FEDERAL LABORATORIES OF DEMOCRACY

Hannah J. Wiseman[1] and Dave Owen[2]

Abstract: Facilitating state policy experimentation is an oft-cited justification for the United States' federalism system. Despite growing recognition of risk aversion, free riding, and other disincentives to state-led experimentation, the mythology of state laboratories still dominates the discourse of federalism. We propose a framework that counters this entrenched assumption and enables more productive analysis of policy experimentation. The Article explores a continuum of experimental approaches that differ in terms of the degree of experimental rigor that they incorporate—such as the extent to which they control for confounding variables—and the governance levels at which they are designed and implemented. We apply this new analytical framework to case studies from divergent policy areas, including agricultural, natural resources, and education law. These examples highlight rigorous experiments designed and largely administered by federal agencies.

Our framework and case studies turn the concept of the "laboratories of the states" on its head, showing that experimentation can, often does, and should occur at multiple levels, including the federal level. In countering and adding nuance to traditional experimentation accounts, the Article reveals the benefits of federal involvement in policy experiments, and thus the perils of weakening federal authority or excluding federal involvement in an effort to enhance core federalism values like experimentation. Federal expertise and resources—and even the simple availability of experimental platforms, such as federally-owned and managed lands—often give the federal government a comparative advantage in the policy experimentation field. This is not to say that the federal government should always lead and implement experiments, but it calls attention to the importance of understanding experimentation as a multi-level endeavor that extends well beyond the states.

---

[1] Attorneys' Title Professor of Law, Florida State University College of Law. J.D., Yale Law School, A.B., Dartmouth College.

[2] Professor of Law, University of California—Hastings College of Law, J.D., University of California, Berkeley School of Law, B.A., Amherst College. The authors thank Erin Ryan for early comments at the brainstorming stage of this article, including thoughts on Forest Service experimentation, and Scott Dodson, Malcolm Feeley, and Jodi Short for comments on drafts. Thanks also to the Florida State University Research Center and University of California, Hastings librarians for locating many of the historic and scientific sources cited in this article.

TABLE OF CONTENTS

INTRODUCTION

An oft-cited justification for federalism is that it induces creative policy experimentation at the state level.[3] According to the standard arguments, limiting federal power and protecting state sovereignty will allow states to function as our "laboratories of democracy," the places where governmental innovations can begin and spread.[4] For courts and federalism scholars, this alleged virtue has remained alluring for decades, and celebrations of state policy laboratories continue to occupy a central place in the discourse of federalism.[5] Similarly, much of the literature on policy experimentation tends to assume, if it confronts questions of federalism at all, that states (and sometimes local government) will be the experimenters.[6]

---

[3] *See, e.g.*, Michael Abramowicz et al., *Randomizing Law*, 159 U. PA. L. REV. 929, 946 (2011) (noting and critiquing "a frequent justification of federalism—that allowing states to make independent choices provides a kind of laboratory to test policies").

[4] *See, e.g.*, New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) (describing states as "laboratories"); Henry M. Hart, Jr., *The Relations Between State and Federal Law*, 54 COLUM. L. REV. 489, 493 (1954) ("The federal system has the immense advantage of providing forty-eight separate centers for . . . [legislative] experimentation.").

[5] *See, e.g.*, Arizona State Legislature v. Arizona Independent Redistricting Comm'n, 135 S.Ct. 2652, 2673 (2015) ("This Court has 'long recognized the role of the States as laboratories for devising solutions to difficult legal problems.'") (quoting Oregon v. Ice, 555 U.S. 160, 171 (2009); Ann Althouse, *Vanguard States, Laggard States: Federalism and Constitutional Rights*, 152 U. PA. L. REV. 1745, 1752-75 (2004) (describing courts' reliance on this justification.

[6] *See, e.g.*, WALLACE E. OATES, FISCAL FEDERALISM 12 (1972) (arguing that a decentralized system of governance can generate "greater experimentation and innovation in the production of public goods"); Hongbin Cai & Daniel Treisman, *Political Decentralization and Policy Experimentation*, 2009 QU. J. OF POL. SCI. 35, 35-36 ("political decentralization has been widely thought to stimulate policy experimentation and innovation"); Christos Kotsogiannis & Roberts Schwager, *On the incentives to experiment in federations*, 60 J. URB. ECON. 484, 484 (2006) ("A commonly held view is that fiscal federalism promotes innovative public programs, [and] speeds up the process of policy experimentation and its diffusion."). *But see* Wallace Oates, An *Essay on Fiscal Federalism*, J. ECON. LIT. 120 (1999) (acknowledging that centralized governments also can experiment). There is a broad literature closely investigating states' experimentation with policy and documenting what the authors believe to be diffusion of that policy, although the extent to which this experimentation and diffusion occurs is disputed. For literature investigating what authors believe to be evidence of experimentation and diffusion, *see, e.g.*, ANDREW KARCH, DEMOCRATIC LABORATORIES: POLICY DIFFUSION AMONG THE AMERICAN STATES (2007) (arguing that effective innovation and diffusion of policy occurs within and across state

Yet there are many reasons to be skeptical of these accounts. While no one disputes the fact that state and local governments sometimes do innovate, a variety of characteristics of state and local governments make it unlikely that they will experiment nearly as often as traditional federalism theory would assume.[7] Even when they do experiment, other characteristics of state and local governments may hinder good policies' paths to wider adoption.[8] Consequently, if we value policy laboratories and cannot count on states to assume that role, then it is important to consider how other elements of our federalist system can enable policy experimentation, or something closer to that ideal. This endeavor is particularly valuable in an era of political upheaval and growing calls for massive reduction in federal governmental "interference." If states are not the optimal experimenters, then broad-based shrinkage of federal involvement could have the perverse effect

---

lines); Frances Stokes Berry & William D. Berry, *State Lottery Adoptions as Policy Innovations: An Event History Analysis*, 84 AM. POL. SCI. REV. 395, 410-11 (1990) (concluding on the basis of an empirical model that states sometimes overcome obstacles to innovation); Michael Mintrom, *Policy Entrepreneurs and the Diffusion of Innovation*, 41 AM. J. POL. SCI. 738, 738-39 (1997) (summarizing the broad political science literature on policy innovation and diffusion).

[7] *See, e.g.*, Susan Rose-Ackerman, *Risk Taking and Reelection: Does Federalism Promote Innovation?*, 9 J. LEGAL STUDIES 593 (1980) (concluding that due to risk aversion, free riding, and other problems there is unlikely to be much efficient innovation purely at the local level); Brian Galle & Joseph Leahy, *Laboratories of Democracy? Policy Innovation in Decentralized Governments*, 58 EMORY L. J. 1333, 1369, 1398 (2009) (concluding that "there is social underprovision of experimentation by small jurisdictions," that "the quality of the information generated . . . is likely below the theoretical ideal," and that "absent outside intervention, state and local governments will on the whole innovate at well below the socially optimal level"), Edward L. Rubin & Malcolm Feeley, *Federalism: Some Notes on a National Neurosis*, 41 UCLA L. Rev. 903, 912 (1994) (arguing that federalism "only makes sense" when individuals in different regions have different rights-based preferences because the federal government could just as easily "choose more effective instrumentalities" for reaching a particular goal and "adapt the selected instrumentalities to local circumstances"); MALCOLM M. FEELEY & EDWARD RUBIN, FEDERALISM: POLITICAL IDENTIFY & TRAGIC COMPROMISE (2008) (challenging the common "laboratories of the states" justification for federalism and other common alleged virtues of federalism); Koleman S. Strumpf, *Does Government Decentralization Increase Policy Innovation?* J. PUB. ECON. THEORY 207, 208 (2002) (noting that "a local policymaker may free-ride off his neighbor's experiment").

[8] *See, e.g.*, Berry & Berry, *supra* note 6, at 401-405 (describing factors that hinder or impede innovation and diffusion in the state lottery context, such as the financial health of the state and the percentage of the population that adheres to fundamentalist religious views as well as whether the proposed adoption is in an election year and involves re-election of incumbents); Virginia Gray, *Innovation in the States: A Diffusion Study*, 67 AM. POL. SCI. REV. 1174, 1185, 1176 (1973) (noting that "diffusion patterns differ by issue area and by degree of federal involvement" and that "[h]ard-to-amend limitations in the state's constitution or values of the political subculture might causes a state's leaders to be practically immune to diffusion").

of reversing critical policy experimentation, and thus undermining a virtue often used to justify state power.

This Article responds to the challenge of constructing useful policy laboratories, and to the inadequacies of traditional theories, through closer attention to the intersections of experimentation and federalism. We craft a conceptual framework to fuse key attributes of policy experimentation with the United States' federalist system. We then flesh out this framework and demonstrate its analytical capacity by discussing several real-world policy initiatives. The governance structures for the policy initiatives we describe are all quite different from the stereotypical "laborator[ies] of the states," and they also are directly at odds with the popular myth that "the central government can examine only one policy at a time and so will slowly uncover superior new policy choices."[9] Instead, these experiments involve the federal government in both designing and implementing experiments, sometimes without much help from the states, and sometimes relying on state and local entities to cooperate in experiments facilitated by the federal government.[10]

Our core thesis is that these arrangements are not anomalous. In a federalist system of hierarchical and decentralized governance, a key driver of experimentation often will, and should, be the federal government.[11] Furthermore, federal initiatives sometimes incorporate attributes that make policy experimentation surprisingly rigorous[12]—far more rigorous than the haphazard patchwork of state policies that arise from the largely decentralized experimentation envisioned by most federalism proponents.

Our primary case study, which has attracted scant attention in the legal literature, explores ambitious experiments in U.S. agricultural policy that evolved over nearly a century. The experiments began with a federally-

---

[9] Strumpf, *supra* note 7, at 208 (summarizing, though not adopting, this widespread view).

[10] There is an extensive literature on federal-local collaboration, but this literature has not tended to focus on how the federal government enlists local governments or works at the local level to conduct policy experimentation. See, e.g., Nestor Davidson, *Cooperative Localism: Federal-Local Collaboration in an Era of State Sovereignty*, 93 VA. L. REV. 959 (2007).

[11] The political science literature has made this observation, but it rarely arises in the legal literature, and courts continue to follow the opposite assumption. For arguments in the political science context suggesting that centralization can generate more—and sometimes too much—experimentation, see, e.g., Cai & Treisman, *supra* note 6, at 36 (observing that "[e]xplicitly experimental local policies occur in both centralized dictatorships and centralized democracies"); *id*. (summarizing similar observations by other political scientists and economists); Kotsogiannis & Schwager, *supra* note 6, at 485 (arguing that a decentralized system is conducive to producing fewer policy innovations than a centralized one"). This literature tends to use models to assess the likelihood of experimentation rather than to discuss governance structures and the mechanics.

[12] *See infra* Part III.B.

designed and federally-implemented approach. Using Congressionally-approved funding, the federal government employed a true boots-on-the-ground system for modifying crop management practices that had contributed to massive dust storms and loss of valuable topsoil. The United States Department of Agriculture sent federal agents to far-flung rural locations to build experimental research stations that tested and demonstrated improved soil conservation techniques to farmers; the federal government also enlisted the Civilian Conservation Corps to implement these techniques around the country, with approaches that varied by region.[13] These early federal efforts have since morphed into a complex federal-state-local program that involves rigorous experimentation, including clear standards for agricultural conservation policies, sophisticated approaches for measuring the results of federally-supported conservation practices, and hundreds of scientific papers reporting on these results and suggesting how conservation practices could improve.[14]

Other examples, such as federal agencies' evolving policies for wildfire management[15] and the United States Department of Education's Race to the Top,[16] demonstrate that the soil conservation story has parallels in other fields.[17] Each program has been different, yet these projects share an activist federal role not just in funding innovation or compiling data, but also in selecting hypotheses and sometimes carrying out the actual experiments. In combination, these examples help illustrate the possibility, and also explore the merits, of policy approaches that fall at previously underappreciated points within our experimentation framework. They show the value, in other words, of taking the laboratories of democracy concept well beyond state (and local) government. And they show the perils of judicial, legislative, and academic tendencies to equate policy experimentation exclusively with subfederal governance.

To begin the project of melding experimental design with federalism and producing meaningful lessons for policy experimentation, Part I describes four literatures that dance around the role of federal experimentation: the traditional federalism literature, which tends to assume that *state* experimentation flows naturally from a federalist system; the experimental design and adaptive management literatures, which focus on experimental systems without grounding their analyses in federalist structures; and the experimental governance literature, which attempts to

---

[13] *See infra* Part III.B.1.

[14] *See infra* Part III.B.1.

[15] *See infra* Part III.B.2.

[16] *See infra* Part III.B.3.

[17] We make no claim, of course, that our examples exhaust the field of federal experimentation.

marry federalism and experimentation yet discounts the experimental potential of the federal government itself.[18] Part II explains why the gaps in these bodies of literature are important. After describing what we mean by experimentation, we explore the challenges of fitting real experimental policy approaches into a state-centric federalist structure. In light of the substantial room for improvement identified in Parts I and II, in Part III we provide an analytical framework of governance and policy experimentation, including approaches that incorporate different degrees of experimental rigor and that rely on varying levels of federal involvement. We then apply this framework to case studies and examples, showing how the typology can shift focus to more productive arrangements for experimental governance.

Finally, Part IV draws generalizable lessons. We offer no magic formula; instead, the analysis in Part III shows that effective governance has flowed from experiments conducted by the federal government and from experiments designed by the federal government and implemented by state and local actors with federal guidance. A single level of government is not universally superior in terms of differentiating the experiment or measuring, collecting and reporting data. Indeed, recent threatened federal intervention in some policy areas threatens to interfere with key state and local experiments.[19] Nevertheless, the federal government brings important advantages to policy experimentation, including, most importantly, the resources at its disposal and its combination of centralized coordination and partially decentralized operations. And even when the federal government lacks the money, staff, and expertise to implement an experiment itself, it can play an important role in harnessing resources at other levels, coordinating the experiment, and initiating productive reporting of lessons learned across local and state borders.[20]

---

[18] For a very rare exception to these generalizations, see Joseph Landau, *Bureaucratic Administration: Experimentation and Immigration Law*, 65 DUKE L.J. 1173 (2016). Landau explains how federal immigration policies illustrate the possibility of federal experimentation, while also noting that [m]any of the mechanisms associated with more experimental regimes remain untapped within federal immigration law." *Id.* at 1238.

[19] *See, e.g.*, Coalition for Competitive Elec. v. Zibelman, appeal from Dynegy v. Zibelman, 2017 WL 3172866 (S.D.N.Y., July 25, 2017); Village of Old Mill Creek v. Star, 2017 WL 3008289 (N.D. Ill., July 15, 2017). Both cases address whether the Federal Power Act preempts state clean energy initiatives.

[20] This supports earlier, similar suggestions by scholars such as Ed Rubin, Malcolm Feeley, and Michael Livermore, who pointed out that the federal government can help coordinate experiments. *See, e.g.*, Michael A. Livermore, *The Perils of Experimentation*, 126 YALE L.J. 636, 644 (2017) (briefly arguing in favor of "managed experimentation," in which the federal government, states, and local governments all play a role); FEELEY & RUBIN, *supra* note 7 (describing how the federal government could coordinate experiments and vary approaches by locality or region).

Those lessons have important implications not just for federalism theorists, but also, more practically, for legislators, judges, and advocates. Legislators often draft statutory provisions—"big waivers" are just one prominent example—designed to facilitate policy innovation, and they often look exclusively to states to supply that innovation.[21] Our analysis reveals that state focus to be overly myopic; legislators also should consider using the federal government itself to pursue policy experiments. Similarly, courts often invoke policy experimentation as a rationale for limiting federal authority.[22] That rationale, we show, is also often misguided. For advocates, the lessons are more nuanced. Persuading any level of government to adopt experimental policies is difficult, and an advocate's best option will usually be the governance level that is willing to try, not the one that would be optimal locus of policy experimentation in some perfect world. Beggars, after all, cannot be choosers. Nevertheless, would-be policy entrepreneurs still should keep the federal government in mind as a possible focus, and sometimes the preferred focus, of their advocacy.

Intentional policy experimentation will rarely match the type of carefully-planned experimentation that occurs in a scientific laboratory. Indeed, Justice Brandeis, in coining the famous laboratories term, likely did not envision policy development processes that would resemble scientific experiments.[23] But defining a new space for policy experimentation broadens opportunities for designing meaningful and more effective experiments at several levels of government, far beyond a simplified state-centric approach.

I. Experimentalism and Federalism: An Overview

For decades, the Supreme Court has spoken of federalism and experimentation in the same breath.[24] Academics often echo the judicial statements and sometimes elaborate upon them. But even as federalism has remained a central focus of legal-academic inquiry, and as sophisticated schools of thought have grown up around the idea of experimental

---

[21] *See* David J. Barron & Todd D. Rakoff, *In Defense of Big Waiver*, 113 COLUM. L. REV. 265 (2013) (describing the granting of waivers to states as a way to induce policy experimentation).

[22] *See infra* notes 31-34 and accompanying text.

[23] *See* Barry Friedman, *Valuing Federalism*, 82 MINN. L. REV. 317, 377 (1997) ("'Innovation' might have been a better word choice for Justice Brandeis than 'experimentation,' saving us all a lot of bother."). *But see* Althouse, *supra* note 5, at 1751 (arguing that "Justice Brandeis does not appear to view 'experimentation' as a metaphor. His government policymakers operate 'in the fields of social and economic science.'").

[24] *See* Livermore, *supra* note 20, at 648 ("Within legal scholarship, experimentation is often understood through the lens of federalism.").

governance, most exploration of the intersections of federalism and experimentalism has remained relatively cursory. Academics and judges do routinely discuss governance innovations instigated by the federal government. But when they turn from specific examples to broader theories, their discussions typically adopt fairly simplistic models of experimentalism or federalism—or both.

This Part introduces this theoretical background. We begin with the Supreme Court's discussions of the laboratories of democracy and then turn to key areas of academic literature that consider the intersections of federalism and experimentalism. Each area, we show, does not grapple with some of the key complexities of the intersecting terrain of experimental design and a federalist governmental structure.

A. Experimentalism and Traditional Federalism Theory

In 1932, a dispute over Oklahoma's regulation of ice companies reached the United States Supreme Court, and a majority of the Court, in a forgettable opinion, set the regulatory controls aside.[25] Justice Brandeis dissented. The nation then was deep in the Great Depression—"an emergency more serious than war," in Brandeis' words.[26] As he acknowledged, the path out of those dark times was far from clear.[27] Brandeis was sure of one thing, however: "[t]here must be power in the states and the nation to remould, through experimentation, our economic practices and institutions to meet changing social and economic needs."[28] That capacity for experimentation, he noted, was closely tied to federalism. "It is one of the happy incidents of the federal system," he wrote, "that a single courageous state may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country."[29]

[25] New State Ice Co. v. Liebmann, 285 U.S. 262, 306 (1932).

[26] *Id.* at 306.

[27] *Id.* at 309 ("Whether [the view that increased regulation of economic competition is necessary] is sound nobody knows. The objections to the proposal are obvious and grave.").

[28] *Id.* at 311.

[29] *Id.* at 386-87.

That sentence has become iconic.[30] The Court often quotes it, or, in more shorthand form, refers to "laboratories of democracy."[31] The references come in opinions by liberals and conservative justices alike and in cases addressing a wide variety of subject matter.[32] On an often-divided Court, the value of state laboratories is one thing everyone can agree on.[33] Yet the judicial references also are strikingly brief. Absent from the Court's opinions is any effort to explore the nuances of these state laboratories, or to define in any detail the conditions that allow them to succeed in fulfilling their celebrated role (let alone to make satisfaction of those conditions a factor as the Court weighs the legality of state action). Nor has the Court done much to extend its discussion of the laboratories of democracy to the many non-state governing entities that populate our federalist system.[34] The implicit assumptions, instead, appear to be that experimentalism will automatically emerge from federalist governance and that the locus of experimentation will be the states.

A similar theme emerges from much of the classic academic work on federalism.[35] Academic scholars repeatedly identify federalism with policy experimentation. This tendency dominates scholarly work by dual federalists, who argue that a strong separation between federal and state

---

[30] Other justices had said similar things, but with less memorable phrasing. *See, e.g.*, Truax v. Corrigan, 257 U.S. 312, 344 (1921) (Holmes, J. dissenting) (lamenting use of the Fourteenth Amendment to "prevent the making of social experiments that an important part of the community desires, in the insulated chambers afforded by the several states").

[31] *See* Fisher v. University of Texas at Austin, 136 S.Ct. 2198, 2214 (2016); Arizona State Legislature v. Arizona Independent Redistricting Comm'n, 135 S.Ct. 2652, 2673 (2015); Oregon v. Ice, 555 U.S. 160, 170 (2009); Chandler v. Florida, 449 U.S. 560, 579 (1981); Arizona v. Evans, 514 U.S. 1, 8 (1995); Reeves v. Stake, 447 U.S. 429, 441 (1980); Whalen v. Roe, 429 U.S. 589, 597 & n.20 (1977); San Antonio Ind. Sch. Dist. v. Rodriguez, 411 U.S. 1, 49-50 (1973); Fay v. People of State of N.Y., 332 U.S. 261, 296 (1947).

[32] *See, e.g.*, Evenwel v. Abbott, 136 S.Ct. 1120, 1141 (2016) (Thomas, J. concurring in the judgment) ("as the Court recently reminded us, States are free to serve as "'laboratories'" of democracy.") (quoting *Arizona State Legislature v. Arizona Independent Redistricting Comm'n*, 135 S.Ct. 2652, 2673 (2015). The Court's *Arizona State Legislature* opinion, which Justice Ginsburg authored, in turn quotes Justice Kennedy's concurring opinion in *United States v. Lopez*, 514 U.S. 549, 581 (1995), for the same proposition. 135 S. Ct. at 2673.

[33] While everyone may agree on the principle, it nevertheless tends to pop up in dissents. *See, e.g.*, Gobeille v. Liberty Mut. Ins. Co., 136 S.Ct. 936, 957 (2016) (Ginsburg, J. dissenting); Gonzalez v. Raich, 545 U.S. 1, 42-43 (2005) (O'Connor, J. dissenting); Boy Scouts of America v. Dale, 530 U.S. 640, 664 (2000) (Stevens, J. dissenting).

[34] For a very rare example of the Court applying this reasoning to a local government, see San Antonio School Dist. v. Rodriguez, 411 U.S. 1, 49-50 (1973) (noting "[a]n analogy to the Nation-State relationship in our federal system").

[35] *See, e.g.*, Friedman, *supra* note 4, at 399 ("Common intuition suggests that the vast majority of techniques used today to govern were developed at the state and local level.").

powers facilitates experimentation,[36] and from "dynamic" or "interactive" federalists, who argue that overlap between federal and state powers similarly enables and strengthens policy experimentation.[37] Yet work by both schools of thought rarely attends to either the actual mechanics of experimentation or the experimental possibilities of federal governance.[38] One could form the impression that state-centered experimentalism will spring naturally, like weeds in well-watered and fertile soil, from the policy differentiation and intergovernmental interaction that federalism creates. There are, of course, exceptions to this generalization, which we discuss in more detail below. But for the most part, traditional federalism theories have celebrated state experimentation while ignoring its mechanics.

B. Experimental Design and Adaptive Management

While federalism theory has tended to devote only fleeting attention to the methods of governmental experimentation, other bodies of theory have made such experimentation their central focus, but often with scant attention to the governance structures at the heart of federalism theory. Two bear mention here.

The first area of work, which for simplicity we will call the experimental design literature, focuses on improving the frequency and rigor of policy experiments.[39] Some articles at the edges of this vein just call for

---

[36] *See* Livermore, *supra* note 20, at 648-49 (summarizing and citing scholarship in this realm); *see also* Gonzalez v. Raich, 545 U.S. 1, 42-43 (2005) (O'Connor, J. dissenting) (arguing that protecting "historic spheres of state sovereignty from excessive federal encroachment" allows states to function as laboratories of democracy).

[37] *See* Heather C. Gerken, *Federalism as the New Nationalism – An Overview*, 123 YALE L.J. 1889, 1902 (summarizing and citing multiple sources in this vein).

[38] *See* Dave Owen, Regional Federal Administration, 63 UCLA L. Rev. 58, 113-14 (2016) (noting that conventional schools of federalist thought often assume effective communication among levels of government). For an exception, see Landau, *supra* note 18 (finding hints of an experimental federal role in immigration policy). Some environmental federalism work has focused on ways in which federal, state, local, and sometimes foreign governments can facilitate the spread of policy innovations, and this work acknowledges the possibility of experiments beginning with the federal government. *See, e.g.*, Kirsten H. Engel, *Democratic Environmental Experimentalism*, 35 J. ENVTL. L. 57, 70-71 (2017); Ann E. Carlson, *Iterative Federalism and Climate Change*, 103 NW. U. L. REV. 1097 (2009) But none of this work focuses on either the historic reality of federal experimentation or the advantages the federal government brings to developing experimental policy.

[39] *See, e.g.*, Daniel E. Ho, *Does Peer Review Work? An Experiment of Experimentalism*, 69 STAN. L. REV. 1 (2017); Abramowicz et al., *supra* note 3. Kenneth Abbott and Duncan Snidal's recent work bridges this area and the experimentalist governance literature, which we discuss in more detail below. *See* Kenneth W. Abbott & Duncan Snidal, Experimentalist Governance 2.0: Taking "Experiments" (More) Seriously,

more governmental experimentation.[40] But the core studies argue that government should conduct rigorous, randomized policy experiments, and some work demonstrates, through real-world examples, how this can be done and the insights such experimentation can produce.[41]

The second body of work considers "adaptive management."[42] In contrast to the experimental design literature, which focuses on developing rigorous and discrete policy experiments, the adaptive management literature makes a sweeping claim: *all* policy interventions are experimental and should be treated as such.[43] Proponents of adaptive management argue that in a world of uncertainty and limited knowledge, policy must and should evolve through learning.[44] That means treating policies as provisional experiments, monitoring their results, and continuously adjusting them.[45]

The adaptive management literature is enormous, and it explores a wide range of sub-issues.[46] But for our purposes, just one point about both the adaptive management and experimental design literatures is particularly important. Each typically focuses on experiments implemented by a single governing entity,[47] and neither has much to say about how its recommended

---

https://www.buffalo.edu/content/www/baldycenter/events/speakers/_jcr_content/par/download/file.res/Abbott%20XG%20Buffalo%209-16.pdf. *See also* ROBERT A. SCHAPIRO, POLYPHONIC FEDERALISM: TOWARD THE PROTECTION OF FUNDAMENTAL RIGHTS 6 (2009) (arguing for a focus on how experimentation within federalism works); Doni Gewirtzman, *Complex Experimental Federalism*, 63 BUFF. L. REV. 241 (2015) (focusing on the need to "identify the traits that allow decentralized systems to innovate effectively" and the extent to which our governance system "contains those traits").

[40] *See, e.g.*, Michael Greenstone, *Toward a Culture of Persistent Regulatory Experimentation and Evaluation*, *in* NEW PERSPECTIVES ON REGULATION 111 (David Moss & John Cisternino, eds., 2009).

[41] *See, e.g.*, Ho, *supra* note 39; Abramowicz et al., *supra* note 3.

[42] *See, e.g.*, Robin Kundis Craig & J.B. Ruhl, *Designing Administrative Law for Adaptive Management*, 67 VANDERBILT L. REV. 1 (2014).

[43] *E.g.* KAI N. LEE, COMPASS AND GYROSCOPE: INTEGRATING SCIENCE AND POLITICS FOR THE ENVIRONMENT 9 (1993) ("Policies are experiments; *Learn from them.*") (emphasis in original).

[44] *See* Craig R.Allen et al., *Adaptive Management for a Turbulent Future*, 92 J. ENVTL. MGMT. 1339, 1339 (2011).

[45] *See, e.g.*, Bradley C. Karkkainen, *Toward a Smarter NEPA: Monitoring and Managing the Government's Environmental Performance*, 102 COLUM. L. REV. 903, 907-08 (2003).

[46] The literature critiquing adaptive management also is extensive. *See, e.g.*, Dave Owen, *Probabilities, Planning Failures, and Environmental Law*, 84 TULANE L. REV. 265, 330-35 (2009) (offering a qualified critique of adaptive management while conceding its value in some circumstances); Holly Doremus, *Adaptive Management, The Endangered Species Act, and the Institutional Challenges of "New Age" Environmental Management*, 41 WASHBURN L.J. 50 (2001) (same).

[47] *But see* Susan Welch & Kay Thompson, *The impact of federal incentives on state policy innovation*, 24 AM. J. POL. SCI. 715, 717-728 (1980) (discussing federal grants-in-aid

governance approaches should be integrated with a federalist system.[48] In our view, this is not a failing; the lack of attention to federalism just reflects the authors' desire to keep their examples simple, and to explore principles that should not be limited to any particular governance structure. But it does mean that these bodies of literature address the intersections of federalism and experimentalism only obliquely.

C. Democratic Experimentalism Theory

In contrast to the traditional federalism, experimental design, and adaptive management literatures, one school of thought has consciously built itself around the intersection of experimentalism and federalist structures. In a series of articles, Charles Sabel, Michael Dorf, and other academic authors have argued for "democratic experimentalism," a system in which governance occurs through continuous processes of goal setting, policy innovation, measurement, reexamination, and adjustment (in other words, setting "benchmarks" based on desired results and changing policy accordingly).[49] Federalism is central to this vision. As Dorf and Sabel explain, "[t]he chief role of Congress in such a system would be to authorize and finance experimental reform by states and other subnational jurisdictions," and federal agencies would help with benchmarking efforts and other aspects of the experiment.[50] They are not alone in proposing this approach. Other advocates of forms of experimental governance have articulated a similar vision, in which the federalist system allows the national government to play a facilitative role in state experimentation.[51]

---

and other incentives and the extent to which they cause state policy innovation and diffusion). Some literature also discusses flexible governmental approaches (not experimentation) through shared agency control at one governmental level. *Cf.* Jody Freeman and Jim Rossi, *Agency Coordination in Shared Governmental Space*, 125 HARV. L. REV. 1131, 1192 (2012) (arguing that agency coordination through memoranda of understanding produces flexibility that "is advantageous because it allows agencies to adapt to new circumstances over time without resorting to elaborate and time-consuming procedures").

[48] *See, e.g.*, Craig & Ruhl, *supra* note 42, at 63-87 (providing "The Model Adaptive Management Procedure Act," which envisions action by a single federal agency).

[49] *See* Charles Sabel & William H. Simon, *Minimalism and Experimentalism in the Administrative State*, 100 GEORGETOWN L.J. 53 (2011); Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267, 288, 345 (1998).

[50] Dorf & Sabel, *supra* note 49, at 345.

[51] *E.g.* Orly Lobel, *The Renew Deal: The Fall of Regulation and the Rise of Governance in Contemporary Legal Thought*, 89 MINN. L. REV. 342, 381 (2004) (identifying experimental governance with devolution to state and local governments); *see also See, e.g.*, Lisa Larrimore Ouellette, *Patent Experimentalism*, 101 VA. L. REV. 65, 121-25 (2015) (envisioning an experimental system with an international body at the coordinating center and nations as the sites of experimentation); Rose-Ackerman, *supra* note 7, at 615-616

This is an intriguing concept, and in some ways, the ideas we express in this article are extensions of this school of thought. But there are two key ways in which experimental governance scholars' visions of experimental federalism differ from the vision we will expound. First, as other commentators have pointed out, the experimental governance literature often treats policy experimentation as such a broad category that it removes much of the meaning from the term.[52] As Kenneth Abbott and Duncan Snidal note, "[a]ny situation in which actors 'try different things' is considered to be experimental, with little reference to the well-developed literature on experimentation in the natural and social sciences.[53]

Second, while the experimentalist governance literature embraces federalism as a source of experimentation, its federalist vision is narrowly cabined. Dorf and Sabel, for example, discuss intriguing examples of innovative federal policy,[54] but in their proposed governance model, "the state and local governments actually do the experimenting."[55] In contrast, the federal government, which they describe as highly centralized, facilitates experimentation (through mechanisms that are not always clear[56]) and does little else.[57] In reality, however, the federal government is itself decentralized in many ways, some of which can facilitate differentiation and

---

(although remaining skeptical of prospects for effective experimentation, noting that "a federal structure can encourage innovation by lower level governments" if national politicians spearhead an "innovation policy," such as awarding "grants to low-level governments on the condition that they carry out a search for new ways of doing things" or issue federal prizes to lower-level governments to reward "new ideas"); Livermore, *supra* note 20, (exploring "managed experimentation"). Brian Galle and Joseph Leahy observe that a federal facilitative role is not necessary in some scenarios, such as when state or local governments are "highly heterogenous" and will produce few positive externalities from which others would benefit; these entities might simply independently innovate. Galle & Leahy, *supra* note 7, at 1361.

[52] *See, e.g.*, Sabel & Simon, *supra* note 49, at 78-92 (treating as examples of experimentalism a wide variety of government programs, many of which do not appear to have some of the key attributes featured in democratic experimentalists' own definition of the term). Some of the experimentalist governance literature qualifies this sweeping use of examples by noting that the programs described exemplify some potential elements of an experimental governance regime. But in other places, those qualifiers seem to disappear.

[53] Abbott & Snidal, *supra* note 39.

[54] Dorf & Sabel, *supra* note 49, at 332-36, 382-88 (describing innovative federal policies).

[55] *Id*, at 428.

[56] *See* Abbott & Snidal, *supra* note 39 at 15-16 (arguing that "XG theory… lacks virtually any central management" and depends on "serendipity" to produce coordination and evaluation). Abbott and Snidal would add a more robust management role, but they still envision that management occurring through a centralized federal government while state governors lead the experiments.

[57] *E.g.* Dorf & Sabel, *supra* note 49, at 428.

experimentation.[58] Indeed, two of our case studies describe the federal government initially carrying out much of the experimentation itself—relying on agents in far-flung rural pockets of the country to test and demonstrate better agricultural conservation[59] and fire management[60] practices. And contrary to the vision of states serving as the decentralized entities that implement a federal experimental vision, state governments can be relatively central, at least in comparison to cities and other units of local government, and they, too can be the top-down initiators and organizers of lower-level experiments. [61] As our Race-to-the-Top example discusses, the federal government enlisted states to act as the organizers of a far more decentralized experiment in education reform—one largely carried out at the school district level.[62]

Finally, our governance systems also contain many actors that do not fit neatly into the traditional hierarchy of federal, state, and local governments typically described in the federalism and democratic experimentalism literatures.[63] The institutional landscape of federalism is thus messy and complex, and that complexity demands more nuance in discussions of the roles of different governing entities in an experimentalist system.

II. THE CHALLENGES OF EXPERIMENTALIST FEDERALISM

Scarce attention to the intersection of federalism and experimentalism would not be a problem if, as the traditional federalism literature seems to presuppose, state-centered experimentation just emerges naturally and

---

[58] *See* Owen, *supra* note 38, at 109-10; Richard Briffault, *What about the 'Ism'? Normative and Formal Concerns in Contemporary Federalism*, 47 VAND. L. REV. 1303. 1308 (1994) (noting that federalism arguments grounded in the virtues of decentralization would favor redistributing power from states to local governments).

[59] *See infra* note 141 and accompanying text.

[60] *See infra* notes 223-227 and accompanying text.

[61] Dave Owen, *Cooperative Subfederalism* (forthcoming).

[62] FUNDAMENTAL CHANGE: INNOVATION IN AMERICA'S SCHOOLS UNDER RACE TO THE TOP XVI, xiv (2015) (noting that "[s]tates are using performance management approaches to help districts support effective interventions in their lowest-performing schools, "[s]tates like Tennessee, North Carolina and Massachusetts created networks of their lowest-performing schools that improved supports for teachers and school and district leaders, and "Ohio districts hired former principals with track records of improving student achievement to coach principals in struggling schools." The Department of Education also distributed money directly to school districts as part of a district-specific Race to the Top program. DISTRICT REFORM SUPPORT NETWORK, TRANSFORMING THE CULTURE OF TEACHING AND LEARNING: FOUR RACE TO THE TOP-DISTRICT GRANTEES 4, https://rttd.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=21503.

[63] *See, e.g.*, Owen, *supra* note 61 (describing air quality management districts and land use planning agencies with territories encompassing multiple cities and counties).

frequently from a federalist structure. But there are many reasons to think it does not. This Part explains why purely state-centric federalism may not be such fertile ground for experimentation, and, therefore, why more careful exploration of arrangements that can produce experimentation is worthwhile.

A. Defining Policy Experiments

Before embarking on our critique of states as the assumed natural leaders of policy experiments, we provide a few words about what we mean by "experiment." In this article, we use the term to refer to processes that share, to at least some degree, several common attributes. These processes need not mirror the sterile halls of a scientific laboratory to count as an experiment, but they must exhibit some attributes of the traditional definition of this term.

- First, a policy experiment should reflect one or more hypotheses.[64] An experiment, at its core, is a test of an idea, and it is difficult to run a meaningful test without first deciding on the idea(s) to be tested.
- Second, experimentation requires policy differentiation. That differentiation might occur by design, as in a controlled, randomized experiment, or researchers may opportunistically exploit policy differences that arise naturally.[65] But in either case, the differentiation should allow a comparison that will put the experimental hypothesis to the test.[66]
- Third, experimentation requires control of confounding variables. In a controlled experiment, this can be done through randomizing the distribution of subjects into groups with different treatments, and through focusing differentiation on a single key attribute.[67] For natural experiments, such control is much more difficult.[68] Nevertheless, techniques like regression analysis can sometimes

[64] *See* Michael J. Saks, *Scientific Method: The Logic of Drawing Inferences from Empirical Evidence, in* MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY Vol. 1, 297, 304 (David Faigman et al., eds. 2016).

[65] *See, e.g.,* Mark R. Rozenssweig & Kenneth I. Wolpin, *Natural "Natural Experiments" in Economics*, 38 J. ECON. LIT. 827, 828 (2000).

[66] *See* Saks, *supra* note 64, at 304 (explaining that an experiment should try to falsify its hypothesis).

[67] *See* Abramowicz et al., *supra* note 3, at 934-37 (explaining how randomization works and the advantages it provides).

[68] *See id.* at 939-43 (describing advantages of randomization relative to regression analyses of natural experiments).

allow for enough control of enough confounding variables to facilitate qualified confidence in experimental conclusions.[69]

- Fourth, experimentation requires observation and data collection, and requires analysis of those observations and data.[70]
- Fifth, experimentation requires documentation. In academic settings, an experimental study often culminates in a write-up that explains the study and its results and analyzes its significance.[71] Typically, that paper will be peer-reviewed prior to publication.[72] In non-academic policy settings, peer review is less prevalent, though still potentially valuable, and we will use the term "experiment" to describe situations in which such review is absent.[73]
- Sixth, effective policy experimentation will require both repetition and adjustment of the experimental design. In most fields, researchers are reluctant to draw firm conclusions from individual experiments.[74] They instead try to reproduce experiments, both to test the validity of the original results and to discern the sensitivity of results to different interventions.[75]

---

[69] *See, e.g.*, Jeffrey J. Rachlinski, *Noah by the Numbers: An Empirical Evaluation of the Endangered Species Act*, 82 CORNELL L. REV. 356 (1997) (reviewing CHARLES C. MANN & MARK L. PLUMMER, NOAH'S CHOICE: THE FUTURE OF THE ENDANGERED SPECIES ACT (1995) (using statistical analysis to undercut claims that the Endangered Species Act is ineffective).

[70] *See* Saks, *supra* note 64, at 307-10 (describing the importance of observation, as well as the complex issues that can arise as experimenters decide what to measure and observe).

[71] One major problem with experimental research is that findings are more likely to be published if they are interesting, and they are more likely to be interesting if they are counterintuitive or surprising. That "publication bias" creates incentives for researches to interpret their results in more interesting ways and means that a skewed subset of results actually gets published. *See* Abramowicz et al., *supra* note 3, at 943 (describing this problem).

[72] *See* Jerome P. Kassirer & Edward W. Campion, *Peer Review: Crude and Understudied, but Indispensable*, 272 JAMA 96 (1994) (describing and critiquing peer review practices).

[73] *See* J.B. Ruhl & Jim Salzman, *In Defense of Regulatory Peer Review*, 84 WASH. U. L. REV. 1, 15 (2006) ("The use of peer review is far more limited and variable by agencies when exercising regulatory responsibilities.").

[74] *See* Saks, *supra* note 64, at 305 (noting that researchers become confident that a hypothesis is probably correct only after it has survived repeated attempts at falsification).

[75] *See* Arturo Casadevall & Ferric C. Fang, *Reproducible Science*, 78 INFECTION AND IMMUNITY 4972, 4972 (2010) (describing reproducibility as "a bedrock principle in the conduct and validation of experimental science").

One other point about our use of the term "experiment" bears mention. While we do not use the term as broadly as many legal authors,[76] we also do not limit it to situations that would score very highly on each of these metrics. It is usually quite difficult to manage policy experimentation with laboratory-style rigor,[77] and if the term "experiment" is to have more than occasional relevance to policy realms, it needs to include messier efforts. For that reason, we think policy experimentation is best thought of as a continuum, not a single, discrete category of action. A carefully designed, randomized experiment, which commentators often describe as the "gold standard," might reach the top of that continuum.[78] But we also extend the term to situations in which hypotheses exist but are somewhat muddy, confounding variables are imperfectly controlled, data collection happens but is uneven, and formal peer review does not exist. These tin-standard experiments still hold value in the real world—they are sometimes the highest level of experimentation that makes sense or the best we can hope for—and we therefore include them in our discussion.[79]

B. Barriers to Combining Policy Experiments and Federalism

While we have defined "experiment" broadly, much of the governing done within a federalist system will not fit even within that capacious definition. There are several reasons why, and we discuss these challenges within the definitional contours introduced in Part II.A.

1. Differentiation and Confounding Variables

[76] To be fair, the pots are calling the kettle black. We both have used the term "experiment" somewhat indiscriminately in our own past work.

[77] Of course, achieving rigor is also difficult in laboratories. Even dedicated scientists can fall victim to a variety of distorting effects, including confirmation and publication biases.

[78] *See, e.g.*, D. James Greiner & Cassandra Wolos Pattanayak, *Randomized Evaluation in Legal Assistance: What Difference Does Representation (Offer and Actual Use) Make*?, 121 YALE L.J. 2118, 2122 (2012). For arguments questioning whether researchers overvalue randomization, see Angus Deaton & Nancy Cartwright, *Understanding and Misunderstanding Randomized Control Trials*, http://www.nber.org/papers/w22595.pdf (working draft; last accessed August 22, 2017).

[79] *See infra* Part II.B (discussing reasons, some of them perfectly valid, why governments might not want to experiment).

Policy differentiation is the reason why a connection between federalism and experimental governance seems obvious.[80] Indeed, a federalist system seems designed, above all else, to allow different subnational jurisdictions to adopt differing policies.[81] But there are several reasons why the policy differentiation produced by federalism may not produce as much experimentation as one might expect.

First, as other commentators have pointed out, states' incentives for experimentation are often weaker than conventional federalism theory presupposes.[82] From a state or local perspective, the optimal policy approach may be not to experiment, but instead to adopt new approaches after some other jurisdiction has demonstrated that the approach works and has sorted out its kinks.[83] Imitation, after all, is usually easier than invention.[84] This means that there may be few to no first-mover innovators due to the prospect of free riding and the risk of losing votes when undertaking a new policy experiment. Or, alternatively, state leaders may focus on conforming their policy initiatives to the expectations of national political parties rather than seeking to forge an independent course.[85] There are countervailing incentives, of course; leaders can make their names as policy entrepreneurs,[86] and sometimes a subnational political climate will reward leaders for differentiating their jurisdiction from its neighbors or from the national

[80] *See, e.g.*, Gobeille v. Liberty Mut. Ins. Co., 136 S.Ct. 936, 957 (2016) (Ginsburg, J. dissenting) (drawing a connection between "state-law diversity" and "the role of States as laboratories").

[81] *See* Heather K. Gerken, *Exit, Voice, and Disloyalty*, 62 DUKE L.J. 1349, 1354-55 (2013) (noting this benefit of a federalist system); Michael W. McConnell, *Federalism: Evaluating the Founders' Design*, 54 U. CHI. L. REV. 1484, 1493 (1997) (reviewing RAOUL BERGER, FEDERALISM: THE FOUNDERS' DESIGN (1987)) ("The first, and most axiomatic, advantage of decentralized government is that local laws can be adapted to local conditions and local tastes, while a national government must take a uniform… approach.").

[82] *See* Rose-Ackerman, *supra* note 7 (arguing that federalism is not nearly as likely to promote experimentation as its proponents suggest); Galle & Leahy, *supra* note 7 (largely agreeing with Rose-Ackerman's conclusions).

[83] *See* Rose-Ackerman, *supra* note 7, at 610-11.

[84] *See* Scott Dodson, *The Gravitational Force of Federal Law*, 164 U. PA. L. REV. 703 (2016) (documenting state's reluctance to differentiate state law from federal even in realms where states have clear authority to chart their own courses).

[85] *See* Jessica Bulman-Pozen, *Partisan Federalism*, 127 HARV. L. REV. 1077 (2014) (arguing that the agendas of national political parties dominate state policy selection).

[86] *See, e.g.*, Michael Mintrom, *Policy Entrepreneurs and the Diffusion of Innovation*, 41 AM. J. POL. SCI. 738, 765 (1997) (arguing that state-level policy entrepreneurs—"people who seek to initiate dynamic policy change"—"play an important role in articulating innovative ideas onto government agendas").

government.[87] But the idea that federalism will lead state and local governments inexorably toward policy differentiation overstates the case.

Second, federalism may not produce the kinds of differentiation that facilitate learning. True experiments require tailored levels of policy differentiation. If a jurisdiction adopts a program that differs from those of its neighbors in four or five key ways, for example, it can be difficult to isolate the differences that actually matter. Experimentalism also requires carefully timed differentiation, and real-world differentiation may not last very long. Two local governments might adopt different responses to the same policy problem, but if one jurisdiction's approach initially seems to work, the other jurisdiction may change course before anyone has gathered enough data to separate signal from noise.[88] Or, alternatively, two jurisdictions may adopt policies that seem to invite comparison, but may do so at different times. A land use policy might succeed during a booming economy while a slightly differentiated policy fails during a recessionary period a few years later, and observers will have a hard time determining whether it was the differentiation or the recession that changed outcomes.

As that last example illustrates, differentiation problems are deeply intertwined with a federalist system's tendency to produce confounding variables. One of the reasons the federalism-experimentalism connection seems intuitive is that subnational governments can do different things. But that capacity for differentiation, even if sometimes overstated, has led to a history of differentiation, and that history—along with immutable differences of geography and population—means that even adjacent state and local governments tend to differ in many ways. A medical researcher probably would not choose to work in a laboratory that houses fifty different research animals, no two of which are the same; she would want her lab rats as identical as possible. Yet the subnational jurisdictions of the United States are more akin to the former menagerie than the latter controlled environment. And that means that any experiment designed to compare policies in different jurisdictions is likely to confront an abundance of confounding variables.

In one other key way, federalism creates problems with confounding variables. According to most researchers, the best way to control for such

---

[87] *See, e.g.*, Hiroko Tabuchi, *U.S. Climate Change Policy: Made in California*, N.Y. TIMES, Sept. 27, 2017; Bruce R. Huber, *How Did RGGI Do it? Political Economy and Emissions Auctions*, 40 ECOLOGY L.Q. 59, 88-91 (2013) (describing policy entrepreneurial behavior in the creation of the northeast's Regional Greenhouse Gas Initiative).

[88] *See* Abbott & Snidal, *supra* note 39, at 17 (noting the tendency of "local authorities to bunch around promising approaches"); Gary King et al., *A "Politically Robust" Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program*, 26 J. POL'Y ANALYSIS & MGMT. 479, 479-80 (2007) (noting that it is often politically infeasible to maintain a control group where a treatment seems to be succeeding).

variables is through randomization.[89] But randomization requires a high level of control from the experimenter, who randomly assigns subjects to different treatments.[90] And a core purpose of federalism is to limit this type of centralized control, and thus to protect the partial sovereignty of state governments.[91] The differentiation of federalism happens because subnational governments choose it, not through random assignment. Consequently, randomization, despite its promise, will generally be antithetical to federalism.

2. Data Collection and Analysis

Meaningful experimentation requires data collection in addition to controlled differentiation. Yet a federalist structure does not necessarily incentivize states to produce or disseminate useful data. There are several reasons why.

First, in any governance system, the sponsors of a policy may not have much incentive to collect data.[92] They will likely have adopted that policy because they believe it will work, and because they have successfully convinced others that it will work. They therefore may not perceive the need to collect data to validate what they think they already know—or to potentially undercut their own prior claims.[93] Or, somewhat similarly, the only data that interest them may be information about the political marketability of their initiative, not its actual success in achieving policy goals.[94] Additionally, data collection is often expensive.[95] If overall funds for the program are limited, every dollar devoted to data collection and analysis is not spent on other aspects of program implementation. For managers who believe—perhaps correctly—that the program is important

[89] *See* Abramowicz et al., *supra* note 3 (arguing that randomization should be used for more policy analyses).

[90] This statement may sound paradoxical— one might ask how control can be the key to randomness—but the alternative to control by the experiment's designer may be experimental subjects sorting themselves into groups in non-random ways.

[91] *See* Gregory v. Ashcroft, 501 U.S. 452, 458 (1991) (extolling the benefits of divided government).

[92] *See generally* Eric Biber, *The Problem of Environmental Monitoring*, 83 U. COLO. L. REV. 1 (2011) (explaining reasons why monitoring is often absent or ineffective).

[93] *See, e.g.*, THOMAS O. MCGARITY, REINVENTING RATIONALITY: THE ROLE OF REGULATORY ANALYSIS IN THE FEDERAL BUREAUCRACY 137 (1991) (quoting an EPA analyst: "How is my career going to be advanced by doing a study that shows that three years ago the agency made a wrong prediction? It is not in my best interest.").

[94] *See* Livermore, *supra* note 20, at 639 (noting that politicians may pursue experiments for reasons other than helping the public).

[95] *See* Biber, *supra* note 91, at 31.

and already is well-designed, and who therefore expect that every dollar spent on implementation will make their state or city a better place, that opportunity cost may be intolerable.

These problems can arise when a single jurisdiction is conducting an experiment, but they are likely to be even more acute when the goal is to produce results that might benefit other jurisdictions within a federalist system. The problem, again, is one of free riding: a jurisdiction will usually want to gather data only to the extent that it benefits from that data collection, not because it offers some potential benefit to its neighbors.[96] Or it might even think that because its neighbors will collect data on some policy experiment, it can skimp on data collection and just use the information others compile.

Even if jurisdictions within a federalist system are committed to data collection, problems of comparability may arise. To compare multiple jurisdictions, researchers generally need data that address uniform metrics and were collected in consistent ways. Otherwise, any attempt at a comparative analysis risks mixing apples and oranges. But just as a federalist system allows some policy differentiation, it also can facilitate differentiation in more technical matters like data collection practices and management platforms.[97] That differentiation won't necessarily occur; in software selection, as in policy development, imitation can bring efficiencies. But real-world examples show that problems with inconsistent data sets do often arise.[98]

3. Repetition and Variation

Beyond data limits, a federalist system also offers mixed prospects for integrating individual experiments into a larger experimental program. That potential does exist; the possibility of imitation may facilitate re-testing of policy experiments, with the imitators either using the same basic policy or using slightly adjusted approaches. Indeed, a large literature argues that policy diffusion, including diffusion within regions, is common.[99] With

---

[96] *See* Hannah J. Wiseman, *Regulatory Islands*, 89 N.Y.U. L. REV. 1661, 1713-14 (2014) (noting disincentives to gather and share information).

[97] *See, e.g.*, NATIONAL ACADEMY OF PUBLIC ADMINISTRATION, EVALUATING ENVIRONMENTAL PROGRESS: HOW EPA AND THE STATES CAN IMPROVE THE QUALITY OF ENFORCEMENT AND COMPLIANCE INFORMATION 24 (2001) (describing state-to-state differences in data collection practices).

[98] *See, e.g., id.*

[99] *See, e.g.*, Daniel A. Farber, *Carbon Leakage Versus Policy Diffusion: The Perils and Promise of Subglobal Climate Action*, 13 CHI. J. INT'L L. 359, 375-76 (2013); *Note, When Do Policy Innovations Spread? Lessons for Advocates of Lesson-Drawing*, 119 HARV. L.

regional diffusion, in particular, neighboring states might have sufficiently similar characteristics to produce a sort of re-test with each jump of the policy across a state line.[100] Such repetition can help experimenters figure out whether a successful experiment is replicable and also how subtle adjustments to the policy program affect its success.

Nevertheless, federalism can create barriers to programmatic experimentation. As one of us has previously explained, it can be exceedingly difficult for states to obtain information about what other states' policies even are.[101] Particularly in rapidly evolving policy realms—which are precisely the areas where experimentation would theoretically be most valuable—state policies may be dynamic and poorly documented, which makes obtaining information about those policies a time-consuming exercise with quickly-outdated results.[102] To gather information about the consequences of those policies will be even more difficult, if such information even exists.[103] And even when subnational jurisdictions do obtain information about their peers' initiatives, they are likely to access that information through intermediaries, and those intermediaries may have powerful motives to slant the information they provide.[104] Consequently, in the political real world, policy propagation is likely to be haphazard and uninformed, and the spread of policies may be driven more by the ideological narratives surrounding a policy than by any empirical measures of their success.

C. Federalism, Experimentation, and Centralized Authority

To all the problems described above, there is an obvious response. These many design problems can be ameliorated if there is a centralized manager coordinating the experiments. Federalism does offer the possibility of such management. While the political and, sometimes, judicial and

---

REV. 1467 (2006); *supra* note 8 (describing some of the political science literature on diffusion).

[100] *See* Abbott & Snidal, *supra* note 39, at 8, 10 (describing how political boundaries can facilitate comparisons); Berry & Berry, *supra* note 8, at 403 ("We would hypothesize that a state's probability of adopting a lottery increases after one or more states with a reputation as a leader within its region adopt it. This definition is consistent with research that has found that there are states to which the other states in a region look most frequently for innovative ideas.").

[101] *See* Wiseman, *supra* note 96.

[102] *See, e.g.*, *id.* at 1694-1704 (describing how hard it can be for states to obtain information about other states' fracking policies).

[103] *See supra* notes 92-98 and accompanying text (describing disincentives to gather and share monitoring data).

[104] *Id.* at 1715-16.

academic rhetoric of federalism often fixates solely on state empowerment, a strong, if also limited, centralized government is an essential element of the United States' federalist system.[105] That centralized government could play the part of manager. Indeed, this is close to the democratic experimentalism scholars' vision: in their proposed system, federal coordination helps state and local experimental governance succeed.[106] And in state-local relationships, the states could play that same centralized coordinating role.[107]

At a basic level, we agree with that prescription. But identifying at a general level the possibility of centralized coordination is only a start, and the federal government can do much more than just coordinate. For that reason, the next part turns to explaining a typology that captures ways in which federal coordination of experimental systems actually has been done, along with other experimental approaches that fall near or far from the federalized model.

III. MODELS OF POLICY EXPERIMENTATION

Once one moves beyond the assumption that sub-federal control organically fosters experimentation, there are several ways to categorize the many types of experimentation that *do* occur and the many governance levels that serve as their locus. We define two major axes as the foundation of an analytical framework. They are the degree of federal involvement in the experiment (or, conversely, state or local leadership) and the extent to which the experiment incorporates the experimental features identified in Part II. After describing these two primary features, we use case studies to demonstrate experiments that fall at various points within this framework.

The case studies, which come from agricultural, natural resources, and education policy, illustrate two overarching points. First, contrary to the suppositions of traditional federalism theories, the federal government can be the engine of experimentation. Democracy's laboratories need not be exclusively state or local, and federalism theory ought to embrace that possibility. Second, experimental federal policymaking can take place through a rich variety of governance structures and with varying degrees of experimental rigor. We make no claims that federal experimentation always will be best, or about which of these structures is best or what degree of

---

[105] *See* U.S. Term Limits, Inc. v. Thornton, 514 U.S. 779, 838 (1995) (Kennedy, J. concurring) (emphasizing the federal government's importance); Akhil Reed Amar, *Of Sovereignty and Federalism*, 96 YALE L.J. 1425, 1456-58 (1987) (discussing the many constitutional provisions that emphasize the supremacy of the national government).

[106] *See supra* notes 49-62 and accompanying text.

[107] *See* Owen, *Cooperative Subfederalism*, *supra* note 61.

experimental rigor is optimal. The answers to those questions will likely be highly contextual, and will depend on the law governing a policy initiative, the nature of the resources or practices being regulated, the relative competence of different agencies and the resources available to them, and, of course, history and politics. The answers also will change as conditions and policies evolve. But we do claim, and our examples illustrate, that federal experimentation is a promising alternative to a traditional and myopic focus on the states.

A. A Typology of Policy Experiments

There are potentially endless ways to approach policy experimentation. In some circumstances, a disheveled patchwork of uncoordinated states independently throws solutions at problems and sees how they stick.[108] Through happenstance, this type of haphazard effort might produce differing results and potential lessons, but it only remotely resembles an experiment. At the other extreme, the federal government sometimes carefully defines a goal and enlists states and local governments to propose a variety of policy approaches to meeting this goal. The federal government—often through a grant mechanism—selects the states to implement these approaches, requires detailed and uniform data reporting, and prepares reports on the results achieved and their transferability.[109] Alternatively, the federal government may just act alone.[110] We characterize these extremes, and the many gray areas in between, in two primary ways, noting that experiments differ in terms of which levels of government design and implement them (the federalism aspect of policy experimentation) and the extent to which the experiments incorporate key features, such as differentiation and control of confounding variables (the design aspect of policy experimentation).

[108] This type of experiment has evolved in the context of regulating the environmental impacts of oil and gas development. *See* Hannah J. Wiseman, *Risk and Response in Fracturing Policy*, 84 U. COLO. L. REV. 729 (2014) (describing substantial variation in state oil and gas policy); Wiseman, *supra* note 96 (describing varied state policies that lack a common goal or approach and that seem unmoored from any sort of broader experimental approach).

[109] *See infra* notes 245-265 and accompanying text (describing the Race to the Top program).

[110] *See infra* notes 211-244 (describing federal experiments with wildfire).

Figure 1 visually depicts this framework.[111] The letters A through I represent nine approaches to policy experimentation that could fall under this framework.[112]

Figure 1. Policy Experiments and their Governance Locus

Experimental rigor

Federal involvement	Minimal	Moderate	Rigorous
Minimal federal involvement	A	B	C
Federal and state involvement	D	E	F
Minimal state involvement	G	H	I

Much of the traditional federalism rhetoric—especially that espoused by courts—hints, though without much elaboration, at Boxes B and C.[113] The courts seem to assume that a hands-off federal approach—for example, avoiding federal preemption of state policy in a particular area—will cause states to take the experimental reins, applying different types of policies and measuring and honestly reporting the achievements or failures associated with these policies.[114] Experimental rigor has never been a focus of judicial discussion, and thus it is hard to discern whether judges envision action in

[111] By "minimal experimental rigor" we refer to policies that are implemented without efforts to incorporate the design features that we identify in Part II as part of a true policy experiment. Experiments with moderate rigor are implemented with some effort at deliberate differentiation and typically include strong data gathering and evaluation of program results, whereas truly rigorous experimentation would incorporate nearly all of the design features identified and would use actual randomized policy experiments or something close to that.

[112] The rigid boundaries of a simplified matrix are not the real world, of course. Many policy experiments likely fall along the borders of these boxes, and policymakers face a continuum of design and governance choices rather than a discrete set of options.

[113] *See supra* notes 25-38 and accompanying text.

[114] *See, e.g.*, Oregon v. Ice, 555 U.S. 160, 170 (2009).

Box B or C—or have even considered the distinction. But the emphasis on state action is clear.

In contrast to the courts and many federalism scholars, the experimental design literature pays attention to experimental rigor (thus focusing on boxes C, F, and I). But unlike the federalism camps, it often ignores the levels of government at which the experiment plays out.[115] Thus, it does not distinguish among boxes C, F, and I, instead melding them and focusing generally on the degree to which a real policy experiment emerges.

In the real world, most of the action is in box D. It is rare for the federal government to truly stay out of the way of states, even where there is no formal federal preemption of state control.[116] Indeed, even in fields lauded for longstanding state independence, such as land use regulation and education, the federal government commonly induces state and local action through grants and other spending mechanisms and may intervene through regulatory controls.[117] And governments often embark upon these projects without an intentional policy experiment in mind—thus failing to produce an even moderately rigorous experiment—but some policy differentiation does emerge along the way. Sometimes the federal government also intervenes to support consistent data collection and dissemination practices, thus edging state-directed projects closer to the D-E boundary.[118] Meanwhile, real-world practices do sometimes fall within Box A, with neither experimental rigor nor meaningful federal involvement.[119]

Missing from the traditional literature is any significant discussion that explores the marriages of federalism and experimentation within boxes E, F, H, and I. Yet those boxes are not null sets, as shown by the case studies

---

[115] *See supra* notes 39-48 and accompanying text.

[116] *See* Jessica Bulman-Pozen, *From Sovereignty and Process to Administration and Politics: The Afterlife of American Federalism*, 123 YALE L.J. 1920, 1932 (2014) (arguing that integration is the dominant theme of modern American federalism).

[117] *See* Eloise Pasachoff, *Agency Enforcement of Spending Clause Statutes: A Defense of the Funding Cut-Off*, 124 YALE L.J. 248, 260-66 (2014) (describing federal grants in these and other areas). Many federal regulatory interventions into education occur through civil rights laws, and many interventions into land use occur through environmental laws.

[118] *See,e.g.*, The White House, Office of the Press Secretary, The White House, Office of the Press Secretary, FACT SHEET: Data by the People, for the People — Eight Years of Progress Opening Government Data to Spur Innovation, Opportunity, & Economic Growth (September 28, 2016),

https://obamawhitehouse.archives.gov/the-press-office/2016/09/28/fact-sheet-data-people-people-eight-years-progress-opening-government (describing multiple data consistency and transparency initiatives).

[119] For example, as one of us has discussed in the energy law context, the federal government has steered clear of certain aspects of oil and gas regulation, leaving significant regulatory decisions to the states. The states, in turn, have implemented a patchwork of regulatory approaches that have little semblance of an experiment. *See supra* note 108.

in section 2. The federal government can design and implement experiments without much state or local assistance and sometimes has done so.[120] It also has played a coordinating and cooperating role in experimental programs partially staffed by local governments and states.[121]

The following case studies—borrowed from the somewhat disparate fields of agricultural, natural resources, and education policy—show that real-world policy experiments, both old and new, have involved the federal government in these ways. We select these examples because they highlight the many governance levels involved in single policy experiments and varying degrees of rigor in experimentation.

B. Agricultural Soil Conservation: A Rigorous Experiment with Extensive Federal Involvement

One might think that no policy arena is *less* likely to produce a program of federal experimentation than the management of private agricultural land.[122] After all, the Supreme Court has repeatedly emphasized the states' "traditional and primary power over land and water use," and has likewise asserted that "regulation of land use is perhaps the quintessential state activity."[123] Both in judicial proceedings and in the political realm, that rhetoric has real bite. The Court has invoked it as a reason for questioning assertions of federal regulatory jurisdiction,[124] and agricultural interests and

---

[120] *See infra* notes 211-244 and accompanying text (discussing wildfire policy).

[121] *See infra* notes 245-265 and accompanying text (discussing Race to the Top).

[122] Indeed, when President Roosevelt first devoted some Civilian Conservation Corps funds to controlling soil erosion on private lands—a program that eventually ballooned into a massive federal endeavor and has endured through the modern-day U.S. Department of Agriculture—this was a controversial move due to "[c]concern about the public's objections to expenditures of federal funds on private lands." Douglas Helms, *The Civilian Conservation Corps: Demonstrating the Value of Soil Conservation*, J. SOIL & WATER CONS. 184, 184-85 (March-April 1985).

[123] Solid Waste Agency of N. Cook Cty. v. U.S. Army Corps of Eng'rs (SWANCC), 531 U.S. 159, 174 (2001); FERC v. Mississippi, 456 U.S. 742, 768 n. 30 *(*1982). The latter statement is inaccurate; most land use regulation is done by local governments.

[124] *See, e.g.*, Federal Power Commission v. East Ohio Gas Co., 338 U.S. 464, 489 (1950) ("Long before the Federal Government could be stirred to regulate utilities, courageous states took the initiative and almost the whole body of utility practice has resulted from their experiences."); Rapanos v. United States, 547 U.S. 715, 738 (2007) (warning that extending federal authority to "immense stretches of intrastate land . . . . stretches the outer limits of Congress's commerce power and raises difficult questions about the ultimate scope of that power."); Solid Waste Agency of N. Cook Cnty. v. U.S. Army Corps of Eng'rs (SWANCC), 531 U.S. 159, 174 (2001). For a thorough exploration of courts' use of the laboratories argument both to justify experimentation by state courts and state policymakers, *see* Althouse, *supra* note 5, at 1752-75.

their elected supporters have aggressively deployed federalist arguments in their opposition to anything that verges on land use regulation.[125] Yet a rigorous, long-lasting, and far-reaching set of policy experiments resides in the area of agricultural policy—specifically, federal policy designed to reduce the loss of valuable agricultural topsoil through erosion.[126] These experiments began in the 1920s, grew in scope as the Dust Bowl turned soil erosion into a national tragedy, and have continued to the present day.

The discussion that follows describes these experiments. It supports two key points, both of which defy core federalism assumptions. First, this program, though integrated with the state and local structures of traditional federalism, was truly federal at its core. The program first developed through direct contact between federal agency employees and private landowners, and state involvement, when it did take place, occurred within a policy framework established by federal agencies. Second, the program was genuinely experimental. Federal scientists took contrasting theories and generated alternative hypotheses, tested those hypotheses through carefully controlled experiments, recorded and publicized results, and then folded those results back into new experiments and policy change.

1. Federal Agencies and Private Dirt

As historians tell it, the USDA's experiments in agricultural policy began in 1928, when Hugh Hammond Bennett, a visionary soil surveyor employed by the federal Bureau of Chemistry and Soils (a division of the USDA), collaborated with a Forest Service inspector to write a USDA report called "Soil Erosion, A National Menace."[127] This document alerted the public, as

---

[125] *See, e.g.*, Dave Owen, *Urbanization, Water Quality, and the Regulated Landscape*, 82 U. COLO. L. REV. 431, 477 & n. 288-89 (2011) (compiling quotes from legislative hearings).

[126] Much of this policy is "voluntary"; because it involves using money, demonstration projects, and direct provision of supplies such as seeds and other erosion control devices to farmers to incentivize them to improve soil conservation practices, but it is nonetheless policy. Congress approves the funding for the projects and grants, which are distributed throughout the country. Further, not all measures are voluntary. Farmers who do not implement USDA conservation practices are ineligible for a variety of federal funds. *See, e.g.*, *Envtl. Quality Incentives Program*. U.S. DEPT. OF AG., https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/programs/financial/eqip/ (describing farmers selected for funding base specifically on the conservation practices they implement).

[127] U.S. DEPT. OF AG., SOIL AND WATER RESOURCES CONSERVATION: 1980 APPRAISAL: SOIL, WATER, AND RELATED RESOURCES IN THE UNITED STATES 9 (1981) (hereinafter "1980 Appraisal). *See also* H.H. Bennett & W.R. Chapline, *Soil Erosion—A National Menace*, USDA Circular No. 33 (Apr. 1928); Douglas Helms, *Two Centuries of Soil Conservation*, OAH MAGAZINE OF HISTORY 25 (Winter 1991).

well as Congressmen and other federal agencies, to the problem of soil loss caused by improperly tilled and managed soils that blew or washed away.[128] The authors estimated that 1.5 billion tons of soil[129] and at least 126 billion "pounds of plant food material" were lost annually.[130] The federal report was distinctly local in nature, drawing from examples around the United States, documenting and providing pictures of problems in specific regions, such as southwestern Wisconsin,[131] and even counties and individual farms, such as "an apple orchard near Lookout Mountain in northeastern Kansas,"[132] and "one place a few miles south of Troy, Kans."[133] The authors starkly concluded that "[a]n era of land wreckage destined to weigh heavily upon the welfare of the next generation is at hand."[134]

Sadly, they were right. During the 1930s, a combination of drought, wind, and ill-advised policies (many of them federal[135]) that had encouraged cultivation of semi-arid lands devastated much of the Great Plains. Year after year, rains failed, and without the natural sod that had once held soils in place, winds took the topsoil aloft; a single 1934 storm sucked up 350 million tons of soil.[136] "The story of the southern plains in the 1930s," historian Donald Worster has written, "is essentially about dust storms, when the earth ran amok. And not once or twice, but over and over for the better part of a decade: day after day, year after year, of sand rattling against the window, of fine powder caking one's lips, of springtime turned to despair, of poverty eating into self-confidence."[137]

While real-life Tom Joads fled west, and while John Steinbeck and other writers penned their laments,[138] federal soil scientists began to respond. That response would not be easy, for when federal experimentation in soil conservation policy began in earnest, there were "more than 6.5 million farms on about 1 billion [privately-owned] acres" scattered around the United

---

[128] The report focused on water erosion but noted in footnote 1 that "much damage is also done by wind erosion." Bennett & Chapline, *supra* note 127, at 2.

[129] *Id*. at 4. This number comes from the estimate of "a yearly discharge of 500,000,000 tons of suspended material into the sea by rivers, plus twice this amount stranded upon lower slopes and deposited" elsewhere on land or in inland waters. *Id*. See *also* 1980 APPRAISAL, *supra* note 127, at 9 (describing this estimate).

[130] Bennett & Chapline, *supra* note 127, at 2.

[131] *Id*. at 11.

[132] *Id*. at 10.

[133] *Id*.

[134] *Id*. at 22.

[135] *See* 1980 APPRAISAL, *supra* note 127, at 9 (describing the acts that encouraged farmers who had recently immigrated from Europe and were unfamiliar with U.S. landscapes, or how to properly till them, to acquire property at low or even no cost).

[136] DONALD WORSTER, DUST BOWL: THE SOUTHERN PLAINS IN THE 1930S 13 (1979).

[137] *Id*.

[138] *See* JOHN STEINBECK, THE GRAPES OF WRATH (1939).

States.[139] Such a dispersed and localized problem also might seem rather ill-suited for a federal response. Nevertheless, a broad federal program emerged.

USDA began its implementation efforts by funding and establishing "soil experiment stations," which it deployed around the United States.[140] At these stations, federal and state agents planted experimental crops designed to trap soil particles and prevent them from washing away, demonstrated modern plowing and growing techniques that prevented erosion, and implemented other practices to ascertain their effectiveness in the particular region and persuade nearby farmers of their value.[141] Early on in the effort, employees of the Civilian Conservation Corps also directly implemented conservation measures at these experiment stations and on public and private lands, such as helping to plant crops and trees to hold soil.[142]

Congress and the executive branch created new administrative structures to advance the soil management program. First, in 1933, Congress funded and created a new agency within the Department of the Interior called the Soil Erosion Service, with Bennett at its head.[143] The enactment of the Soil Conservation Act of 1935 quickly followed, which renamed Bennett's new agency as the "Soil Conservation Service" and moved it to the U.S. Department of Agriculture. The goal of the service was to "provide permanently for the control and prevention of soil erosion."[144] That would be accomplished through diverse research, demonstration, and other projects; the act directed the Secretary of Agriculture to "conduct surveys, investigations, and research relating to the character of soil erosion and the

---

[139] 1980 APPRAISAL, *supra* note 127, at 9 (providing statistics from the 1930s). As of 1980—when the federal government conducted a comprehensive survey of its agricultural soil conservation program—there were approximately "1.5 billion acres of nonfederally owned" U.S. land, 27 percent of which was devoted to rangeland for livestock, an equal percentage of which was crop land, and 9 percent of which was pasture. *Id*. at 2.

[140] Douglas Helms, *Hugh Hammond Bennett and the Creation of the Soil Conservation Service*, 8 HISTORICAL INSIGHTS at 2 (2008). These were later renamed soil conservation experiment stations. *Id*.

[141] Helms, *supra* note 127, at 25:

> In the new Soil Erosion Service Bennett located soil conservation projects in the watersheds near erosion experiment stations so that directors of the stations could utilize the research in-formation. Farmers in the watersheds signed five-year cooperative agreements to install conservation measures. The Soil Erosion Service furnished equipment, seed, seedlings, [and] assistance in planning the measures, and labor from the Civilian Conservation Corps or the Works Projects Administration.

[142] *Id.* at 2-3.

[143] Helms, *supra* note 140, at 11-12. In 1935 President Roosevelt moved the agency to the USDA and renamed it as the Soil Conservation Service.

[144] Pub. L. 74-76, 59 State. 163, An Act: To Provide for the Protection of Land Resources Against Soil Erosion, and for Other Purposes. *See also* Douglas Helms, *SCS: 50 Years Young*, THE FARMER 48, 48 (Mar. 16, 1985).

preventive measures needed" and to "publish and disseminate" the results of these surveys; to "conduct demonstrational projects in areas subject to erosion by wind or water; and "[t]o carry out preventive measures, including, but not limited to, engineering operations, methods of cultivation, the growing of vegetation, and changes in use of land," among other measures.[145] Those ambitious tasks would be accomplished in part through the network of soil experiment stations that the USDA had already begun to build.

While the federal government led key parts of the effort, it did not act entirely alone. The states, through federally-supported university extension services,[146] state forestry boards,[147] and other divisions of state government, had already begun to study erosion problems and experiment with solutions. Indeed, as early as 1887 the federal government had donated lands to states (creating "land grant" colleges and universities[148]) and funded state agricultural experiment stations, "having due regard to the varying conditions and needs of the respective States or Territories."[149] Based on work at these experiment stations and elsewhere, states had issued reports about erosion and its causes, and some had formed their own agricultural experiment stations at which government officials implemented and assessed the effectiveness of various farming and ranching practices. Bennett's 1928 report on the national soil erosion crisis cited to examples from these state experiment stations. For example, it contrasted two approaches at the "Spur substation of the Texas Agricultural Experiment Station,"[150] and also drew upon reports from similar experiment stations in Missouri and North Carolina.[151] Another key portion of the 1935 Soil Conservation Act set the stage for the federal government to directly collaborate with the states in experimenting and disseminating results, [152] in part because the federal

[145] Pub. L. 74-76 §1.

[146] *See, e.g.*, G.E. MARTIN, TERRACING IN OKLAHOMA, Okla. Agr. Col. Ext. Co. Agt. Work Cit. 218 (1927).

[147] *See, e.g.*, EROSION AND FLOOD PROBLEMS IN CALIFORNIA, Calif. State Bd. Forestry Rpt. to Legislature (1921).

[148] A land grant university is a university built on land donated by the federal government to a state. Many of our nation's most prominent public universities are land grant institutions. *See Land-Grant Colleges and Universities*, U.S. DEPT. OF AG., https://nifa.usda.gov/sites/default/files/resource/lgu_map_6_25_2014_0.pdf.

[149] Hatch Act, 24 Stat. 440, §§ 1-2 (1887).

[150] Bennet & Chapline, *supra* note 127, at 6.

[151] *Id*.

[152] *Id*. § 1(3) (allowing the USDA to "cooperate or enter agreements with, or to furnish financial or other aid to, any agency, governmental or otherwise, or any person . . . for the purposes of this Act").

government recognized that securing effort and commitments from individual farmers would be easier with state and local support.[153]

The federal government solidified the involvement of state and local actors in the soil conservation experiment by providing a model act through which states would enable the creation of local soil conservation districts.[154] This model law, adopted by all states in some form,[155] provided for federal involvement in these districts and enabled close collaboration between the districts and the federal government.[156] Conservation districts continue to operate and exercise these powers. Three thousand districts[157] now emphasize three soil conservation strategies[158] initially championed by Hugh Hammond Bennett, applying and differentiating these practices across "nearly all private rural land" in the United States.[159]

---

[153] As one source explains, "[w]idespread local leadership was required to motivate and guide" the thousands of private landowners operating farms and ranches. CHARLES REAGAN WILSON & WILLIAM FERRIS, EDS., ENCYCLOPEDIA OF SOUTHERN CULTURE 363 (1989).

[154] Under the model act, each state formed a soil conservation district as a "governmental subdivision" of the state. U.S. Dept. Ag., Soil Conservation Serv., A Standard State Soil Conservation Districts Law 3 (1936). The act provided that the state should first establish a soil conservation committee—a state agency—of which a federal USDA representative could also serve as a member. *Id*. at 5. Soil conservation districts could then be formed when "[a]ny twenty-five" individuals occupying land within the area proposed to be a district petitioned the committee requesting district formation. The committee next held a hearing to determine the need for the district and convened a referendum for the formation of the district, in which all "occupiers of land" within the proposed district boundaries could vote. *Id*. at 7-9. After a favorable vote and a determination by the committee that the district could be feasibly administered, the committee appointed two supervisors of the district, and three additional supervisors were elected—again by the "occupiers of land" within the district, thus creating a five-supervisor governing body of the district. *Id.* at 10, 15.

[155] 1980 Appraisal, *supra* note 127, at 14.

[156] These districts—the distinctly local bodies that became important players in implementing diverse federal soil conservation efforts—had extensive powers under the model act, which mirrored the powers granted to the USDA under the federal Soil Conservation Act. They included, for example, conducting "surveys, investigations, or research" relating to soil erosion and its prevention and disseminating the results (but avoiding duplicative research by requiring district coordination with the state or USDA); conducting "demonstrational projects within the district"; directly carrying out erosion prevention and control measures; and obtaining and taking over U.S. and state soil erosion control and conservation projects within its district or acting as an agent of the United States in carrying out these projects, among other powers. A Standard State Soil Conservation Districts Law, *supra* note 154, at 15-17.

[157] *About NACD*, NATL. ASSOC. OF CONSERVATION DISTRICTS, http://www.nacdnet.org/about-nacd/.

[158] *What We Do*, Natl. Assoc. of Conservation Districts, http://www.nacdnet.org/about-nacd/what-we-do/ (noting that the districts focus on "crop rotation, cover crops and no or minimum tillage systems).

[159] 1980 APPRAISAL, *supra* note 127, at 14.

In addition to this reliance on state and local structures, the new federal program continues to take advantage of its own geographic decentralization. The national soil conservation agency—the same agency originally headed by Bennett but now called the Natural Resources Conservation Service[160]—has built from the original strategies emphasized by Bennett and has developed more formal, detailed federal standards for soil conservation (and other conservation practices).[161] But these are not uniform standards. The NRCS has field offices in all fifty states[162] and also operates local service centers,[163] and these offices tailor federal conservation standards to local conditions.[164] Farmers who meet the local conservation standards receive federal funding for implementing soil conservation practices.[165]

Federal soil conservation policy also utilizes another national program called the Cooperative Extension Service, which has been active since

---

[160] The agency's name changed from the Soil Erosion Service to the Soil Conservation Service and then to its current name of Natural Resources Conservation Service. History of NRCA, https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/about/history/.

[161] *Conservation Practices*, Natural Resources Conservation Serv., https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/technical/cp/ncps/?cid=nrcs143_026849

[162] *State Offices Directory*, Natural Resources Conservation Serv., https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/contact/states/. (Employees of the field offices have "USDA.gov" e-mail addresses.)

[163] *Id*.

[164] Each conservation standard is accompanied by a technical guide describing how to implement federal conservation practices, and the guides "used in each field office are localized so that they apply specifically to the geographic area for which they are prepared." Conservation Practices, https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/technical/cp/ncps/?cid=nrcs143_026849.

[165] *Natl. Conservation Practice Standards*, Natural Resources Conservation Serv., https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/technical/cp/ncps/; Envtl. Quality Incentives Program, https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/programs/financial/eqip/ ("States offer a variety of EQIP funding opportunities to address priority local or state resource concerns. Producers may apply at the local NRCS field office where the eligible land to be enrolled is located."). Farmers wishing to receive Environmental Quality Incentive Program Funds must prepare and have approved a Conservation Activity Plan, which are "developed for producers to identify conservation practices needed to address a specific natural resource need." FY 2016 EQIP Conservation Activity Plan (CAP), https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/programs/financial/eqip/?cid=nrcseprd401472. As noted above, the USDA defines national conservation practices that are then localized through field specific manuals. Conservation Practices, https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/technical/cp/ncps/?cid=nrcs143_026849 (noting to farmers that "[y]ou must have the conservation practice standard developed by the state in which you are working to insure that you meet all state and local criteria, which may be more restrictive than national criteria").

1914.[166] Through this service, the federal government provides funding and partners with state and local institutions—most typically land grant universities but also other entities—to form educational "extensions" around the country.[167] These extensions, through federal-state partnerships, conduct localized research[168] and then disseminate it in order to encourage farmers to adopt practices with demonstrated beneficial results.[169] A primary state extension office (staffed with USDA employees) is housed within the land-grant university, and each state also has a "network of local or regional offices."[170] The Extension Service has "an office in or near most of the nation's approximately 3,000 counties."[171]

In summary, the U.S. soil conservation program and broader federal efforts to improve agriculture were both integrally tied to federalism and predominantly federal. Although USDA policy experimentation relies heavily on sub-federal involvement, the federal government has played a major role throughout the long history of soil conservation efforts. Indeed, many of the federal initiatives described above coexist with highly localized government offices, including county offices.[172] And the program's reach is striking. The data from just the first year of the Civilian Conservation Corps's work show the intensely local, diffuse scale of this national effort, with the Corps improving more than "950,000 acres of forest stands," which help to prevent erosion, and building "420,000 erosion control check-dams" and 4,000 miles of fence.[173] Later, in assessing results achieved by the 1935 Soil Conservation Act, the USDA and its state partners surveyed the owners of "67 percent of all private land" in the United States, contacted more than 37,000 "individuals, partnerships, and corporations" and examined approximately 200,000 soil samples taken by a state program using federal

---

[166] *Cooperative Extension History*, U.S. Dept. of Ag., Natl. Inst. Of Food & Ag., https://nifa.usda.gov/cooperative-extension-history.

[167] U.S. Dept. of Ag., Natl. Inst. Of Food & Ag., Extension, https://nifa.usda.gov/extension; Renewable Resources Extension Act, Sustaining the Nation's Forest and Rangeland Resources for Future Generations 2-7 (2016), https://nifa.usda.gov/sites/default/files/resource/RREA_Strategic_Plan_2012_2016.pdf (describing federal-state-local partnerships involved in a USDA program designed to support sustainable rangeland and grassland).

[168] This research is not limited to soil conservation. Extension services also address issues such as developing plant strains tailored to certain climates and hydrologic conditions and improving pest management.

[169] *How We Work*, Natl. Institute of Food and Ag., https://nifa.usda.gov/how-we-work.

[170] USDA Local Offices, U.S. Dept. of Ag., https://www.outreach.usda.gov/USDALocalOffices.htm.

[171] *Cooperative Extension History*, *supra* note 166.

[172] *USDA Local Offices*, U.S. Dept. Ag., https://www.outreach.usda.gov/USDALocalOffices.htm.

[173] 1980 APPRAISAL, *supra* note 127, at 11.

funding.[174] The agency also held "about 9,000 public participation meetings, attended by over 164,000 persons."[175] The long history of federal involvement in agricultural policy experimentation shows that the federal government can operate, independently and in concert with sub-federal entities, at a distinctly localized scale.

2. Soil Conservation and Experimentation

In addition to including substantial federal participation, many of the USDA's soil conservation initiatives were genuinely experimental. The USDA took a body of theory, developed competing hypotheses, tested them, repeated similar tests in other places, and collected mountains of data. Those data supported an outpouring of written studies, which in turn informed additional policy development and experimentation.

a. Hypotheses and Policy Differentiation

At the time Bennett wrote his report, there were competing hypotheses about how soil conservation policy should be handled. Another agency competing for the funds—the Bureau of Agricultural Engineering—preferred one national, uniform plan that solely implemented terracing strategies around the United States[176] Bennett, however, was firmly committed to a menu of practices that would differ by region—practices that could be tested through experimentation and then expanded to farms that operated under similar conditions.[177] Congress chose the latter approach. Consequently, when the USDA began its national soil conservation experiment in the late 1920s, it was acutely aware of the need to vary conservation practices to match them with unique soils, rainfall levels, and other conditions around the country. The federal government therefore decided to test this range of hypotheses, rather than a singular approach, through its experiment stations and demonstration projects across the United States.

Federal agencies did this testing through a deliberate program of differentiation. In part, this differentiation flowed from the geography of the new geographic structures; a key point of locating the experiment stations and then the soil conservation district offices around the country was to allow implementation of different policies in different places. For example, in desert areas the Corps collected native seeds that would best grow in an arid

[174] *Id*. at 1, 35 (describing the Iowa State-led National Resource Inventories).
[175] *Id*. at 5.
[176] Helms, *supra* note 140, at 9-10.
[177] *See infra* notes 178-180.

climate and would help to stabilize grazing areas.[178] And the Soil Conservation Service demonstration projects in the Pacific Northwest promoted winter cover crops because of the heavy winter rainfall in this area.[179] Indeed, the overall approach at this time, according to USDA historians, was to use the experiment stations and projects to develop a conservation system "tailored to the individual farm."[180]

In addition to varying soil conservation approaches among localities, USDA also varied its approaches within localities, testing alternative hypotheses and then emphasizing successful approaches. For example, at the San Dimas watershed experiment station in California, Civilian Conservation Corps workers tested the hypothesis that keeping vegetative cover on fields to trap water was important to replenish groundwater supplies—and specifically, that the water-trapping benefits were more important than the soil health benefits achieved from burning crops.[181] CCC workers built special structures to capture surface water and allow it to percolate through soil to groundwater, thus helping to measure the importance of cover.[182] Similarly, Walter Lowdermilk, a prominent agency scientist, completed several studies designed to isolate the factors that contributed to erosion in different types of watersheds and to transfer this knowledge into policy—providing what he described as a "basis for enlightened management."[183] Similar studies at other research stations tested hypotheses regarding how to best measure soil runoff and erosion (and thus how to best measure the effectiveness of soil conservation policies).[184]

b. Control of Confounding Variables

As the previous examples suggest, USDA scientists thought carefully about using structured differentiation to produce meaningful experimental results. Often, that care included setting up control groups. At the original ten research stations established by the federal government, for example, staff

[178] Helms, *supra* note 127, at 3.

[179] Helms, *supra* note 122, at 187.

[180] Helms, *supra* note 144, at 48. In other circumstances, the Corps applied national strategies deemed to be effective in most regions, such as fencing cattle grazing areas so that cattle would not trample and disturb extensive soil areas. 1980 APPRAISAL, *supra* note 127, at 11.

[181] Douglas Helms, *Walter Lowdermilk's Journey: Forester to Land Conservationist*, 8 ENVTL. REV. at 4 (summer 1984).

[182] *Id*.

[183] *Id.* (quoting Walter C. Lowdermilk, *Further Studies of Factors Affecting Surficial RunOff and Erosion*, Proc. Intl. Congr. of Forestry Experiment Stations 625 (Stockholm 1929).

[184] *Id.* at 5.

deliberately created "control plots"[185] to account for variables other than different soil conservation practices that might affect erosion.[186] Similarly, Lowdermilk pioneered techniques to control the many factors that confound efforts to measure the causes of soil erosion at watershed[187] scales. At the experiment stations, and in studies preceding these stations, Lowdermilk accordingly endeavored to "isolate various factors at work" in soil erosion and "measure their influences separately" in individualized plots with carefully-controlled variables.[188] The specialized structures built by CCC workers in the San Dimas watershed were also part of an effort to demonstrate the independent effects of certain crops and other techniques. Similar control techniques to isolate variables and identify the most effective erosion control practices were deployed at other experiment stations.[189]

c. Data Collection, Analysis, and Documentation

The most rigorous experimental aspects of federal soil conservation involve data collection, analysis, and documentation. Relatively early on, executive and congressional directives required the USDA to try to determine whether its experiments were working—in other words, to collect and analyze data on the effectiveness of soil conservation practices.[190] Staff at

[185] 1980 APPRAISAL, *supra* note 127, at 10.

[186] Lowdermilk was inspired in part by carefully-controlled experiments that he observed at forestry experiment stations abroad. These stations involved carefully controlled "sample plots" with measurements of different treatments, such as different amounts of tree thinning, and various results from these treatments, such as sizes of trees and soil composition. Swedish Institute of Experimental Forestry, Guide to the Excursions of the International Congress of Forestry Experimental Stations, Tables and Figures at 1-19 (1929); Readings in the History, *supra* note 122, at 38 (noting Lowdermilk's participation in the Stockholm meeting).

[187] A watershed is an area in which all surface water runoff flows into one water source, such as a stream, river, or lake. Measuring or modeling soil erosion at this scale can be complex, particularly if the watershed contains a variety of soil and land cover types. *See* Readings in the History, *supra* note 122, at 38 ("In an open [watershed] setting there were too many variables.").

[188] Walter C. Lowdermilk, *Studies in the Role of Forest Vegetation in Surficial Run-Off and Soil Erosion*, 12 AG. ENGINEERING 107, 108 (1931).

[189] Helms, *supra* note 183, at 4.

[190] Over time, just as the focus of the National Resources Conservation Service has expanded to issues beyond soil conservation, so, too, has the focus of data collection and analysis. However, we focus on the soil conservation aspect of the effort because of its long history. For a discussion of the expansion of USDA *The Conservation Effects Assessment Project benchmark watersheds: Synthesis of preliminary findings* attention to conservation issues beyond the preservation of soil, *seem e.g.*, C.W. Richardson et al.,63 J. SOIL & WATER CONSERVATION 590, 590 (2008).

federal and federal-state experiment stations embarked upon careful data collection and reporting efforts, describing experimental techniques and results in a series of detailed reports published both through federal agencies[191] and in peer-reviewed science journals.[192]

Widespread data collection and analysis in this area continue to the present day. USDA uses its wide network of state and local offices to help with the massive task of determining how and where soil is eroding, and why, and plug these data into models. For example, local Natural Resource Conservation Field Offices help to identify the farms that participate in USDA conservation programs and the specific conservation practices used on individual farms.[193] Independently and in collaboration with state and local governments, the Soil Conservation Service has completed hundreds of detailed studies of reduced soil erosion.[194] The results from these studies provide quantitative support for effective soil conservation practices[195] and

---

[191] *See, e.g.*, USDA, Bureau of Public Roads, Second Progress Report on Erosion and Run-Off Experiments in Piedmont, North Carolina (1929) (written "in cooperation with the North Carolina Department of Agriculture).

[192] *See, e.g.*, J. Cho. et al., *Water quality effects of simulated conservation practice scenarios in the Little River Experimental watershed*, 65 J. SOIL & WATER CONS. 463 (Nov./Dec. 2010); G.W. Feyereisen et al., *Long-term stream chemistry trends in the southern Georgia Little River Experimental Watershed*, 63 J. SOIL & WATER CONS. 475 (Nov./Dec. 2008); R.A. Kuhnle et al., Conservation practice effects on sediment load in the Goodwin Creek Experimental Watershed, 63 J. SOIL & WATER CONS. 496 (Nov./Dec. 2008); R.N. Lerch et al., Overview of the Mark Twain Lake/Salt River Basin Conservation Effects Assessment Project, 63 J. SOIL & WATER CONS. 345 (Nov./Dec. 2008); G.W. McCarty et al., *Water quality and conservation practice effects in the Choptank River watershed*, 63 J. SOIL & WATER CONS. 461 (Nov./Dec. 2008); R.D. Harmel et al., *Conservation Effects Assessment Project research in the Leon River and Riesel watersheds*, 63 J. SOIL & WATER CONS. 453 (Nov./Dec. 2008); R.B. Bryant et al., *Cannonsville Reservoir and Town Brook watersheds: Documenting conservation efforts to protect New York City's drinking water*, 63 J. SOIL & WATER CONS. 339 (Nov./Dec. 2008); Douglas L. Karlen et al., *Is No-Tillage Enough? A Field-Scale Watershed Assessment*, 7 ELECTRONIC J. INTEGRATIVE BIOSCI. 1, (May 2, 2009); A. Simon & L. Klimetz, *Relative magnitudes and sources of sediment in benchmark watersheds of the Conservation Effects Assessment Project*, 63 J. SOIL & WATER CONS. 504 (Nov./Dec. 2008).

[193] Maurice J. Mausbach & Allen R. Dedrick, *The Length We Go: Measuring Environmental Benefits of Conservation Practices*, J. SOIL AND WATER CONSERV. 96A, 100A (2004).

[194] *See, e.g.*, Douglas L. Karlen et al., *Is No-Tillage Enough? A Field-Scale Watershed Assessment of Conservation Effects*, 7 ELECTRONIC J. OF INTEGRATIVE BIOSCI. 1 (2009); R.F. Cullum et al., *Effects of Conservation Reserve Program on Runoff and Lake Water Quality in an Oxbow Lake Watershed*, 5 J. INT. ENVTL. APPLICATION AND SCI. 318 (2010); R.A. Kuhnle et al., *Conservation practice effects on sediment load in the Goodwin Creek Experimental Watershed*, 63 J. SOIL & WATER CONSERV. 496 (Nov./Dec. 2008).

[195] In 1989 USDA studies triggered by an executive initiative began to "quantify environmental effects of conservation practices at the field scale." Richardson et al., *supra* note 190, at 590. And in 2002, when Congress significantly increased funding for agricultural

inform and improve detailed and frequently-updated national models, which provide broader predictions about the effectiveness of soil conservation practices where a particular watershed has not been studied.[196]

The data collected as part of this effort are extensive, to say the least. For example, to assess the effects of USDA conservation practices on croplands, USDA now relies on samples from 71,000 to 72,000 different land segments—a subset of a larger inventory of samples originally collected by "thousands" of Soil Conservation Service agents.[197] USDA staff also rely on data from USDA field offices to obtain specific information about crops and "unique landscape features," among other data not included in the samples,[198] and they conduct farmer surveys and other "one-time special studies" to supplement the sample information.[199] It is rare for agencies to be this committed to meaningful data collection and analysis of the policy tools that they implement, but the USDA is somewhat uniquely positioned in this endeavor. The ARS, which employs more than 2,000 scientists,[200] collects

---

conservation practices, Congress also mandated a study with similar quantification goals, *id.* at 590, leading to the creation of the USDA's Conservation Effects Assessment Project. *Id*. at 591. The CEAP, implemented through the USDA's Agricultural Research Service (ARS) and cooperative extensions, *id*., will provide "in-depth quantification of water quality and soil quality impacts of conservation practices at the local level," such as "tons of soil saved." Lisa F. Duriancik et al., *The first five years of the Conservation Effects Assessment Project*, 63 J. SOIL AND WATER CONSERVATION 185A, 18A (Nov/Dec 2008); Mausbach & Dedrick, *supra* note 193 at 97A.The CEAP also aims to improve and validate existing models and help "[d]evelop policy-planning tools to aid selection and placement of conservation practices for optimal environmental quality," among other goals. Richardson et al., *supra* note 190, at 591. Thus, an express function of this extensive data-driven effort is to improve soil conservation policy. Fourteen watersheds have been identified as "benchmark watersheds" for the CEAP; in many of these watersheds the ARS already had begun the work of quantifying soil erosion and practices designed to limit it, and thus these areas serve as "benchmarks" or comparison points for how to best measure the effects of conservation practices in other watersheds. *Id.* at 591.

[196] *See, e.g.*, P.W. Gassman et al., *The Soil and Water Assessment Tool: Historical Development, Applications, and Future Research Directions*, 50 AM. SOC'Y OF AG. AND BIOLOGICAL ENGINEERS 1211 (2007).

[197] UNITED STATES DEPARTMENT OF AGRICULTURE, NATURAL RESOURCES CONSERVATION SERVICE, 2012 NATURAL RESOURCES INVENTORY SUMMARY REPORT 7-2 (2015), https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcseprd396218.pdf. The USDA used existing NRI soil sample points but had to survey farmers to get more information about these samples and the conservation practices at the farms from which the samples were taken. Cropland National Assessment, USDA, https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/ceap/na/?cid=nrcs143_014144.

[198] NATURAL RESOURCES CONSERVATION SERVICE, *supra* note 197, at 7-2 to 7-3.

[199] *Id*. at 1-2.

[200] United States Department of Agriculture, Agricultural Research Service, https://www.ars.usda.gov/about-ars/ (last visited December 21, 2017).

detailed data on the impacts of various erosion control practices and other conservation practices incentivized through federal programs, and it has done so for quite a long time. [201]

The data also do not just sit on shelves. Many of the scientific reports describing the controls used and results collected in USDA studies conducted in watersheds around the United States end with recommendations for applying these practices elsewhere or observations about the specific effectiveness of specific conservation approaches.[202] And the USDA incorporates these lessons learned directly into policy, modifying its guidance provided to farmers who wish to receive federal funding by implementing USDA-approved conservation measures.[203]

d. Repetition and Variation of Experiments

In carrying out its soil conservation experiments, the federal government has also made a concerted effort to test and re-test hypotheses.[204] It has

[201] Since the 1980s, scientists within the service have developed and applied complex models that attempt to measure the amount of soil and other substances eroding from farms and ranches and entering waters within a particular region. Gassman et al., *supra* note 196, at 1212. These models allow different conservation practices to be plugged in and "simulated," thus demonstrating the likely effectiveness of the practices in particular regions. Gassman et al., *supra* note 196, at 1213. A 2007 survey of scientists' uses of and improvements to the models, both within the United States and internationally, lists more than 115 scientific papers, with several of the papers using the model in multiple U.S. watersheds. *Id.* at 1217-1224. A widely-used ARS model—the Soil and Water Assessment Tool—has been specifically applied to "assess the benefits of [USDA] conservation practices" at the national level and "for watersheds of varying sizes that are representative of different regional conditions and mixes of conservation practices." *Id*. at 1215.

[202] *See, e.g.*, Kuhnle et al., *supra* note 194, at 502 (noting that in watersheds subject to a particular type of erosion ("channel erosion"), both the amount of loose soil on the surface (due to a lack of crop cover, for example), and water carrying that soil through runoff contribute to erosion, and that these factors must be considered in designing conservation practices); Cullum et al, *supra* note 194. at 325 ("All physical and chemical water quality data from the runoff from these drainage ditches provided support for the hypothesis that improvement in edge-of-field water quality can be demonstrated via land placed in the Conservation Reserve Program.").

[203] *See, e.g.*, Richardson at al., *supra* note 190, at 590 (noting that the Conservation Effects Assessment Project provides data that inform programs like the Conservation Reserve Program, through which the USDA pays farmers to, for example, plant trees rather than corn, and the Environmental Quality Incentives Program, in which farmers follow USDA guidance in implementing conservation measures and receive federal funding for this implementation).

[204] *See, e.g.*, Karlen et al., *supra* note 194 (noting that "although previous studies had suggested that reduced tillage and extended cropping systems would be more sustainable

followed this practice since the formation of the first experiment stations, when government scientists tested results on farms near the research stations.[205] In addition to retesting hypotheses, the USDA uses experiments throughout the United States to validate and refine its models.[206]

* * * * *

In summary, a birds-eye view of the federal effort to encourage sound soil conservation policies reveals a surprisingly rigorous and highly localized set of experiments, in which the federal government has consistently been a central player. And while skeptics might argue that this is merely an unusual subject area—one that has more to do with uncontroversial physical science than with policy, and where federal experimenting therefore would come abnormally easily—that characterization would be wrong. Here, physical science and policy were closely linked. The ultimate goal was not just to let farmers know how to keep dirt on their property, but also to inform legal decisions on the criteria for allocating massive sums of federal money.[207] Policy, in other words, was centrally at issue. Additionally, this sort of linkage between science and policy is hardly atypical. In fields as diverse as education policy,[208] financial regulation,[209] and endangered species protection,[210] policy decisions grounded in technical research, and often in science, are routine. The real lesson here is that a robust, federally directed program of policy experimentation can exist where one might least expect it.

---

than the continuous corn grown on the site since the early 1960s, quantitative evidence was lacking").

[205] Gassman et al., *supra* note 196, at 1215.

[206] Richardson et al., *supra* note 190, at 591.

[207] *See supra* note 195 and accompanying text.

[208] *See, e.g.*, Center for Education Policy Research, Harvard University, https://cepr.harvard.edu/ (last visited November 14, 2017) (describing research initiatives). The idea of relying on empirical research to inform important education policy decisions is not new. *See, e.g.*, Brown v. Bd. of Educ. of Topeka, 347 U.S. 483, 494 n.11 (1954) (turning to education research to support the Court's finding that separate educational facilities are inherently unequal).

[209] *See, e.g.*, Board of Governors of the Federal Reserve System, Economic Research, https://www.federalreserve.gov/econres.htm (last visited November 14, 2017) (providing links to policy-relevant economic research and describing the Fed's research support efforts); U.S. Securities and Exchange Commission, About the Division of Economic and Risk Analysis, https://www.sec.gov/dera/about (last visited November 14, 2017).

[210] *See* Holly Doremus, *The Purposes, Effects, and Future of the Endangered Species Act's Best Available Science Mandate*, 34 ENVTL. L. 397, 399 (2004) (noting the importance of science to endangered species policy).

C. "Playing with Fire"[211]

The federal program of soil experimentation upends some of the assumptions of the traditional federalism literature, and it was not a unique outlier. Over the course of the twentieth century, American wildfire policy underwent a dramatic transformation.[212] That transformation ultimately involved both Congressional policymaking and state actions, but the primary driver of change was a program of deliberate policy experimentation within the federal bureaucracy. Indeed, in the terms of our matrix in section III.A, this was almost a box H/I experiment. Experimental rigor was high, and the program was even more federally-centered than the USDA's work with soils.

For most of the twentieth century, American wildfire policy was dominated by a single, clear idea: put it out, and quickly.[213] Bolstered by Congressional funding, a management ethos partially forged in the wet forests of Europe, a desire to preserve trees for harvesting rather than combustion,[214] and, eventually, surplus military gear, and cheered on by Smokey the Bear, federal agencies launched an all-out effort to suppress every fire that appeared on the public lands.[215] By the 1930s, that ambition hardened into a simple rule: every fire should be out by 10:00 AM on the day after it was spotted.[216] The United States Forest Service led the crusade against wildfire, but other federal agencies followed suit, as did many states.[217] Only in the southeast did widespread tolerance for both wildfire and prescribed burning remain.[218]

---

[211] John McLaughlin, *Restoring Fire to the Environment in Sequoia and Kings Canyon National Parks*, 12 PROC. OF THE TALL TIMBERS FIRE ECOLOGY CONFERENCE 391, 394 (1973) ("I suppose one could say we are playing with fire….").

[212] *See generally* Stephen J. Pyne, *Between Two Fires: The Past and Future of Fire in America*, 18 PENN ST. ENVTL. L. REV. 129 (2010) (providing an overview of this history).

[213] See Jan W. van Wagtendonk, *The Evolution of National Park Service Fire Policy*, 52 FIRE MGMT. NOTES 10, 10-11 (1991).

[214] This desire was not shared by the National Park Service, but it was central to Forest Service culture.

[215] *See* Scott L. Stephens & Neil G. Sugihara, *Fire Management and Policy Since European Settlement*, in FIRE IN CALIFORNIA'S ECOSYSTEMS 431, 433-34 (Neil G. Sugihara ed. 2006); George Busenberg, *Wildfire Management in the United States: The Evolution of a Policy Failure*, 21 REV. POL'Y RESEARCH 145, 148-52 (2004) (describing this evolution).

[216] Stephens & Sugihara, *supra* note 215, at 434.

[217] *See* Jim Brenner & Dale Wade, *Florida's Revised Prescribed Fire Law: Protection for Responsible Burners*, in PROCEEDINGS OF FIRE CONFERENCE 2000: THE FIRST NATIONAL CONGRESS ON FIRE ECOLOGY, PREVENTION, AND MANAGEMENT 132, 132 (K.E.M. Galley et al., eds. 1992) ("During the early part of the 20th Century, the use of fire as a management tool by state and federal agencies was seen as anathema.").

[218] *Id.* at 133 ("Florida has led the nation in acreage treated with prescribed fire every year since records have been kept.").

From the outset, however, some forest managers expressed doubt about this policy, and by the 1960s, the doubts were beginning to swell. [219] Particularly at the University of California, Berkeley and at San Jose State University, forest scientists began resurrecting the old view that fire was a natural and desirable part of many ecosystems.[220] It could not be suppressed forever, they argued, for fuel buildup would only increase, and the fires that eventually did burn would be enormous, dangerous, and ecologically catastrophic.[221] Their arguments found a sympathetic audience in Starker Leopold, a prominent ecologist who, in 1963, wrote a report for the Department of the Interior recommending more ecology-based management of public lands.[222] But turning fires loose on lands where they had been suppressed for decades—and suppressed with almost religious zeal—was not the kind of radical policy change that could happen all at once. Instead, scientists and federal land managers needed sites for experimentation.

They found those sites scattered around the west.[223] In Sequoia, Kings Canyon, and Yosemite National Parks, managers began experimenting with controlled burns.[224] The Forest Service's attitude toward this new approach was more tepid, but managers in the Selway-Bitterroot and Gila National Forests also conducted prescribed burns.[225] And while authors often throw the word "experiment" around somewhat indiscriminately, in this context it is accurate. Forest managers took a growing body of theory, developed

---

[219] *See* Stephens & Sugihara, *supra* note 215, at 433 (describing opposition in the 1880s); Stephen J. Pyne, *Resistance, Restoration, Resilience: A Survey of Fire's American Century*, 48 ARIZ. ST. L.J. 53, 55 (2016) ("One side argued that fire suppression was misguided and that the proper strategy was to 'light burn' the montane forests as the Indians had."); Jan W. van Wagtendonk, *The History and Evolution of Wildland Fire Use*, 3 FIRE ECOLOGY 3, 4-5 (2007) (describing skepticism from park managers even during the height of "the fire protection years").

[220] *See* van Wagtendonk, *supra* note 213, at 11-12.

[221] *See* Scott L. Stephens & Lawrence W. Ruth, *Federal Forest-Fire Policy in the United States*, 15 ECOLOGICAL APPLICATIONS 532, 533 (2005) (describing the consequences of, and growing skepticism about, fire suppression).

[222] *Id*. at 12; A. STARKER LEOPOLD ET AL., WILDLIFE MANAGEMENT IN THE NATIONAL PARKS (1963).

[223] In the early 1950s, the Park Service had experimented with controlled burns in the Everglades. *See* Bruce M. Kilgore, *Fire Management in the National Parks: An Overview*, 14 PROC. OF THE TALL TIMBERS FIRE ECOLOGY CONFERENCE 45, 46 (1976). But until the early 1970s, the Everglades experiment remained exceptional. *Id.*

[224] *See* Carol Miller, *The Contribution of Natural Fire Management to Wilderness Fire Science*, 20 INT'L J. WILDERNESS 20, 20-21 (2014); Peter H. Schuft, *A Prescribed Burning Program for Sequoia and Kings Canyon National Parks*, 12 PROCEEDINGS OF THE TALL TIMBERS FIRE ECOLOGY CONFERENCE 377 (1973); *see U.C. Foresters Aid Fire Ecology Program at Yosemite Park*, CAL. AG., Feb., 1971, at 3, ucanr.edu/repositoryfiles/ca2502p3-63835.pdf.

[225] Stephens & Ruth, *supra* note 221, at 533.

hypotheses from it, and tested those hypotheses through carefully observed trials.[226] Rather than just running their experiments once, they repeated them at other sites.[227] And because adjacent forest did not burn, and there were many areas where fire suppression remained the dominant policy, there was no shortage of control areas with which scientists could draw comparisons. The resulting burns generated an abundance of data, which federal land managers and external scientists used to generate a huge outpouring of written research studies.[228] Land managers also began hosting annual conferences to discuss their findings.[229] And ideas spread. A let-it-burn strategy went from being the hypothesis for a few isolated experiments on a few acres in a few national forests and parks to a key part of national fire policies.[230]

That transition was not always smooth. Public outrage after 1988's enormous fires in Yellowstone National Park, and after a prescribed burn ran out of control and rampaged through Los Alamos, New Mexico, undercut, for a while, the agencies' embrace of fire.[231] And many fire scientists agree that the embrace—particularly by the Forest Service—still is not nearly enthusiastic enough.[232] Across the American West, forests still contain a dangerous overabundance of fuel, and that is partly a result of continued fire suppression, which in turn arises from a complex set of incentives that reward land managers more for reactive firefighting than for preventive management.[233] Nevertheless, real changes have taken place, even if they have not gone as far as they should. And the process does bear nearly all the hallmarks of a deliberate program of policy experimentation. Double-blind, randomized controlled trials did not happen, but the program included hypotheses grounded in theory, deliberate policy differentiation, repeated

---

[226] *See* Kilgore, *supra* note 223, at 51 (describing the importance of "carefully controlled laboratory and field studies" carried out in conjunction with university and Forest Service scientists).

[227] *See id.* at (4-5) (describing burning experiments in a variety of different ecological settings).

[228] For a compilation of papers, see National Park Service, Sequoia and Kings Canyon National Parks, Fire Research Papers, https://www.nps.gov/seki/learn/nature/fic_fireres.htm (last visited October 2, 2017).

[229] Many of the papers cited here come from the proceedings of the Tall Timbers Fire Ecology Conference.

[230] *See* Stephens & Ruth, *supra* note 221, at 534 (describing changes to federal fire policies); Kilgore, *supra* note 223, at 47-48 (describing the spreading of the Park Service's let-it-burn policy).

[231] *See* van Wagtendonk, *supra* note 213, at 8, 11.

[232] *See, e.g.*, Scott L. Stephens et al., *U.S. Federal Fire and Forest Policy: Emphasizing Resilience in Dry Forests*, 7 ECOSPHERE 1, 2 (2016), http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1584/full.

[233] *See id.* at 6-7.

testing, observation and data collection, and analysis and dissemination of outcomes, and policy changed as a result.

Another feature of the story of fire experimentation merits emphasis: most experiments took place on, and the vast majority of the research studies emerged from, the federal lands.[234] Federal lands also were the initial focus of policy reform.[235] And the reforms came from agencies, not Congress or the White House. While elected officials did eventually get in on the policy reforms, they waited until 2003, over three decades after the federal bureaucracy's experimental program began.[236] That does not mean the federal government acted entirely independently; during this same period, there were some state experiments with controlled burns. Florida continued to be a leader, and California set prescribed burns in several of its state parks.[237] Some of the scientists who helped instigate the process also worked at state universities.[238] But many state policies, particularly in the West, still reflect the old suppress-everything ethos, as do private land management strategies.[239]

---

[234] *See, e.g.*, National Park Service, *supra* note 228. The compilation of papers at this site is just a partial compendium of the large literature on federal forest fire management, yet it still contains many more papers than we have been able to find about state law. *See also* Lauren Wishnie, *Fire and Federalism*, 17 N.Y.U. ENVTL. L.J. 1006, 1015 (2008) (noting "the preeminent position of the federal agencies" in wildfire policy, though also discussing state involvement).

[235] *See* Stephens & Ruth, *supra* note 221. States also have become increasingly involved in fire planning, particularly for "wildland-urban interface" areas where dispersed rural settlement brings people into fire-prone areas.

[236] *See* Robert B. Keiter, *The Law of Fire: Reshaping Public Land Policy in an Era of Ecology and Litigation*, 36 ENVTL. L. 301, 312-13 (2006) ("Remarkably, the [Healthy Forests Restoration Act of 2013) represents the first significant federal legislation on the role and management of fire on the public lands.").

[237] W. James Barry & R. Wayne Harrison, *Prescribed Burning in the California State Park System*, in PROCEEDINGS OF THE SYMPOSIUM ON FIRE IN CALIFORNIA ECOSYSTEMS: INTEGRATING ECOLOGY, PREVENTION AND MANAGEMENT 203, 204-06 (1997). Interestingly, the state park prescribed burning program was encouraged by Harold Biswell, a UC Berkeley scientist who had also worked with the National Park Service, and National Park Service scientists helped train state agency staff. *Id.* at 206.

[238] *See* Kilgore, *supra* note 223, at 46.

[239] *See* Karen Bradshaw, *A Modern Overview of Wildfire Law*, 21 FORDHAM ENVTL. L. REV. 445, 453 & n. 30 (2010) (describing differences between federal and state policies); Michelle Steen-Adams et al., *Historical Perspective on the Influence of Wildfire Policy, Law, and Informal Institutions on Management and Forest Resilience in a Multiownership, Frequent-Fire, Coupled Human and Natural System in Oregon, USA*, 22 ECOLOGY & SOC'Y __, __ (2017; pagination forthcoming) (finding, in a comparative study of public and private forest management in Oregon, that "[i]n general, private owners have responded to increased wildfire hazard in frequent-fire forests with comparatively limited adaptation… yet have not shifted away from the past practices of stages I–II that historically contributed to fuel

The reliance on federal lands for this experiment provides one obvious explanation for federal agency leadership in this area. Historical contingency also played a role; by the time the program of experimentation began, the federal government had long since established its leadership in the field of forest fire management—with increasingly poor results.[240] But those explanations, though accurate, are incomplete. Private forests (which states regulate) and state-owned forests are also abundant and sometimes are more fire-prone than nearby federal lands.[241] State and private owners therefore had incentives to undertake similar reforms, even though most did not do so.[242] Similarly, Congress, which funds a shockingly expensive fire suppression program, had ample reason to pursue innovation, but did not.[243] That suggests that some distinctive advantages of the federal bureaucracy—specifically, its combination of decentralized local management units and a national coordinating superstructure—made it particularly well suited for carrying out an experimental program. As one National Park Service scientist later explained, "NPS fire policies have evolved in a pattern of leaps forward followed by experimentation and refinement. The decentralized nature of the agency allows it to take advantage of new philosophical ideas and translate them into policy."[244]

The story of playing with fire thus underscores some of the same basic points as our discussion of soil management. The first is that while state and local governments may sometimes be laboratories of democracy, they are not the only laboratories. For some policy arenas, federal agencies may be much more promising crucibles of experimental reform. Second, while federal experimental programs can be carried out in close coordination with the states, as was the case with key parts of the USDA's soil management program, the federal government also can act fairly independently. In other words, our laboratories of democracy may be intertwined with, or largely outside of, the structures of federalism.

---

accumulation."); *id* at __ (describing how Oregon state law emphasizes complete fire suppression and slowed the pace of federal reform).

[240] *See* Busenberg, *supra* note 215 (describing this federal role and its consequences).

[241] *See, e.g.*, Steen-Adams et al., *supra* note 239, at 36 (finding, in a study of Oregon forests, that "private corporate forests are more vulnerable to wildfire than are other large ownerships because of the comparatively high proportion of frequent-fire forest").

[242] *See* Amanda Hemmerich, *From Fire Comes Life: Why Courts Assessing Forest Fire Damages Should Recognize Ecological Benefits*, 46 ENVTL. L. REP. NEWS & ANALYSIS 10608, 10608 (2016) (describing a costly settlements of litigation over forest fires that began on private land).

[243] *See* Bradshaw, *supra* note 239, 449-50 (2010) (providing distressing statistics on the costs of firefighting).

[244] Van Wagtendonk, *supra* note 213, at 14.

D. Race to the Top

Our third case study involves a very different policy realm. In 2009, the Department of Education launched a reform program called Race to the Top.[245] Like the federal experiments with soil and fire, Race to the Top was a federally-instigated program designed to spur policy innovation.[246] But unlike the soil and fire programs, the field staff implementing the program were state and local teachers and administrators, not federal employees.[247] Race to the Top thus exploited, rather than worked around, the political structures of federalism.[248]

The Race to the Top program sought to improve the quality of primary and secondary school education across the United States.[249] Its creators had ambitious ideas about improving teacher training, school administration, transparency, and accountability.[250] But unlike the Park Service or the Forest Service, which could experiment on federal lands, DOE did not have federal schools to use as its laboratories. And in contrast to the USDA, which began its soil experimentation during a period when the Great Depression had shattered state credibility and opened new possibilities for federal intervention, DOE did not have the political capital to simply take over the field of education reform.[251] Indeed, the most recent federal expansion into the field of education reform—the No Child Left Behind Act—had already become deeply unpopular.[252]

Consequently, DOE turned to a program of competitive and conditional grants to influence policy.[253] On its own, that approach is nothing new; using the federal spending power to influence state and local action is now

---

[245] *See* Fundamental Change, *supra* note 62.

[246] Overview Information, Race to the Top Fund: Notice Inviting Applications for New Awards for Fiscal Year (FY) 2010, 74 Fed. Reg. 59,836, 59,836 (Nov. 18, 2009) (describing the program purposes).

[247] *Id*. a vi-vii.

[248] *See generally* Martin Kurzweil, *Disciplined Devolution and the New Education Federalism*, 103 CALIF. L. REV. 565 (2015) (analyzing Race to the Top as part of an emerging new federalism model).

[249] *Id.*

[250] William G. Howell, *Results of President Obama's Race to the Top*, EDUCATION NEXT, Fall, 2015, p. 2 (noting the degree of control exercised by the Department of Education).

[251] *See* CASS SUNSTEIN, AFTER THE RIGHTS REVOLUTION: RECONCEIVING THE REGULATORY STATE 22-24 (1990).

[252] *See* Kurzweil, *supra* note 248, at 601-02 (describing opposition).

[253] Overview Information; Race to the Top Fund; Notice Inviting Applications for New Awards for Fiscal Year (FY) 2010, 74 Fed. Reg. 59,836 (2010); *see* SETTING THE PACE: EXPANDING OPPORTUNITY FOR AMERICA'S STUDENTS UNDER RACE TO THE TOP (2014).

commonplace.[254] But this particular grant program was distinctive in a few ways. First, the federal government carefully defined the goals of the program. Rather than simply authorizing state and local governments to apply for money to improve education, it demanded programs to improve outcomes in four specific subject areas.[255] DOE did not explain exactly how it expected state and local governments to achieve all of those goals,[256] or how much improvement it expected them to achieve; the idea, instead, was to let state and local governments test out and learn from different approaches.[257] But by incorporating each of those goals into its scoring metric for grant applications, DOE ensured that states that received grants would direct their experimentation toward goals of the federal government's choosing.[258]

Second, DOE created an elaborate system designed to facilitate data-gathering and learning. Grantee states, DOE explained,

> must make available, through formal (e.g., peer-reviewed journals) or informal (e.g., newsletters, Web sites) mechanisms, the results of any evaluations they conduct of their funded activities. In addition…, Race to the Top States, [local educational agencies], and schools are expected to identify and share promising practices, make work available within and across States, and make data available in appropriate ways to stakeholders and researchers so as to help all States focus on continuous improvement in service of student outcomes.[259]

---

[254] *See* Eloise Pasachoff, *Agency Enforcement of Spending Clause Statutes: A Defense of the Funding Cut-Off*, 124 YALE L.J. 248, 260-66 (2014) (describing the scope of federal grantmaking programs).

[255] 74 Fed. Reg. at 59,836. Those goals were:

> (a) Adopting internationally-benchmarked standards and assessments that prepare students for success in college and the workplace;
>
> (b) Building data systems that measure student success and inform teachers and principals in how they can improve their practices;
>
> (c) Increasing teacher effectiveness and achieving equity in teacher distribution; and
>
> (d) Turning around our lowest-achieving schools.

[256] Some of the criteria were so specific that they left little room for flexibility. *See* Kurzweil, *supra* note 248, at 603-04 (noting examples).

[257] *See* FUNDAMENTAL CHANGE, *supra* note 62, at vii.

[258] *See* 74 Fed. Reg. at 59,842-59,845 (defining criteria); *see* Howell, *supra* note 250, at 2 (noting the degree of control exercised by the Department of Education).

[259] 74 Fed. Reg. at 59,838 (parentheses in original).

DOE also called for the Institute of Educational Services to perform several national evaluations of the effects of Race to the Top grants.[260]

Despite those features, describing Race to the Top as a rigorously experimental program would be inaccurate. DOE did not require its grantee states to establish control groups—instead, it expected statewide policy changes—and in that key sense, the program architects prioritized getting reforms in place over facilitating experimental learning.[261] Nor did DOE itself advertise its program as experimental. That word is prominently absent from the Federal Register notices soliciting grant applications, and while it does appear occasionally in DOE's own reports upon the program, "innovation" instead was DOE's term of choice.[262] But the commitment to data collection and results dissemination, along with the reservation of latitude for interstate variation, still sets Race to the Top far apart from a program that simply attempts to push states toward a new set of federally-favored policies.

The resulting program thus bears some resemblance to the governance architecture called for by democratic experimentalism scholars, some of whom have leaped to claim Race to the Top as adoption reflection of their ideas.[263] But there are also key differences. Democratic experimentalism, in the classic account, involves the federal government primarily as a facilitator and as a vector for learning, while state and local governments take the lead in actually defining policy goals and selecting benchmarks.[264] In Race to the Top, in contrast, the federal government was more than just a facilitator, though it certainly did play that role. It also defined the goals.

Whether this program has improved educational outcomes is hard to say. Early retrospective studies have generally found that Race to the Top was effective in spurring states to adopt new policies, but determining whether those policies actually improved student learning will take more time.[265] And while that question is crucially important to the teachers and students

---

[260] *Id.*

[261] *See* 74 Fed. Reg. at 59,842 (demanding plans for "statewide" improvement). DOE also favored applications from states that were adopting the Common Core curricular standards, and thus encouraged some uniformity among as well as within states. 74 Fed. Reg. at 59,843.

[262] *See, e.g.*, 74 Fed. Reg. at 59,836 ("The purpose of Race to the Top… is to encourage and reward States that are creating the conditions for education innovation and reform."); Fundamental Change, *supra* note 62, at 40 ("States experimented with new approaches to provide districts with the supports and tools needed to assist low-performing schools and sustain improvements in teaching and learning.").

[263] *See* Sabel & Simon, *supra* note 49, at 56 (asserting that "the Race to the Top education program [] can only be understood in experimentalist terms").

[264] *See supra* notes 49-57 and accompanying text.

[265] *See* LISA DRAGOSET ET AL., RACE TO THE TOP: IMPLEMENTATION AND RELATIONSHIP STUDENT OUTCOMES (2016); Howell, *supra* note 250.

involved, for our purposes it is somewhat beside the point. Whether or not Race to the Top turns out to have improved education, it already has demonstrated the possibility of a program of experimentation—albeit of the moderately rigorous variety—instigated, directed, and overseen by the federal government and implemented by local governments and states.

IV. EXPLAINING FEDERAL EXPERIMENTATION

In American politics and law, the balance of federal and state authority is a subject of never-ending debate.[266] Congress is constantly deciding the degree to which federal governance should extend itself into policy realms both new and old, the extent to which the federal government should preserve state primacy, and how much state discretion Congress should preserve within the boundaries of federal programs. Consider, for example, our ongoing debates over health care policy: the federal-state balance (in a realm crying out for governmental experimentation) has been a central focus.[267] Administrative agencies repeatedly confront the same questions, and in a variety of different contexts. So do the courts. Beneath preemption,[268] Commerce Clause,[269] and 10th Amendment cases,[270] among others, and beneath statutory interpretation cases involving the shadows of constitutional law,[271] a key underlying question lurks: how powerful should the states and the federal government be? The answers to that question rarely turn solely on theories of federalism, and the desire for policy experimentation is just one part of federalism theory. But it is an important part. Its prominence within federalism debates justifies moving beyond the classic mantras of federalism and exploring how and why governmental experimentation actually happens.

---

[266] See New York v. United States, 505 U.S. 144, 149 (1992) (describing federalism as "our oldest question of constitutional law").

[267] *See* Michael S. Sparer, *Federalism and the Patient Protection and Affordable Care Act of 2010: The Founding Fathers Would Not Be Surprised*, 36 J. HEALTH POLITICS, POL'Y & L. 461, 463-66 (2011).

[268] *See* Roderick M. Jr. Hills, *Against Preemption: How Federalism Can Improve the National Legislative Process*, 82 N.Y.U. L. REV. 1, 3 (2007).

[269] *E.g.* Nat'l Fed. of Ind. Bus. v. Sebelius, 567 U.S. 519, 533-37 (2012).

[270] *E.g.* Printz v. U.S., 521 U.S. 898 (1997).

[271] *E.g.* Solid Waste Agency of Northern Cook Cnty. v. U.S. Army Corps of Eng'rs, 531 U.S. 159, 172-74 (2001) (invoking principles of constitutional avoidance to reject a statutory interpretation that would have allowed federal Clean Water Act jurisdiction over isolated wetlands); Gregory v. Ashcroft, 501 U.S. 452 (1991) (using avoidance principles to narrowly interpret a statute that would have invalidated state laws mandating age-based retirement for judges).

The case studies in Part III focus on the *how* questions. They demonstrate that the federal government can be involved to varying degrees in programs of policy experimentation, that those programs can involve varying levels of experimental rigor, and that, despite that variation, the federal roles can be crucially important. In this final Part, we ask *why* the federal government can and often should take on these roles. In contrast to the traditional federalism literature, which has generally assumed state primacy in the experimental field or has critiqued states' experimental potential, we consider whether the federal government brings distinctive advantages to governmental experimentation. We conclude that it does. There are structural features of federal governance—all revealed, to at least some extent, by our case studies—that explain why lawmakers and advocates should look to the federal government as a source of experimental policy.

A. Perspective

One important comparative advantage of the federal government—and one that might call for continued federal involvement in policy areas often viewed as areas of traditional sub-federal control— is its ability to see a broad problem that requires differentiated approaches. The Race to the Top and fire management programs both involved federal recognition of national problems, and our soil conservation example is perhaps the best illustration of this point. Decisions by thousands of individual farmers scattered around the United States contributed to a massive problem—dust storms caused by widespread soil erosion—that states and local governments had not adequately identified or addressed. It took a highly motivated federal official to draw attention to the problem and trigger an experimental program, with a central motivator being the aptly-named report entitled "a *national* menace."[272]

B. Differentiation, Coordination, and Communication

When it comes to implementing the experiment and ensuring differentiated approaches, the federal government again has important and underappreciated advantages. For reasons outlined in Part II, state and local government officials are often too risk averse, or tempted by free riding, to take on this differentiation effort.[273] Even with federal incentives, such as grants or prizes, the federal government might not be able to inspire systematically differentiated efforts. In many circumstances, therefore, the federal government itself will be the best experimenter. In other words,

[272] *See* 1980 Appraisal, *supra* note 127 (emphasis added).
[273] *See supra* notes 82-87 and accompanying text.

differentiation is sometimes best accomplished when a partially centralized actor like the federal government, which has both national headquarters offices and geographically distributed staff, sends agents to different regions rather than simply enlisting states and local government as agents.[274] The federal government can directly implement varied approaches without trying to persuade sub-federal entities to do so. And even if state or local agents are amenable to trying different policies—including those that risked lower success—sending federal employees to run agricultural experiment stations or to implement different fire management practices can substantially reduce the transaction and monitoring costs that accompany reliance on sub-federal implementers.

The federal government also has a relatively unique ability to ensure uniform measurement and data collection, as well as to inspire data sharing across sub-federal borders.[275] States and local governments have little incentive to collectively agree upon a uniform system or to share results—particularly negative ones.[276] Where the federal government has identified a national problem, it has greater incentives, and the advantage of a semi-centralized regime, when it comes to taking on this data-intensive effort. This is true even for problems that play out differently across geographic areas, as evidenced by forest fire management. The federal government is incentivized to identify these differences and to tailor policies to address them. Further, as demonstrated by Race to the Top, the government can enlist states and local governments in the data collection effort and can also lead conferences involving federal experts and sub-federal experimenters, inducing these officials to share their lessons learned more broadly.[277]

There are potential counterarguments to these claims. Perhaps the most obvious is that federal agencies will struggle to differentiate policy because of their greater removal from regulated entities and because those efforts would face a legitimacy deficit. The former argument, though it echoes themes common in federalism discourse, is often overblown; as the soil and fire examples illustrate, federal agency staff are just as close to many problems, if not closer, than their state counterparts.[278]

---

[274] *See* Owen, *supra* note 38, at 109. Parallel relationships are also possible between state governments, which often have field or regional offices within their states, and local governments.

[275] While states do not share this ability, some private organizations are heavily involved in semi-governmental activities like standard-setting. *See* Emily S. Bremer, *American and European Perspectives on Private Standards in Public Law*, 91 TULANE L. REV. 325, 326 (2016).

[276] *See supra* notes 97-98 and accompanying text.

[277] *See supra* notes 245-260 and accompanying text.

[278] *See supra* notes 140-175 and accompanying text (describing the localized staffing of federal soil conservation initiatives).

The latter argument is more nuanced and requires a lengthier response. The argument would start from the premise that, despite all the rhetoric celebrating policy differentiation, such differentiation is actually deeply worrisome, for it means deliberately introducing inconsistency into governance.[279] Usually we value treating like cases alike, and if an entity is going to establish a different practice, perhaps that entity should be an elected state legislature rather than an appointed federal administrator—or that administrator's civil service staff. That argument thus reflects one of the largest debates in administrative law, which pits the skeptics of administrative legitimacy against its advocates, and a full explication of that debate is beyond the scope of this article.[280] Our summary response is simply that we are persuaded by the many arguments in favor of administrative legitimacy. If a federal agency, which will be subject to legislative, executive, and judicial branch oversight, and whose actions will be bounded by governing statutory and constitutional law and by a generalized requirement of reasoned decision-making, decides to deliberately differentiate policy, and does so in ways consistent with principles of ethical research, that is a legitimate course to pursue.

C. Resources

Effective experimentation also requires a massive amount of resources, and here, too, the federal government often has an advantage. Because of the deep and broad bureaucratic structures at the federal level, in numerous policy areas the federal government can often contribute more expertise, money, and labor than its state and local counterparts. Both the soil and fire examples illustrate this point. As shown by our USDA case study, when there was relatively broad public support for enhanced federal involvement in a policy area and funding to match this support, the federal government deployed its extensive resources at the most local of levels, sending government scientists around the country to set up and run agricultural

[279] *See* Owen, *supra* note 38, at 76-77 (noting recurring concerns about inconsistent federal policies).

[280] *Compare, e.g.*, EPA v. EME Homer City Generation, LLP, 134 S.Ct. 1584, 1617 (2014) (Scalia, J. dissenting) ("Too many important decisions of the Federal Government are made nowadays by unelected agency officials exercising broad lawmaking authority, rather than by the people's representatives in Congress.") *with* Mark Seidenfeld, *A Civic Republican Justification for the Bureaucratic State*, 105 HARV. L. REV. 1511 (1992) (arguing that the deliberative processes of agencies compare favorably with those of legislatures). For a relatively recent distillation of arguments on both sides, see the majority and dissenting opinions in Free Enterprise Fund v. Public Co. Accounting Oversight Bd., 561 U.S. 477 (2010).

experiment stations.[281] The Civilian Conservation Corps also helped to carry out the experiments—installing the metal vats that collected soil to measure the effectiveness of conservation approaches and then disseminating the proven approaches, such as tree planting.[282] Later on, even when federal resources diminished somewhat, the government's role in soil conservation policy and experimentation remained quite sticky. There are still USDA field offices and thousands of USDA staff operating around the country.[283] The forest fire management case study also demonstrates that the federal government can provide not just the staffing, but also the medium in which the experiment plays out—in that case, the vast reserves of federally-owned land and trees necessary to investigate varied management practices.

The federal government will not always have this advantage. As the Race to the Top example illustrates, there are some policy realms in which state and local governments have a near-monopoly on facilities and staff. Often, this will simply be an accident of federalism and path dependence—governments at the local, state, regional, or federal levels will have built up expertise and resources in a particular area simply because they historically exercised the most responsibility in that particular field. Consequently, in designing an experiment in a particular policy area, officials should pay close attention to the government level or levels with the most amassed resources and expertise in this area. But often, that level will be federal. And even when money, expertise, and labor tend to be concentrated at a lower level, the federal government can also leverage those resources and help to coordinate them. In the education area, states and school districts collect most of the funds that support education and employ nearly all of the people involved in the education system.[284] Yet by granting funds based on a checklist of factors that that states and local governments could experiment with and enlisting experts to periodically check in with states and school districts and make them talk to each other on conference calls, the federal government helped to differentiate the experimental approach, collect important data, and spread lessons learned.[285]

---

[281] The government did, however, typically rely on donation of state lands, and later partnered with the states to conduct the experiments.

[282] *See supra* notes 141-142 and accompanying text.

[283] Local Service Centers Directory, USDA, https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/contact/local/.

[284] U.S. Dept. of Education, The Federal Role in Education (last visited December 20, 2017) ("[T]he Federal contribution to elementary and secondary education is about 8 percent" of total budgets.); *see also* Brown v. Bd. of Educ. of Topeka, 347 U.S. 483, 493 (1954) (describing education as "perhaps the most important function of state and local governments").

[285] *See supra* notes 259, 265 and accompanying text.

* * * * *

We make no claim, of course, that the federal government should always be a part of policy experiments. Some of the benefits of centralized coordination can be reproduced in the interactions between local governments and states, and the federal government need not always take the lead.[286] And often policy entrepreneurs cannot wait for the perfect institutional arrangements to arise. If the federal government is uninterested in supporting an effort at policy reform, as it sometimes will be, the more sensible course will be to proceed with those governmental entities whose leaders are willing to innovate, even if they would not be central players in an idealized experimental effort. Our core claim, instead, is that, as our examples show, the federal government has been, and should continue to be, a key player in policy experimentation. As legislators, executive branch officials, and judges adjust the boundaries of federal and state power, they should bear that lesson in mind.

CONCLUSION

For decades, American political discourse has placed a premium on policy experimentation. Sometimes, that emphasis may go too far; not every policy experiment is a worthy expenditure of public resources or is carried out for salutary ends.[287] But experimentation is often useful. To the extent that experimentation is an important value of our governance systems, it makes sense to ask what government structures will produce good experiments. Because American governance is inextricably intertwined with federalism, the answers to that question necessitate an inquiry into the intersection of policy experimentalism and federalist governance.

That inquiry, as this article has shown, leads to an unexpected outcome. The federal government, we have argued, has been and should be at the center of experimentalist policy. It also should sometimes be at the geographic periphery; in addition to coordinating experiments and communicating their results, federal staff across the nation sometimes should also carry them out. We make no claim that this will always be true; there is no single magic formula for experimentation. But, as our examples show, the often-important federal role deserves greater attention. And in a world of growing calls for broad-based federal deregulation, the importance of the federal government in supporting a central value of federalism should not be overlooked. There

[286] *See* Owen, *Cooperative Subfederalism*, *supra* note 61.

[287] *See generally* Livermore, *supra* note 20 (discussing downsides of experimental policy).

is ample room, in other words, for additional use of, and inquiry into, our federal laboratories of democracy.