the Top

The Obama Administration’s efforts to revamp federal education policy began with the ARRA.\textsuperscript{208} ARRA allocated nearly \$100 billion to the Education Department, the vast majority of which was to be allocated to states to fill

\begin{itemize}
\item 203. Frank Newport, Americans Doubt Effectiveness of “No Child Left Behind,” GALLUP (Aug. 19, 2009), http://www.gallup.com/poll/122375/americans-doubt-effectiveness-no-child-left-behind.aspx. The more familiar a respondent was with the law, the more negative the view, with 50 percent of those reporting to be very familiar with the law saying that it had made education worse. Id.
\item 205. Id.
\item 206. MANNA, supra note 168, at 144.
\end{itemize}
budgetary holes as state revenues declined during the recession. But the Act also set aside $5 billion for the Secretary of Education to award competitive grants to support educational innovation and reform. Of this amount, Secretary Duncan used $4.35 billion to create the Race to the Top grant program.

Race to the Top offered substantial grants to states on the basis of “demonstrated success in raising student achievement and . . . the best plans to accelerate their reforms in the future.” Not all states that applied were guaranteed to receive grants; only those judged by expert evaluators drawn from government, academia, and nonprofits to have the best combination of reform records and reform plans would “win” a grant.

In contrast to NCLB, whose state “plans” merely described how the state would implement the detailed requirements of NCLB, Race to the Top invited applicants to present a “systemic approach to education reform” and to “demonstrate in its application sufficient . . . participation and commitment” by local school districts and other stakeholders “to successfully implement and achieve the goals in its plans.” State applicants were required to address four broad areas “comprehensively and coherently”:

1) Adopting standards and assessments that prepare students to succeed in college and the workplace and to compete in the global economy;
2) Building data systems that measure student growth and success, and inform teachers and principals about how they can improve instruction;
3) Recruiting, developing, rewarding, and retaining effective teachers and principals, especially where they are needed most; and
4) Turning around our lowest-achieving schools.

A detailed rubric that assigned points to different elements supplemented these broad general criteria for proposals. Some of the evaluation criteria

---

211. RACE TO THE TOP EXECUTIVE SUMMARY, supra note 24, at 2; see also American Recovery and Reinvestment Act § 14006.
212. RACE TO THE TOP EXECUTIVE SUMMARY, supra note 24, at 2.
213. President Barack Obama, Remarks By the President on Education (July 24, 2009), available at http://www.whitehouse.gov/the_press_office/Remarks-by-the-President-at-the-Department-of-Education (“We will use the best evidence available to determine whether a state can meet a few key benchmarks for reform—and states that outperform the rest will be rewarded with a grant. Not every state will win and not every school district will be happy with the results.”).
214. RACE TO THE TOP EXECUTIVE SUMMARY, supra note 24, at 4.
215. Id. at 2, 4. These goals are consistent with the guidelines for competitive grants under ARRA. See American Recovery and Reinvestment Act §§ 14005(d)(2)–(5), 14006(a)(2).
were quite prescriptive. For example, states with laws or regulations that barred the linkage of student achievement data to individual teachers and students were ineligible to receive grants.\(^\text{217}\) Additionally, evaluators were to award points to states that demonstrated a commitment to state standards by joining “a consortium of States,” a clear reference to the Common Core State Standards consortia.\(^\text{218}\) And states with caps on the number of charter schools would lose points.\(^\text{219}\) Because a state’s cumulative score would determine its success, however, there was still flexibility to prioritize some criteria over others.

Forty-six states and the District of Columbia submitted Race to the Top proposals in two phases.\(^\text{220}\) After each phase, a group of finalists was selected based on the ranking of their rubric score, and representatives of those states made presentations to the reviewers.\(^\text{221}\) Of forty-one applicants and sixteen finalists in phase one, only Delaware and Tennessee received grants.\(^\text{222}\) Of thirty-six applicants and nineteen finalists in phase two, nine states and the District of Columbia obtained grants to implement their proposals.\(^\text{223}\)

In addition to the grants to states, the Education Department allocated $350 million to a competition to create next-generation assessments aligned to internationally benchmarked common standards—functionally, this meant the Common Core State Standards.\(^\text{224}\) Two organizations, the SMARTER Balanced Assessments Consortium, with twenty-eight partner states, and the Partnership for the Assessment of Readiness for College and Career, with twenty-three partner states, each received grants.\(^\text{225}\)
2. **ESEA Flexibility**

Notwithstanding the policy developments described above, Race to the Top left NCLB in place, creating several problems.\(^{226}\) Federal, state, and local officials had to divide their time and attention between multiple funding streams with different sets of rules. Moreover, Race to the Top commitments and NCLB requirements could conflict. For example, states that were attempting to reorient instruction toward the Common Core standards still had to administer and hold schools accountable to annual state tests that were not yet aligned to those standards. Additionally, the federal, state, and local bureaucratic structures built up to manage compliance with NCLB (and ESEA) remained in place, impeding a reorientation of institutional culture toward performance.\(^{227}\)

NCLB’s AYP ratchet also continued to operate, designating more and more schools for improvement status. In the 2010–11 school year, about thirty-eight thousand of the country’s approximately one hundred thousand schools failed to make AYP.\(^{228}\) Secretary Duncan called the problem “a slow-motion train wreck.”\(^{229}\) and began to negotiate with the states to allow them to refrain from raising their AYP thresholds.\(^{230}\)

In August 2011, with no prospect that Congress would soon revise NCLB, the Administration announced that the Secretary of Education would use his statutory waiver authority to offer broad but structured relief from NCLB.\(^{231}\) NCLB includes a provision that permits the Secretary of Education to “waive any statutory or regulatory requirement of this chapter for a State educational agency, local educational agency, Indian tribe, or school through a local educational agency, that... requests a waiver.”\(^{232}\) There are ten specific exclusions from the scope of the waiver authority, including civil rights provisions and funding allocation formulas.\(^{233}\) The core accountability elements of the statute are not among the exclusions, however. Indeed, the provision requires waiver applicants to describe how waiving specific provisions will “increase the quality of instruction for students” and “improve the academic achievement of students” relative to the statutory scheme.\(^{234}\)

\(^{226}\) See Manna & Ryan, supra note 216, at 528.


\(^{228}\) Dillon, supra note 207. Estimates for the number of schools that would miss AYP in 2012 were as high as eighty thousand. \(\text{Id.}\)

\(^{229}\) \(\text{Id.}\)

\(^{230}\) \(\text{Id.}\)

\(^{231}\) \(\text{Id.}\)


\(^{233}\) \(\text{Id.}\) § 9401(c).

\(^{234}\) \(\text{Id.}\) § 9401(b)(1)(B)(i)–(ii).
The ESEA Flexibility program introduced by the Administration offered states flexibility with respect to ten elements of NCLB on the condition that they devise alternative education reform plans consistent with four broad principles. The ten areas of flexibility included the 2014 deadline for all students achieving proficiency, the method of determining AYP, the implementation of school improvement interventions, the use of funds that were previously designated for specific remedial purposes, and the “highly qualified teacher” requirements. The four principles states had to address in their requests were similar to those of Race to the Top:

1) The adoption and implementation of college- and career-readiness standards and aligned assessments in literacy and math;
2) The development of a differentiated recognition, accountability, and support system;
3) The development and implementation of a system of teacher and principal evaluation that incorporates both measures of student growth and measures of professional practice; and
4) The reduction of duplicative and burdensome administration for local school districts.

First, as to the college- and career-readiness standards and assessments, many states used the Common Core State Standards and the assessments being developed by the consortia. Alaska, Texas, and Virginia, however, proposed standards and assessments that each state developed with their state higher education institutions. Second, the differentiated accountability goal asked states to use a mix of measures to define and identify schools that are performing particularly well (“reward” schools), schools with the lowest overall student performance (“priority” schools), and schools with the largest achievement gaps (“focus” schools), and to develop incentives or interventions tailored to the schools’ circumstances. States also had to explain how they would support and improve schools that did not fall within one of the three categories. Third, the teacher and principal evaluation goal aimed to replace

---

235. ESEA FLEXIBILITY, supra note 27, at 4–6.
236. Id. at 1–2.
237. Id. at 4–6.
240. ESEA FLEXIBILITY, supra note 27, at 4–5.
241. Id. at 5.
the current pro forma evaluation process with one that is more closely tied to effectiveness and supports teachers’ continuous growth as educators. 242

In addition to guidelines for the substantive content of flexibility proposals, the Education Department required states to “engage diverse stakeholders and communities” in developing their proposals and to describe in detail that engagement in their submissions. 243 Depending on the time of application, an approved proposal would be in effect through the 2013–14 or the 2014–15 school year. 244 The Department laid out major deadlines and required states to describe and commit to an implementation timeline for each component of their proposal within that window. 245

ESEA Flexibility also includes a three-part monitoring and renewal process. Part A, conducted from August to October 2012, was intended to assess each state’s “goals and approach to implementing ESEA flexibility” and ensure that the state had the “critical elements of ESEA flexibility in place to begin implementation of its plan in the 2012–2013 school year.” 246 Part B, which began in mid-2013, reviews implementation during the first year as well as corrective actions identified in Part A. Part C will monitor ongoing implementation. 247 In November 2014, the Education Department released guidance for longer-term renewal of the waivers. 248 In the meantime, the Secretary has granted one-year extensions to a number of states whose waivers otherwise would have expired. 249

Forty-six states and the District of Columbia submitted proposals for ESEA Flexibility. 250 So far, forty-three states, the District of Columbia, and Puerto Rico have been approved. 251 Two states’ proposals have been

242. Id. at 6.
243. Id. at 7.
244. Id. at 11–12.
245. Id. at 11–25.
250. The only states that have not applied are Montana, Nebraska, and North Dakota. See ESEA Flexibility Index, supra note 28.
rejected,252 one state withdrew its application,253 and two are still pending.254 The Secretary initially granted Washington State’s application for a waiver, but later revoked the waiver after the state did not meet its commitment to consider student state exam results in teacher evaluations.255 Oklahoma’s waiver request was initially granted but was revoked after the state repealed its adoption of the Common Core State Standards. The waiver was reinstated when Oklahoma passed new standards.256

C. A Break from the Past

Race to the Top and ESEA Flexibility significantly alter the roles of federal, state, and local governments in the policy-making structure of K–12 education. The existence of this shift has been widely recognized, though most commentators have focused on the controversial substantive policies—such as teacher evaluation based partly on student test results257 and participation in the

Puerto Rico, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Washington, West Virginia, and Wisconsin have approved applications. Id.


253. Vermont withdrew its proposal, with officials suggesting that they wanted more flexibility in the scheduling and use of standardized tests than the U.S. Education Department was willing to offer. See Lisa Rathke, Vermont Opt Out of No Child Left Behind Waiver, ASSOCIATED PRESS (June 3, 2012), http://www.boston.com/news/education/articles/2012/06/03/vermont_opts_out_of_no_child_opt_behind_d_waiver.

254. Iowa, Wyoming, and the Bureau of Indian Affairs have pending applications. See ESEA Flexibility Index, supra note 28.


Common Core State Standards—that the federal government has endorsed and that most states have adopted.

These substantive policies are important, but they are secondary to the structural changes that have taken place. Changes in the way that the federal, state, and local governments interact with one another have opened the door to more dynamic and targeted policy making with greater participation and collaboration by local stakeholders. Accordingly, the substantive changes are provisional—although the overarching goals, structure, and process are established by the federal policy, the determinative details are set by states in collaboration with their local stakeholders and other states, and are iteratively scrutinized and revised.

Most commentators who do discuss the structural impacts of the Obama initiatives have focused on whether the programs increase or decrease federal authority over schooling. Some argue that these new federal programs usurp state and local governments’ traditional control of education policy by imposing the Obama Administration’s education preferences on the states. Others contend that the initiatives have devolved excessive authority to the states, dangerously relaxing accountability for student learning outcomes, particularly for poor and minority children. As discussed above, legal commentators have emphasized particular aspects of the initiatives—the use of waiver, the significant state policy-making role, and experimentalist features—without recognizing how those components fit together.


261. See supra Part LB-D.

262. See Barron & Rakoff, supra note 1, at 279–80 (describing ESEA Flexibility as an example of big waiver); Richard A. Epstein, Government by Waiver: No Child Left Behind Follows the Health
In fact, the initiatives have been neither a federal power grab nor an unaccountable devolution. And they are not simply examples of big waiver, cooperative federalism, or experimentalism. In the new education federalism, federal, state, and local actors participate intensely in the development and implementation of policy. The Education Department empowers and motivates states with invitations to deviate from the default legislative scheme. Through initial conditions and monitoring, it holds states accountable to federal goals and a commitment to iteratively evaluate policy in collaboration with stakeholders. In short, the initiatives have transformed education federalism from a regime of bureaucratic governance to a regime of disciplined devolution.

Historically, the Education Department has interacted with the states through a series of command-and-control style rules regarding uses of federal funding and the features of school accountability systems. This approach in turn promoted a compliance mentality on the part of state and local governments with respect to the rules. But it also left important policy areas such as the creation of standards and assessments as well as many implementation details to the states, which allowed them to avoid or undermine the underlying goals of the federal programs.265

In the new education federalism, the Education Department establishes general goals and default policy options that are the starting point for state planning, rather than the final, specific process rules of the past. After states make their proposals, the Department confirms that the process by which the states devised their plans included meaningful collaboration with local stakeholders and enforces that expectation by delaying or denying approval if engagement is insufficient.266 It then evaluates the state plans, pushing for justification and detail on the policies selected and credible expectations for implementation and outcomes. If the states’ plans or explanations are unsatisfactory, the Department requests revision or denies an application outright.267

Once the state plans are approved, the Department monitors implementation. Unlike the process-focused monitoring of the old regime, monitoring under the new education federalism focuses on whether the results of the states’ initiatives advance the goals that were defined at the outset by the Department and refined in the state plans. For example, if a state has deviated

---

263. See Metzger, supra note 6, at 590 (describing Race to the Top as an example of progressive federalism).
264. See Sabel & Simon, supra note 7, at 56, 85 (describing Race to the Top as an example of experimentalism).
265. See supra Part II.A.
266. See infra Part III.A.2.
267. See supra notes 252–54 and accompanying text.
from its original plan, the Department does not automatically view that as noncompliance, but instead seeks to understand the rationale for the change in light of the state’s initial experience implementing its plan and whether the revised plan is consistent with the Department’s overarching goals. Additionally, to enable comparisons of state policies and outcomes, the Department compiles and publishes the information it gains from the monitoring process.

The new education federalism invites states to develop comprehensive plans for education governance. Under ESEA and NCLB, state education departments played the role of bureaucratic middle men: they were responsible for demonstrating to the Department of Education that they were in compliance with federal rules, which they accomplished by shifting the burden of demonstrating compliance to local school districts. Now, state education departments are the designers of their own education policy subject to the general goals and continuing results-oriented scrutiny of the federal government. Further, states must collaborate with local school districts and other education stakeholders in designing education plans. The states are no longer accountable for compliance with specific federal rules, but rather for the development and success of their own plans in achieving broad goals.

The states are also accountable to one another in a way they previously were not. The Common Core State Standards Initiative, reinforced by state commitments in ESEA Flexibility and by the Race to the Top grants, has introduced a framework for interstate collaboration as one of the central components of the new initiatives.

Finally, the new education federalism affects the relationship between state and local actors. While in the past local school districts were the objects of state and federal initiatives, they are now collaborators in the development of state plans. Further, the differentiated accountability systems under ESEA Flexibility allow states to focus interventions on a smaller number of districts and schools and give more flexibility in the type of intervention than the uniform interventions prescribed by NCLB.

III. ASSESSING DISCIPLINED DEVOLUTION IN ACTION

The Obama education initiatives are still in the early stages of implementation. It was not until June 2012 that ESEA Flexibility proposals

268. See, e.g., ESEA FLEXIBILITY PART A MONITORING PROTOCOL, supra note 246; ESEA FLEXIBILITY PART B MONITORING PLAN, supra note 247.

269. See, e.g., ESEA Flexibility Index, supra note 28 (containing links to webpages with downloadable monitoring reports for each state).

270. See supra note 240 and accompanying text.
were approved for a majority of the states. The proposals for change were ambitious, often requiring changes in state law and bureaucratic organization, addition of new state and local employees, extensive training, data systems upgrades, and other scale-ups. While some states have begun to orient their own standardized tests to the Common Core Standards, the two Common Core Standards assessment consortia will not produce the first set of tests for use by their member states until the 2014–15 school year.

States and districts that were early adopters of substantive and process elements of the new education federalism—such as higher standards, combining school autonomy with more precise accountability systems, and structured diagnostic observations—have shown more improvement than other jurisdictions across various measures of student outcomes, including NAEP results, graduation rates, and measures of college readiness. For now, however, it is too early to assess whether the new education federalism as a governance system has produced and is likely to continue to produce better outcomes for students. Nonetheless, the early stages of implementation have generated some evidence of the ways in which state and federal governments are operating under the new regime, which can be used to test the theoretical benefits of disciplined devolution.


A. Promising Signs

On the whole, it appears that many of the abstract benefits of disciplined devolution have been realized in the new education federalism. First, states have been highly motivated to participate in the programs and in many cases have done so in the face of opposition from those with a stake in maintaining the status quo. Second, the initial planning brought together intrastate stakeholders with a diverse array of views about the substantive policies under consideration, including representatives of the local school districts, principals, and teachers who would be expected to implement the plan, and the students and parents who would be affected by the educational changes. Third, the model appears to have produced dynamic policy making in the states, with tailoring to local conditions, experimentation, and adaptation to changing circumstances. Finally, the federal monitoring process and the interstate Common Core consortia have kept the states on course with most of their waiver commitments.

1. Motivation

Critics of experimentalism worry that local actors will not be motivated to take the continuous learning processes seriously and that it will be difficult to build consensus around end goals. The federal role in disciplined devolution—offering states opportunities to deviate from a uniform legislative scheme, setting goals as conditions for doing so, and reinforcing a collaborative process—is a putative answer to this critique. It appears that this theoretical benefit is borne out in the new education federalism. Not only did nearly every state take the federal government up on its offer to deviate from the NCLB framework and enter the new regime, state governments and local stakeholders have expended significant effort in doing so, including by making legislative changes and reorganizing state institutions to better align to the federal goals. Moreover, they did so despite opposition from politically powerful teachers’ unions.

With forty-one states submitting Race to the Top applications and forty-six submitting ESEA Flexibility proposals, it is clear that the incentives of additional funding and relief from NCLB were effective at motivating states to opt into the new regime. Of course, the mere fact that states opted in does not necessarily mean that they participated sincerely. Indeed, some have contended that participation in Race to the Top was superficial, with states merely checking the Education Department’s boxes in order to obtain the federal funds. The relatively short period during which states could apply (just three
months for the first phase and another six months for the second phase)\textsuperscript{276} adds some credence to this critique.

Weighing against this critique of superficiality are the significant legislative changes that many states undertook to improve their application for Race to the Top, as well as the time and effort put into preparing the applications. Twenty-eight of the forty-one states that submitted applications in phase one of Race to the Top passed laws related to teacher evaluation, certification, or professional development either just before or just after the Race to the Top deadline.\textsuperscript{277} A number of states also passed laws either raising or eliminating caps on the number of charter schools that could be opened in the states.\textsuperscript{278} Furthermore, state officials reported spending “thousands of hours” preparing Race to the Top applications, an unlikely investment of effort if the goal were merely to “check boxes.”\textsuperscript{279} Some state officials have cited Race to the Top as either political cover for pursuing potentially controversial reforms or accelerating reforms that were underway—in other words, for those officials, Race to the Top facilitated meaningful reform, even if it did not prompt it.

The superficiality claim has little purchase on ESEA Flexibility participation. With a longer period to prepare applications and a broader scope, each state produced a unique and detailed proposal that typically ran to hundreds of pages.\textsuperscript{281} In addition to the substantive policies, the proposals included descriptions of the consultation that states had done with local school districts, teachers, principals, unions, and other stakeholders, as well as background on reforms the states had previously undertaken and how they primed the state for success.

Further bolstering the case that participation in ESEA Flexibility was sincere, several states used their ESEA Flexibility proposals as an impetus for changing the structure of their education institutions. For example, in contrast

\textsuperscript{276} The Race to the Top rules were published in the Federal Register on November 18, 2009, the phase one applications were due January 19, 2010, and the phase two applications were due June 1, 2010. U.S. DEP’T OF EDUC., RACE TO THE TOP PROGRAM: GUIDANCE AND FREQUENTLY ASKED QUESTIONS 3, 7 (updated May 27, 2010), available at http://www2.ed.gov/programs/racetotop/faq.pdf.

\textsuperscript{277} LEARNING POINT ASSOC’S., STATE LEGISLATION: EMERGING TRENDS REFLECTED IN STATE PHASE 1 APPLICATIONS 2–3 (June 2010), available at http://www.learningpt.org/pdfs/RttT_State_Legislation.pdf.

\textsuperscript{278} Race to Top Applications In, Impact Already Felt, ELECTRONIC EDUC. REP., Jan. 25, 2010, at 4–5.


\textsuperscript{280} Manna, supra note 227, at 6.

to its prior, programmatic structure, Connecticut reorganized its state department of education around strategic priorities, with a cabinet-level official responsible for each priority. It also created a “Commissioner’s Network” to provide support to the twenty-five lowest performing schools in the state. New Jersey also moved away from an organizational structure oriented around individual programs to a system of seven field offices called Regional Achievement Centers, which will support struggling schools and districts with curriculum and instructional practice, teacher recruitment and development, and data analysis, among other things.

Additionally, schools and districts implemented these changes notwithstanding skepticism on the part of politically powerful teachers’ unions. The unions have long opposed the proliferation of charter schools and teacher evaluation systems that incorporate student academic results, two of the most prominent policies encouraged by the Obama Administration. Both of the major national unions have criticized aspects of Race to the Top and ESEA Flexibility. State and local unions continue to criticize and in some cases stymie reform efforts undertaken pursuant to the state plans. And yet, as described below, states were able to convince many local unions not only to formally sign on to the proposals, but also to actively engage in the design of policies they opposed. Negotiation around the substantive policies, which locked in some aspects of the reforms and left others open ended, and the value

283. See id. at 27.
286. See, e.g., Stephen Sawchuk, NEA’s Delegates Vote ‘No Confidence’ in Race to the Top, EDUC. WK. TEACHER BEAT BLOG (July 4, 2010, 12:38 PM), http://blogs.edweek.org/edweek/teacherbeat/2010/07/neas_delegates_vote_no_confide_2.html; Sam Petulla, None of the Above, AM. PROSPECT (June 20, 2011), http://prospect.org/article/none-above (“The NEA and the American Federation of Teachers (the second-largest teachers’ union) came down hard [in response to Secretary of Education Arne Duncan’s announcement that he would offer broad waivers from NCLB requirements], with the NEA saying that the reforms attached to the waivers would create ‘more unmanageable hurdles for schools and students.’ The AFT expressed unhappiness with what it considers a half-measure, saying that only Congressional authorization can accomplish real reform.”).
288. See infra notes 286–87 and accompanying text.
of the flexibility and funding offered by the Obama Administration’s initiatives likely contributed to this.\textsuperscript{289}

2. \textit{Legitimacy and Transparency}

Waiver authority threatens democratic legitimacy and transparency by giving unelected bureaucrats discretion to cancel the law selectively. Disciplined devolution mitigates this concern by shifting the policy-making process to politically accountable state governments and requiring the states to demonstrate collaboration with the affected parties in designing their plans. The initial planning stages of Race to the Top and ESEA Flexibility provide evidence that the disciplined devolution model can foster meaningful stakeholder engagement and transparency.

Race to the Top provided a focal point for planning within states and an incentive for sometimes-fractious groups within states to come together to consider state education policy in a comprehensive way. Endorsement by local school districts, teachers’ and principals’ unions, businesses, and community leaders factored in to about seventy points in the five-hundred-point rubric, in categories such as whether the state had secured local school district commitment, whether that commitment would translate to statewide impact, and whether there was broad stakeholder support.\textsuperscript{290} By phase two of the competition, states had, on average, secured endorsement from 61 percent of local school districts and 68 percent of local unions.\textsuperscript{291} And in states where legislation was passed in anticipation of Race to the Top, the ordinary dynamics of legislative politics likely included input from interested (and politically capable) stakeholders.

Experience with Race to the Top, a longer period to formulate proposals, and the requirement of describing the consultation process in detail appears to have produced robust intrastate collaboration in the ESEA Flexibility process too. New York’s experience provides an illustrative example. From August 2011 until it submitted its proposal in February 2012, the New York State Education Department convened monthly meetings with interested stakeholders such as local school district representatives; union representatives; student and civil rights advocacy groups; parent, school board, and superintendent associations; and local and national education researchers.\textsuperscript{292} During these meetings, the group—known as the ESEA Waiver Think Tank—established goals for the waiver plan, discussed each element of the proposal,

\textsuperscript{289} \textit{See}, e.g., Michele McNeil, \textit{Race to Top Buy-In Level Examined}, \textit{EDUC. WK.}, June 16, 2010, at 1 (describing various strategies states used to increase union endorsement of their Race to the Top applications).

\textsuperscript{290} Michele McNeil, \textit{Applicants for Race to Top Puzzle over ‘Buy-In,’} \textit{EDUC. WK.}, May 12, 2010, at 25.

\textsuperscript{291} McNeil, \textit{supra} note 289, at 1.

2015] DISCIPLINED DEVOLUTION 617

and edited drafts of proposal documents.\textsuperscript{293} Once the proposal was drafted, the State Education Department opened a ten-day public comment period, hosted multiple public meetings to describe and solicit questions and feedback, and hosted meetings between the state education commissioner and the state’s major education groups.\textsuperscript{294} The final version of the flexibility proposal summarized and responded to this feedback and reflected substantive changes influenced by it.\textsuperscript{295}

The federal Education Department reinforced the engagement norm during the application review for both Race to the Top and ESEA Flexibility. After phase one of Race to the Top, reviewers and the Education Department focused attention on the level of “buy-in” achieved by the winners.\textsuperscript{296} As a result, states intensified their negotiations with unions, and union endorsement increased from a state average of 46 percent in phase one to an average of 68 percent in phase two.\textsuperscript{297} Similarly, in initial feedback on the second round of ESEA Flexibility proposals, nearly every state was encouraged to do more to solicit the views of stakeholders on their plans.\textsuperscript{298} In some cases, reviewers and the Department told states that they had to describe their engagement with more specificity or undertake additional outreach before their plans could be approved.\textsuperscript{299}

3. Dynamic Policy Making

By incorporating state differentiation and experimentalist processes, disciplined devolution should yield policy making that is more dynamic than the mandate-based compliance model of a regime like NCLB or ESEA. Policies should reflect local conditions and capacities, test innovative ideas, and evolve with experience. The early implementation of Race to the Top and ESEA Flexibility demonstrates this tailoring, experimentation, and adaptation.

Race to the Top financed dozens of small-scale and experimental reform initiatives in the states that won grants. Georgia has used its funds to launch a series of competitions for schools and districts to test new approaches in areas such as teacher recruitment and the cultivation of students’ problem-solving skills.\textsuperscript{300} A number of Florida districts are using Race to the Top funds to implement and evaluate “lesson study,” a practice originating in Japan in which

\begin{itemize}
\item \textsuperscript{293} Id.
\item \textsuperscript{294} Id. at 14–16.
\item \textsuperscript{295} Id. at 12–13, 16–20.
\item \textsuperscript{296} McNeil, \textit{supra} note 289, at 1.
\item \textsuperscript{297} Id.
\item \textsuperscript{299} Id.
\item \textsuperscript{300} Sean Cavanagh, \textit{Race to Top Funds Nurturing Grassroots Efforts}, EDUC. WK., Apr. 20, 2011, at 26.
\end{itemize}
teachers co-plan lessons before class and then debrief afterwards. Maryland is developing a new training and certification program for elementary teachers in science, technology, engineering, and math.

Within a year of winning Race to the Top, six states and the District of Columbia received approval from the Department of Education to adjust their proposals’ timelines and scope of work. As these states moved to implementation, capacity or policy details that seemed appropriate in the abstract proved not to be so in reality. For example, Massachusetts delayed a plan to adopt regulations for a tiered-principal-licensure and career-ladder system for a year as it began to work out the details of the plan. North Carolina changed a plan to provide retention bonuses to every teacher in low-performing schools into a pilot in which only a limited number of teachers were eligible. In Hawaii, negotiations with the state teachers’ union over details of a comprehensive evaluation plan broke down, and the state turned to smaller-scale pilot efforts to implement an evaluation system. Delaware, one of the original winners, was approved to amend its plan by delaying consequences attached to a student growth measure, giving the state more time to develop measures for grades and subjects without standardized tests.

Some commentators have cast these adjustments as failures by either states or the Department of Education. Yet resetting timeframes and undertaking pilots before fully implementing new programs may also be understood as adapting plans to the experiences of implementation. Indeed, even Hawaii, which the Education Department labeled at “high risk” of losing

301. Id.
302. Id.
304. Id.
306. Cavanagh, supra note 257.
308. See Ulrich Boser, CTR. FOR AM. PROGRESS, RACE TO THE TOP: WHAT HAVE WE LEARNED FROM THE STATES SO FAR?: A STATE-BY-STATE EVALUATION OF RACE TO THE TOP PERFORMANCE 12 (Mar. 2012) (“[I]n many ways, some changes and delays are not surprising. Few social programs—whether it’s education or health care—are implemented in exactly the way that they were conceived and it can take time to build capacity. Moreover, one of the goals of Race to the Top was to figure out what works when it comes to education reform, thus some mid-course corrections are to be expected.”).
grant funding in 2011,\textsuperscript{309} is now on track to be the first state to meet all of its Race to the Top commitments.\textsuperscript{310}

Dynamic policy making under ESEA Flexibility is evident in both the states’ detailed proposals and the early efforts to put the policies in place. Although state plans were oriented toward federally defined reform principles, their details differ significantly. For example, while nearly all of the approved plans embrace the Common Core State Standards, the Department approved Virginia’s plan to set its own college- and career-ready standards and create assessments working with the state university system.\textsuperscript{311} There is also diversity in states’ performance goals. Only nine states selected one of the Education Department’s two default goal-setting methodologies: cutting the percentage of students below proficiency in half by 2016 or achieving 100 percent proficiency by 2020.\textsuperscript{312} Every other state chose to define achievement goals according to a formula of its own devising.\textsuperscript{313}

Each state also has a unique methodology for evaluating schools. While every state includes high school graduation rates and performance on standardized tests in literacy and math in the evaluation, some also include tests in other subjects such as science, social studies, or the English language proficiency of English language learners.\textsuperscript{314} A number of states weigh graduation outcomes by the difficulty level of the diploma received, and some include measures such as ACT or Advanced Placement test results to gauge college readiness.\textsuperscript{315} Still other states include the results of student surveys about the school learning environment.\textsuperscript{316} States use different methodologies for calculating students’ longitudinal learning growth.\textsuperscript{317} And each state combines all of these elements in different ways.\textsuperscript{318}

\textsuperscript{309} Michele McNeil, Race to Top Winners Plug Away at Promises, EDUC. WK., Sept. 19, 2012, at 23; see McNeil, supra note 305, at 19.
\textsuperscript{310} Michele McNeil, In the Race to the Top, Will Hawaii Get There First?, EDUC. WK. POLLS. K–12 BLOG (June 28, 2013, 11:41 AM), http://blogs.edweek.org/edweek/campaign-k-12/2013/06/in_the_race_to_the_top_will_Ha.html.
\textsuperscript{312} JEREMY AYERS & ISABEL OWEN, CTR. FOR AM. PROGRESS, NO CHILD LEFT BEHIND WAIVERS: PROMISING IDEAS FROM THE SECOND ROUND OF APPLICATIONS 22 (July 2012).
\textsuperscript{313} Id. at 23.
\textsuperscript{314} Id. at 22.
\textsuperscript{315} See, e.g., LA. DEP’T OF EDUC., LOUISIANA’S ESEA FLEXIBILITY REQUEST 52–56 (Apr. 5, 2013), available at https://www2.ed.gov/policy/eseaflex/approved-requests/larquestamended042313.pdf (proposing high school evaluation methodology that reflects ACT results and weighted graduation index, including additional points for success on Advanced Placement exams).
\textsuperscript{317} Compare, e.g., COLO. DEP’T OF EDUC., COLORADO ESEA FLEXIBILITY REQUEST 53–54 (Nov. 28, 2012), available at https://www2.ed.gov/policy/eseaflex/approved-requests/co_amend121912.pdf (describing percentile growth and growth-to-standard methodology),
Several states have embarked on efforts to develop new tools or approaches to school evaluation and intervention. For example, New York created a rubric for observational diagnostic reviews of schools’ organization and capacity, contributing qualitative insights to help interpret and react to quantitative findings.\textsuperscript{319} And Tennessee created a competitive grant program for schools to design, implement, and evaluate interventions that address achievement gaps within schools and disseminate best practices.\textsuperscript{320}

And some states have already made adjustments to policies described in their ESEA Flexibility proposals as implementation revealed consequences inconsistent with goals. For example, Virginia changed its approach to setting improvement targets for schools after tests of its originally proposed methodology revealed that success meeting the initial targets would “not result in closing achievement gaps between student subgroups.”\textsuperscript{321} Similarly, Colorado is in the process of rethinking its monitoring of local school districts to “go beyond compliance and look at program quality.”\textsuperscript{322}

The Education Department has encouraged states to make these substantive adjustments. In its guidance on waiver renewal, the Department notes that renewal “provides an opportunity for [a state education agency] to amend its approved request as part of a continuous improvement process to address challenges and build on successes from current implementation.”\textsuperscript{323} As part of their renewal process, Connecticut and several other states have begun to consider more substantial changes to their accountability systems, including

\begin{itemize}
  \item \textit{with Conn. St. Dep’t of Educ., supra} note 281, at 90–92 (describing vertical scale growth methodology).
  \item \textit{Compare, e.g., La. Dep’t of Educ., supra} note 315, at 50–51, 53 (describing 150-point School Performance Score index and criterion-referenced conversion to A–F letter grades), \textit{with S.D. Dep’t of Educ., ESEA Flexibility Request} 33–34, 45 (June 25, 2012), available at https://www2.ed.gov/policy/eseaflex/approved-requests/sd.pdf (describing 100-point School Performance Index and identification of top 10 percent, bottom 5 percent, and the 10 percent with the greatest achievement gaps).
  \item \textit{U.S. Dep’t of Educ., supra} note 248, at 3.
\end{itemize}
adding measures to cover additional subjects and give more weight to qualitative measures.  

4. Interstate and Federal Accountability

Relative to other federal programs that rely on state policy making to carry out federal objectives, disciplined devolution should lead to greater accountability for states to adhere to their commitments. The early implementation of the new education federalism suggests that the accountability of the disciplined devolution model has bite.

After receiving substantial prodding by the Education Department, Race to the Top winners are mostly on track with their commitments. The Department publicly released reports evaluating each state’s progress in fulfilling its commitments one year and two years into the grant period. The reports candidly assessed whether states had hit promised milestones and the likelihood of their hitting future milestones. While the Department found most states to be on track with their plans, it reprimanded New York, Florida, and Hawaii for missing significant deadlines in the first year of evaluations. By the time the second-year reports were released, those three states had improved their standing, while the District of Columbia, Georgia, and Maryland were called out for delays or shortcomings in implementation. Each time, the Education Department warned the lagging states that they risked losing funding if they did not get back on track. The improvement of the lagging states from the first round of reports, and a flurry of activity by the states singled out in the second round of reports, suggest that this public shaming had the desired effect of spurring the states onward.

Similarly, the federal monitoring for ESEA Flexibility has been helpful in keeping states on track with their commitments. Like the Race to the Top evaluations, the Part A monitoring reports are published on the Education Department’s website and include critical assessments of each state’s early

---

326. See Armario, supra note 325.
327. See McNeil, supra note 325.
328. See Armario, supra note 325; U.S. DEP’T OF EDUC., supra note 246; McNeil, supra note 325.
329. See McNeil, supra note 325.
330. See ESEA Flexibility Index, supra note 28.
implementation of its ESEA Flexibility plan. For several states that have missed early implementation deadlines, the Department has released public extension letters noting that the states were at “high risk” of losing their waivers and federal administrative funding. 331 As noted above, the Department has revoked the waiver of one of those states, Washington, after it missed a final deadline for amending state law to require the consideration of students’ state test results in teacher evaluations. 332 It revoked Oklahoma’s waiver after the state dropped the Common Core State Standards and then reinstated the waiver after the state adopted new college and career readiness standards. 333

In addition to federal monitoring, the interstate consortia developing the Common Core Standards and assessments have largely kept the project on track, despite rumblings of discontent from some conservatives and some in the education community. Conservative politicians and activists have increasingly attacked the Common Core Standards, citing concerns that they represent a federal incursion on state authority, will end up nationalizing curriculum, and will be costly. 334 These conservative critics have found unusual common ground with a number of liberal education commentators who worry about standardization of curriculum and more testing. 335 Teachers’ unions, initially supportive of the standards, have begun to object to the new tests as the prospect of their inclusion in teacher evaluation looms. 336 Three states have dropped the standards entirely, and others have initiated processes to consider revisions. 337 Several states have kept the standards but withdrawn from the Common Core testing consortia. 338

332. See supra note 255 and accompanying text.
333. See supra note 256 and accompanying text.
Yet for the most part, the participating states have hung together and the implementation of the Common Core Standards is moving forward. The standards for math and literacy and supporting policies have been released,

and the tests are on track to be administered at the end of the 2014–15 school year. In the meantime, several states have revised their own state tests to align to the Common Core Standards, even though doing so has led to politically damaging declines in student results. Participating states and school districts are also in the midst of training teachers in the new standards, and using the standards as the basis for the observational and professional development aspects of their new teacher evaluation systems. Even the states that have dropped the standards or dropped out of the testing consortia are still required by their waivers to develop college- and career-readiness standards and tests aligned to those standards. The states that develop their own standards and tests are likely to borrow heavily from the Common Core.

The collaborative nature of the Common Core consortia’s work, the organizational structures that have built up around it, and “safety in numbers” all likely contribute to the resilience of the project. Representatives of states in the same consortia regularly meet and hold conference calls for work groups on


implementation and governance issues. The consortia receive funding from Race to the Top and private grants, and they have contracted for professional management. The Council of Chief State School Officers and the National Governors Association continue to provide support for the work and forums for additional interaction among state representatives. And despite a few defections, participating states can still say they have signed on to standards that nearly every state agrees are rigorous and necessary, framing potential withdrawal from the standards as abandoning high expectations.

B. Shortcomings

There are several important ways in which the reality of the new education federalism has not lived up to the theoretical advantages of disciplined devolution. The dynamic policy making has been largely siloed within states because the model has not yet facilitated the spread of best practices. Furthermore, the engagement with intrastate stakeholders evident during the drafting stage has not continued through to the early implementation stages in many states. This decline in collaboration poses a risk to the ongoing legitimacy of the policy-making process and could also affect its ultimate success. At this stage, however, neither of these shortcomings appears to be locked in, and adjustments at the federal level should be able to remedy them.

1. Siloed Policy Making

As discussed above, the disciplined devolution structure of the new education federalism appears to promote state policy experimentation, responsiveness to implementation signals, and self-correction in a continuous process of learning and adaptation. But while the federal monitoring and interstate collaboration predict that the dissemination of policy ideas should also flow across state lines, in practice, the process has largely been confined within states. Best practices have simply not spread as quickly or as widely as anticipated.

The fairly technical area of evaluation methodologies provides a stark illustration of the effects of siloed policy making. For example, year-over-year changes in proficiency rates have been shown to be unreliable indicators of school improvement. Relatively easy-to-implement alternative methodologies, such as standardized growth measures, have been used by a

345. See McDermott, supra note 204.
347. See Kane & Staiger, supra note 192, at 61–65.
number of jurisdictions for several years and were adopted explicitly because of concerns about the change in proficiency methodology. Yet many states continue to make proficiency rates the centerpiece of their school evaluation systems under ESEA Flexibility, while less than half of states approved for waivers incorporate standardized measures of student growth.

One reason for the limited spread of policy lessons across states is the Department of Education’s failure to emphasize or facilitate the sharing of best practices among states or to exercise its own judgment in benchmarking the quality of different state approaches. Neither the Race to the Top nor the ESEA Flexibility monitoring process includes review of whether and how a state has incorporated information from other states or shared its own lessons learned with other states. Furthermore, the Department’s efforts to convene states around particular policy problems have been inconsistent. For example, a June 2011 Government Accountability Office Report found that while the Education Department occasionally convened Race to the Top grant recipients electronically or in person to discuss common challenges and share materials, states that did not win grants did not have access to those meetings or materials. While the Department’s publication of its monitoring reports and evaluations on its website does make accessible a set of resources for the states, it is a passive step: there is no systematic mechanism for providing states with relevant information from other states’ reports.

2. Decline in Collaboration

Participation by a broad range of local stakeholders in the initial formulation of policy under disciplined devolution may mitigate concerns about the legitimacy of deviation from a congressional scheme. By the nature of the model, however, the substantive policies are likely to change frequently with the experience of implementation. Ongoing policy reformulation requires ongoing stakeholder participation to retain its legitimacy. Unfortunately, the stakeholder participation in formulating Race to the Top and ESEA Flexibility plans in many states has not always persisted into the implementation stage.

Some states have simply failed to keep local districts, unions, and advocates party to—and, in some cases, even informed of—implementation of Race to the Top and ESEA Flexibility plans. A 2012 review of Race to the Top implementation, for example, found that only five of the twelve Race to the Top winners made public the results of their monthly status calls with the


Education Department. See BOSER, supra note 308, at 11.

352. See id.

353. Tennessee’s grant program for local innovations in school intervention is one example. See supra note 320 and accompanying text. New York has also developed some experimentalist features at the state level. It has vested discretion in local school districts by identifying “Focus Districts” with the largest achievement gaps rather than individual “Focus Schools,” see N.Y. STATE EDUC. DEP’T, supra note 292, at 104–08; fostered an iterative learning process in the districts and at the state level through its diagnostic reviews, see supra note 319 and accompanying text; and has incorporated the experience of districts into state default policy, which it has imposed as needed. For example, when New York City and its teachers’ union could not reach an agreement with regard to a teacher evaluation system, the state’s education commissioner imposed one. See Channing Joseph, New York to Evaluate Teachers with New System, N.Y. TIMES, June 1, 2013, http://www.nytimes.com/2013/06/02/nyregion/new-evaluation-system-for-new-york-teachers.html. In that system, among other things, the commissioner created a scoring algorithm different from that required by state law after New York City and other districts pointed out design flaws that came to light as they prepared for implementation. See Geoff Decker, N.Y.C.’s Evals Include Scoring Fix That District’s Lacked This Year, CHALKBEAT N.Y. (June 3, 2013, 10:15 AM), http://gothamschools.org/2013/07/03/nycs-evals-include-scoring-fix-that-districts-lacked-this-year.


355. In Maryland, for example, the state teachers’ union and several school districts sued to enjoin implementation of the state’s teacher evaluation system, claiming that the state usurped the districts’ authority to decide, with the unions, what should factor into the evaluations. See Andrew Ujifusa, Teacher-Evaluation Troubles Sprout in Maryland; Court Spat Looms, EDUC. WK. ST. EDWATCH BLOG (June 10, 2013, 10:34 AM), http://blogs.edweek.org/edweek/state_edwatch/2013/06/teacher_evaluation_troubles_sprout_in_maryland_and_court_spat_looms.html.
3. Decentralization, Centralization, and Federal Solutions

Where the new education federalism falls short of the disciplined devolution model, an excess of one of the components of the model is to blame. Siloed policy making is a consequence of decentralized policy making. In disciplined devolution, federal monitoring is meant to provide a counterweight by promoting information sharing and enforcing a state process for integrating it. But the federal role in the new education federalism has been too weak in this regard. Tapering of stakeholder collaboration, however, is a consequence of too much centralization. Even if policy-making authority shifts from the federal government to state governments, if local governments and nongovernmental stakeholders are not allowed to take some ownership of the policy-making process, tailored and adaptive experimentation will be limited. In disciplined devolution, it is the federal government’s role to reinforce appropriate decentralization as well. Both shortcomings—siloed policy making and decreased stakeholder collaboration—may therefore be remedied through adjustments at the federal level.

To promote the sharing of lessons across states, the Education Department could enhance the monitoring process in two ways: by better deploying its own expertise and through state self-assessment. Under the first approach, the Department would aggregate, synthesize, and distill the information gathered during the monitoring into a set of best practices that are published and distributed to the states. Going a step further, during monitoring the Department could offer information on effective practices in other states that address challenges similar to those identified by the state under review and require the state to incorporate and adapt those policies or explain why it does not do so.

The second approach would facilitate states learning from one another directly. As part of the monitoring process, the Department could require states to compare their policies and implementation processes to those of other, similarly situated states and explain why the options they have chosen are preferable. To assist with the interstate comparisons, the Department could organize a peer review process that convenes multiple state officials to undertake or participate in the monitoring of another state.

With regard to flagging intrastate stakeholder participation, making ongoing stakeholder engagement a component of the monitoring protocol should be an effective strategy. The requirement to demonstrate stakeholder engagement at the application stage of Race to the Top and ESEA Flexibility prompted states to undertake significant efforts to inform and solicit input from local officials, unions, and advocates; enforcing such a requirement during monitoring should do the same. The Department of Education may already be moving in this direction: the protocol for Part B of ESEA Flexibility monitoring includes several questions seeking detailed information about
efforts to inform and engage teachers, parents, local districts, and other “diverse stakeholders.” 356

Another potential solution to flagging intrastate collaboration is the Education Department’s direct engagement with school districts. The Department has awarded one round of Race to the Top grants to districts and is preparing a second round. 357 It also granted ESEA Flexibility waivers to eight California school districts after California’s state application was rejected. 358 While there are risks to “dissecting the state” in this way, these efforts at district engagement could be directed toward reinforcing a degree of district autonomy and experimentation in states that have been more prescriptive toward their districts. 359

IV.
A FRAGILE BALANCE

The new education federalism demonstrates the potential of disciplined devolution. It has generated sincere and intensive participation by states, fostered collaboration among intrastate stakeholders (at least at the initial policy formulation stage), yielded policy diversity and experimentation (at least within states), and demonstrated strong accountability mechanisms. It is a promising model for education governance and supports the promise of disciplined devolution as a model in other policy domains.

Yet achieving the full potential of disciplined devolution may prove elusive. The ways in which the education example fell short of disciplined devolution’s predicted benefits reflect a precarious balance of centralizing and decentralizing forces. While the internal pressure of these competing forces has disrupted implementation, exogenous forces pose a potential systemic risk. Political change, premature legislative lock-in of substantive policy changes, a decrease in the federal government’s leverage, or a failure to recognize the nature and benefits of the new structure could each tip the model into a mandate-based, top-down governance structure or leave policy making to states without adequate oversight.

The political risk to disciplined devolution takes three forms. First, a change in administration could lead to a change in the agency’s approach to policy or interstate relations. Second, a continuing administration may change its approach as the political calculus changes. Finally, changes in the political climate in states may disrupt the model. The potential for each of these political

356. See ESEA FLEXIBILITY PART B MONITORING PLAN, supra note 247, at 17–18.
changes exists in the education context, as national Republican politicians have positioned themselves against the Obama Administration’s policies,\textsuperscript{360} teachers’ unions and other Democrat-leaning advocates continue to pressure the Administration to roll back elements of the initiatives,\textsuperscript{361} and aspects of the reforms remain controversial in the states.\textsuperscript{362}

A second risk for disciplined devolution is federal legislation that prematurely locks in substantive policies without preserving the governance structure. Success of some policies in some contexts may convince Congress to require those policies in all contexts. Legislative incorporation of the lessons learned through policy experimentation is helpful to disciplined devolution, but only if flexibility to deviate is retained. Such flexibility is necessary to motivate critical policy analysis on the part of state and local actors and to permit continued experimentation and adaptation. Legislators are liable to mistake success of a particular policy for success of the policy-making process and act on that mistake to cripple the policy-making process. There is some evidence of this risk in existing bills to reauthorize ESEA.\textsuperscript{363}


\textsuperscript{361} In 2013, Secretary Duncan issued guidance that states may apply to delay the inclusion of student test results in teacher evaluations beyond the deadline in their ESEA Flexibility plans, a move that is directly responsive to calls by the teachers’ unions for a “moratorium” on attaching stakes to test results. See Alyson Klein, \textit{Arne Duncan Allows Waiver States Extra Time on Teacher Evaluation}, \textsl{Educ. Wk. Polis. K–12 Blog} (June 18, 2013, 2:30 PM), http://blogs.edweek.org/edweek/campaign-k-12/2013/06/arne_duncan_allows_waiver_stat.html. More recently, Duncan granted all waiver states a yearlong reprieve from deadlines for incorporating student test results into evaluations, announcing the change in a speech that decried the negative impact of standardized tests on instruction and school climate. See Catherine Gewertz, \textit{U.S. Ed. Sec. Duncan: Too Much Testing Costs Teachers and Students ‘Precious Time’}, \textsl{Educ. Wk. Curriculum Matters Blog} (Aug. 21, 2014, 4:46 PM), http://blogs.edweek.org/edweek/curriculum/2014/08/us_ed_sec_duncan_slams_excessi.html.


\textsuperscript{363} For example, the Senate Democrat’s ESEA reauthorization bill requires that states identify a percentage of schools for intervention according to a specific accountability formula, issue school reports that include specific information, and create teacher evaluation systems that incorporate student test results. While these requirements are similar to some of the substantive expectations of ESEA Flexibility, they are more detailed and prescriptive, and they will not be the subject of state-by-state
Although premature legislative incorporation of substantive policy runs a risk of shifting the balance to uniform centralization, excessively delaying an update to default legislation may also undermine the framework. It is critical for disciplined devolution that the default statutory scheme remains credible. If it is not, states will have no reason to take it seriously, and the federal agency’s leverage will be diminished. The risk of legislative obsolescence is illustrated by NCLB’s requirement that all students achieve grade-level proficiency on statewide tests by the end of the 2014 school year.\(^{364}\) The 2014 deadline, and its related funding penalties, was a major motivator for states to participate in ESEA Flexibility. Now that the deadline has passed, states with waivers may assume that the provision is moot, or that the Education Department’s threats to withdraw funding are not credible, and cease to take their commitments under ESEA Flexibility seriously. The NCLB default has also been weakened by the Education Department’s willingness to negotiate alternative arrangements with the handful of states that did not apply for or were denied a waiver.\(^{365}\)

A final, pervasive risk is that the relevant federal and state actors fail to recognize the nature of the new governance structure and therefore unwittingly take steps that weaken it. In other words, disciplined devolution might disappear without anyone realizing it existed in the first place. There is a high risk of such a category error in the education case study. As discussed above, most education commentators have not focused on the structural changes of Race to the Top and ESEA Flexibility at all, and those who have generally mistake them for federal incursion or unaccountable decentralization.\(^{366}\) If the Administration or Congress takes action on the basis of either of these

---


\(^{365}\) For example, the Department granted California a limited waiver from NCLB’s testing requirements that permitted the state to administer Common Core–aligned field tests to a sample of its students in 2014, rather than administering its traditional state tests to all students. Michele McNeil, California Wins Prized NCLB Testing Waiver, EDUC. WK. POLLS. K–12 BLOG (Mar. 7, 2014, 2:31 PM), http://blogs.edweek.org/edweek/campaign-k-12/2014/03/california_wins_prized_nclb_te.html.

\(^{366}\) See supra notes 257–62 and accompanying text. It is not clear whether the architects of the initiatives in the Obama Administration recognize the dynamics of the governance structure. On one hand, Secretary of Education Duncan has elaborated and defended the ways that ESEA Flexibility has allowed states to tailor their policies to local conditions, sparked continuous improvement, and permitted the Department to hold states accountable to their commitments and the broad goals of NCLB. See Arne Duncan, U.S. Sec’y of Educ., No Child Left Behind: Early Lessons from State Flexibility Waivers, Testimony to the U.S. Senate Committee on Health, Education, Labor, and Pensions (Feb. 7, 2013), available at http://www.ed.gov/news/speeches/no-child-left-behind-early-lessions-state-flexibility-waivers. On the other hand, Secretary Duncan has indicated that the Administration would prefer reauthorization to the waiver approach. See Joy Resmovits, Arne Duncan Defends No Child Left Behind Waivers in Senate Committee Hearing, HUFFINGTON POST (updated Feb. 7, 2013, 6:51 PM), http://www.huffingtonpost.com/2013/02/07/arne-duncan -waivers_n_2641155.html.
characterizations, rather than the disciplined devolution understanding, it would jeopardize the regime. Lessening the rigor of evaluation and monitoring to mitigate a perceived federal incursion would lead some states to shirk their substantive commitments and reduce the incentive for collaborative policy experimentation. Conversely, making the federal role more prescriptive to combat a perceived lack of accountability or standardization would limit states’ flexibility to tailor policies and experiment. It might also lead them to demonstrate formal compliance while ignoring or undermining the underlying federal goals (as under NCLB and ESEA).

Disciplined devolution’s instability in the face of these exogenous threats is concerning, but it does not diminish the model’s advantages over the other alternatives to bureaucratic governance. Big waiver authority vested solely in a federal official, without state or local involvement and without an expectation of collaborative input, will lack the legitimacy and transparency of disciplined devolution. Cooperative federalism with limited federal oversight and no process for institutional learning would be the sort of unaccountable decentralization that, for disciplined devolution, is merely a risk. An experimentalist regime without either the institutions of state and local government or the leverage of a federal exceptions process could be even more ephemeral than disciplined devolution.

Moreover, the exogenous changes that threaten disciplined devolution also pose a risk to the other models, which similarly depend on self-aware, motivated actors and a particular legislative and regulatory context to function. Indeed, the structure of disciplined devolution makes it possible to inoculate against these risks in a way that the other models cannot.

The education case study illustrates how disciplined devolution might protect itself from political change, substantive lock-in, obsolete default schemes, and category error. Two strategies, which are possibly complementary, may serve to shore up the new education federalism. First, rather than locking in substantive policies, the reauthorization of ESEA could seek to lock in the disciplined devolution governance structure. To do so, legislation would have to establish a default scheme and include not only the authority for the Secretary of Education to waive its provisions for states whose alternative schemes better meet the statute’s goals, but also include an incentive for states to seek waivers. The legislation should also require the Education Department to undertake ongoing monitoring of the states’ efforts—focused on success in achieving the statute’s underlying goals and maintenance of a collaborative, continuous policy experimentation process—and to facilitate the

---

367. Financial incentives to innovate could serve this purpose. An intentionally burdensome default scheme would also provide an incentive to deviate. That may be too much of a risk, however, as political conditions or lack of capacity may result in some states implementing policies that were designed for their punitive value, not for their effectiveness in achieving the statute’s educational goals.
sharing of best practices across states, as suggested above. Furthermore, the legislation should include either a streamlined reauthorization process or a regulatory process by which the Education Department, representatives of state and local governments, and other stakeholders periodically review and jointly develop updates to the default scheme.

A second strategy would leverage the existing interstate collaboration of the consortia working on the Common Core to reinforce other elements of the governance structure of the new education federalism. Specifically, the consortia could establish a monitoring process among their member states much like the Department of Education’s monitoring process. Consortia staff could manage the monitoring, but ideally it would include an element of peer review, with representatives of member states participating in the monitoring of other states. The Department of Education could support this expanded role for the consortia by providing additional funding, oversight of the consortia’s monitoring process, and deferral of federal monitoring to states that undergo a validated consortium or peer review.

CONCLUSION

American public governance is in transition. As traditional legislative and bureaucratic approaches to solving public problems become less effective, various alternatives have emerged. This Article has identified one such alternative, disciplined devolution, which has been obscured by its relationship to other strategies. The early implementation of disciplined devolution in the education context suggests that it holds promise as a governance model, balancing flexibility and experimentation with accountability. This balance fosters stakeholder collaboration and motivates adherence to national goals and the governance model, thereby curing some of the deficiencies of the big waiver, cooperative federalism, and experimentalism models with which disciplined devolution overlaps.

But the education case study also shows that disciplined devolution is fragile. Internal tensions between centralization and decentralization have

---

368. Neither of the bills that have passed out of committee effectively reinforces disciplined devolution. The Senate Democrats’ reauthorization bill incorporates elements of structured flexibility with respect to some of its requirements, such as setting student achievement and growth targets and devising interventions in struggling schools, but not others. See Klein, supra note 363. The House Republicans’ bill does not envision an ongoing governance role for the federal government at all. See Klein, supra note 360.

369. Having states monitor one another through the Common Core consortia is analogous to the reporting and monitoring regimes that some industries have established through their trade associations. Studies of such self-policing regimes have found effective continuous learning processes and improvement in key outcome measures. See, e.g., Dorf & Sabel, supra note 1, at 371–73 (describing Institute of Nuclear Power Operations); Charles F. Sabel & William H. Simon, Contextualizing Regimes: Institutionization as a Response to the Limits of Interpretation and Policy Engineering, 110 Mich. L. Rev. 1265, 1278–85 (2012) (describing California Leafy Greens Products Handler Marketing Agreement).
disrupted the balance of the model, causing the Obama Administration’s education initiatives to fall short of theoretical expectations of ongoing collaboration and sharing of best practices across states. External political shocks pose a more systemic risk, threatening a retrenchment to bureaucracy or an abandonment of federal oversight. I have offered some tentative strategies for fortifying the disciplined devolution model against these risks, including interstate peer review, federal-local waivers, a disciplined devolution statute, and informal interstate collaboration. Each of these merits further exploration.

Throughout this Article, while analyzing closely only the education context, I have maintained that disciplined devolution has more general applicability. The Article points to programs under the ACA and welfare-to-work in which features of disciplined devolution are evident or incipient.370 On the one hand, it may be that the dynamic I have described in the education context is not feasible in these or other contexts. On the other hand, the structure may be achieved by recognizing its potential applicability, incorporating any missing components, and adapting them as required by the context. As increasingly fractured national politics push more and more policy making to federal administrative agencies and down to the state and local level, experiments in alternative governance structures—some more successful than others—are likely to blossom. The benefits identified by this Article’s analysis of the new education federalism suggest that disciplined devolution has the potential to be one of the successful alternatives, and that it would be worthwhile to seek out opportunities to apply and reinforce nascent instances of it.

370. See supra notes 114–17 and accompanying text.