
Federal Laboratories of Democracy

Hannah J. Wiseman^{†*} and Dave Owen^{**}

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 experimental rigor they incorporate 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, education, and welfare 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. Thus, it highlights 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

[†] Copyright © 2018 Hannah J. Wiseman and Dave Owen.

^{*} Attorneys' Title Professor of Law, Florida State University College of Law, J.D., Yale Law School, A.B., Dartmouth College.

^{**} 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 David Adelman, Scott Dodson, Malcolm Feeley, Michael Livermore, Jodi Short, and participants in the Maurer School of Law, Indiana Bloomington joint Maurer-Ostrom Colloquium series 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.

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.

TABLE OF CONTENTS

INTRODUCTION	1121
I. EXPERIMENTALISM AND FEDERALISM: AN OVERVIEW.....	1128
A. <i>Experimentalism and Traditional Federalism Theory</i>	1129
B. <i>Experimental Design and Adaptive Management</i>	1131
C. <i>Democratic Experimentalism Theory</i>	1134
II. THE CHALLENGES OF EXPERIMENTALIST FEDERALISM	1136
A. <i>Defining Policy Experiments</i>	1137
B. <i>Barriers to Combining Policy Experiments and</i> <i>Federalism</i>	1140
1. <i>Differentiation and Confounding Variables</i>	1140
2. <i>Data Collection and Analysis</i>	1142
3. <i>Repetition and Variation</i>	1144
C. <i>Federalism, Experimentation, and Centralized Authority</i> .	1145
III. MODELS OF POLICY EXPERIMENTATION.....	1146
A. <i>A Typology of Policy Experiments</i>	1146
B. <i>Agricultural Soil Conservation: A Rigorous Experiment</i> <i>with Extensive Federal Involvement</i>	1152
1. <i>Federal Agencies and Private Dirt</i>	1153
2. <i>Soil Conservation and Experimentation</i>	1161
a. <i>Hypotheses and Policy Differentiation</i>	1161
b. <i>Control of Confounding Variables</i>	1163
c. <i>Data Collection, Analysis, and Documentation</i>	1164
d. <i>Repetition and Variation of Experiments</i>	1167
C. <i>“Playing with Fire”</i>	1168
D. <i>Race to the Top</i>	1173
E. <i>Jobs-Plus</i>	1177
IV. EXPLAINING FEDERAL EXPERIMENTATION	1182
A. <i>Perspective</i>	1184
B. <i>Differentiation, Coordination, and Communication</i>	1184
C. <i>Resources</i>	1187
CONCLUSION.....	1191

INTRODUCTION

An oft-cited justification for federalism is that it induces creative policy experimentation at the state level.¹ According to the standard arguments, limiting federal power and protecting state sovereignty allow states to function as “laboratories of democracy,” places where governmental innovations can begin and spread.² For courts and federalism scholars, this alleged virtue has remained alluring for decades, and celebrations of state policy laboratories remain a central theme in the discourse of federalism.³ 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.⁴

¹ 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”). Of course, scholars and courts point to a variety of other potential reasons to concentrate or decentralize power, such as efficiency and responsiveness to local concerns. See, e.g., Heather K. Gerken, *Foreword: Federalism All the Way Down*, 124 HARV. L. REV. 4, 6, 47-48 (2010) (describing the many traditional justifications for federalism). This Article focuses solely on the oft-cited “laboratories” virtue of federalism.

² See, e.g., *New State Ice Co. v. Liebmann*, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) (describing states as “laborator[ies]”); 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.”).

³ See, e.g., *Ariz. State Legislature v. Ariz. Indep. 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).

⁴ 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*, 4 Q.J. POL. SCI. 35, 35-36 (2009) (noting that “political decentralization has been widely thought to stimulate policy experimentation and innovation”); Christos Kotsogiannis & Robert 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 E. Oates, *An Essay on Fiscal Federalism*, 37 J. ECON. LITERATURE 1120, 1133 (1999) (acknowledging that centralized governments can also 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,

Yet there are many reasons to be skeptical of these accounts. While no one disputes 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.⁵ Even when they do experiment, other state and local government characteristics may hinder good policies' paths to wider adoption.⁶ Consequently, if we value policy laboratories, 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

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 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).

⁵ See, e.g., MALCOLM M. FEELEY & EDWARD RUBIN, *FEDERALISM: POLITICAL IDENTITY & TRAGIC COMPROMISE* 26-29 (2008) [hereinafter *FEDERALISM*] (challenging the common "laboratories-of-the-states" justification for federalism and other common alleged virtues of federalism); Brian Galle & Joseph Leahy, *Laboratories of Democracy? Policy Innovation in Decentralized Governments*, 58 *EMORY L.J.* 1333, 1370, 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"); Susan Rose-Ackerman, *Risk Taking and Reelection: Does Federalism Promote Innovation?*, 9 *J. LEGAL STUD.* 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); Edward L. Rubin & Malcolm Feeley, *Federalism: Some Notes on a National Neurosis*, 41 *UCLA L. REV.* 903, 913 (1994) [hereinafter *Federalism: Some Notes*] (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"); Koleman S. Strumpf, *Does Government Decentralization Increase Policy Innovation?*, 4 *J. PUB. ECON. THEORY* 207, 208 (2002) (noting that "a local policymaker may free-ride off his neighbor's experiment").

⁶ See, e.g., Berry & Berry, *supra* note 4, at 401-05 (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, 1176, 1185 (1973) (noting that "diffusion patterns do 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 from interaction").

in federal governmental “interference.”⁷ If states are not the optimal experimenters, then broad-based shrinkage of federal involvement could reverse critical policy experimentation, thus undermining a virtue often used to justify state power.⁸

This Article responds to the challenge of constructing useful policy laboratories and 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.” 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 governance structures for experimentation 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

⁷ See, e.g., Exec. Order No. 13,771, 82 Fed. Reg. 9339 (Jan. 30, 2017) (announcing a broad initiative to reduce government involvement in numerous policy areas).

⁸ For example, in the context of controlling water pollution, Michael Livermore argues that reducing federal jurisdiction over waters “would hamper, rather than facilitate, beneficial experimentation.” Michael A. Livermore, *The Perils of Experimentation*, 126 YALE L.J. 636, 644 (2017).

⁹ Strumpf, *supra* note 5, 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 M. 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 4, at 36 (observing that “[e]xplicitly experimental local policies occur in both centralized dictatorships and centralized democracies”); Kotsogiannis &

Furthermore, federal initiatives sometimes incorporate attributes that make policy experimentation surprisingly rigorous¹² — far more rigorous than the haphazard patchwork of state policies that arise from the 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-designed and federally-implemented approach. The federal government used congressionally-approved funding to employ a true boots-on-the-ground system for modifying crop management practices that 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.¹³ The federal government also enlisted the Civilian Conservation Corps to implement these techniques around the country, with approaches that varied by region.¹⁴ The scientific elements of this program were thoroughly intertwined with policy experimentation. Federal on-the-ground learning extended to political and social questions, like determining how to achieve farmer buy-in once effective techniques had been identified.¹⁵ Further, these efforts informed the development

Schwager, *supra* note 4, at 484-85 (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.

¹³ See *infra* Part III.B.1.

¹⁴ See *infra* Part III.B.1.

¹⁵ See, e.g., Douglas Helms, *The Civilian Conservation Corps: Demonstrating the Value of Soil Conservation*, 40 J. SOIL & WATER CONSERVATION 184, 186 (1985) [hereinafter *The Civilian Conservation*] (noting early efforts to persuade farmers to adopt recommendations for soil conservation measures through cooperative agreements, in which the government supplied some of the materials necessary for implementing the measures and the farmers committed to implement the measures); *id.* at 185-86 (noting a local soil conservation demonstration district director’s intent to “demonstrate proper farm management,” and an early soil conservation leader’s goal of demonstrating “the values of conservation on an area larger than the individual farm”); Douglas Helms, *Two Centuries of Soil Conservation*, 5 OAH MAG. HIST. 24, 26 (1991) [hereinafter *Two Centuries*] (noting policy efforts by the Assistant Secretary of Agriculture to form an alternative to the solely federally-run demonstration projects in order to spread conservation practices “nationwide and have an impact on the way people farmed” by getting farmers “more interested and involved in . . . the work”); H.H. Bennett, *Conservation Work of the Department of Agriculture*, Speech at Raleigh, North Carolina 10-11 (Nov. 1, 1940) [hereinafter *Conservation Work*]

of agency policies and laws. These early federal initiatives 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 results and suggesting how conservation practices could improve.¹⁶

Other examples that we explore here, such as federal agencies' evolving policies for wildfire management¹⁷ and the United States Department of Education's Race to the Top,¹⁸ demonstrate that the soil conservation story has parallels in other fields.¹⁹ 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 illustrate the possibility, and 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 sub-federal governance.

To begin the project of melding experimental policy design with federalism and producing meaningful lessons for policy experimentation, Part I reviews existing literature. We describe 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 marry federalism and experimentation yet discounts the experimental potential of the federal government itself.²⁰ Part II

(noting Soil Conservation Service field stations, at which the government was investigating “[u]nder actual field conditions . . . the economic aspects of conservation farming” in addition to scientific questions such as “the relation between farming practices and sedimentation of stream channels [and] reservoirs”) (transcript available in the Iowa State University Special Collections and University Archives Department).

¹⁶ See *infra* Part III.B.1.

¹⁷ See *infra* Part III.C.

¹⁸ See *infra* Part III.D.

¹⁹ We make no claim, of course, that our examples exhaust the field of federal experimentation.

²⁰ For a very rare exception to these generalizations, see Joseph Landau,

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. This Part includes 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 could interfere with key state and local experiments.²¹ Nevertheless, the federal government brings key 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.²²

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.

²¹ See, e.g., *Coal. for Competitive Elec. v. Zibelman*, 272 F. Supp. 3d 554, 567 (S.D.N.Y. 2017), *appeal docketed*, No. 17-2654 (2d Cir. Aug. 25, 2017); *Vill. of Old Mill Creek v. Star*, No. 17 CV 1163 & No. 17 CV 1164, 2017 WL 3008289, at *6 (N.D. Ill. July 14, 2017). Both cases address whether the Federal Power Act preempts state clean energy initiatives. Although the court in *Zibelman* do not find preemption, plaintiffs have appealed the case. *Zibelman*, 272 F. Supp. 3d at 567.

²² 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., Livermore, *supra* note 8, at 644 (briefly arguing in favor of “managed experimentation,” in which the federal government, states, and local governments all play a role); Rubin & Feeley, *Federalism: Some Notes*,

The lessons from Part IV have important implications not just for federalism theorists, but also, more practically, for legislators, judges, and advocates. Legislators often draft statutory provisions designed to facilitate policy innovation, often looking exclusively to states to supply that innovation. “Big waivers” are just one prominent example of this phenomenon.²³ 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.²⁴ 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. Thus, an advocate’s best option will usually be the governance level that is willing to try, not the one that would be the 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 advocacy.

Intentional policy experimentation — no matter the governance level at which it occurs — rarely matches the type of carefully-planned experimentation that occurs in a scientific laboratory. Indeed, although Justice Brandeis, in coining the famous laboratories term, likely envisioned something close to a true experiment,²⁵ the case law and literature that followed used the term increasingly loosely. But defining a new space for policy experimentation broadens opportunities for designing meaningful and more effective

supra note 5, at 924 (describing how the federal government could coordinate experiments and vary approaches by locality or region).

²³ See David J. Barron & Todd D. Rakoff, *In Defense of Big Waiver*, 113 COLUM. L. REV. 265, 271 (2013) (describing the granting of waivers to states as a way to induce policy experimentation).

²⁴ See *infra* notes 35–38 and accompanying text.

²⁵ See Barry Friedman, *Valuing Federalism*, 82 MINN. L. REV. 317, 399 (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 3, at 1751 (“Justice Brandeis does not appear to view ‘experimentation’ as a metaphor. His government policymakers operate ‘in the fields of social and economic science.’”). Indeed, Justice Brandeis introduced this term during a period of close attention to policy experimentation that borrowed directly from science. See generally JOHN DEWEY, *LOGIC: THE THEORY OF INQUIRY* (1938) (arguing for scientifically-driven and empirically-supported social policy); Foreword, *Symposium on Cooperative Federalism*, 23 IOWA L. REV. 455 (1938) (discussing policy experimentation in a traditional experimental light).

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.²⁶ Academics often echo the judicial statements and sometimes elaborate upon them.²⁷ But most exploration of the intersections of federalism and experimentalism has remained relatively cursory, 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 governance. Academics and judges do routinely discuss governance innovations instigated by the federal government, including the federal government's facilitation of sub-federal experimentation.²⁸ But when they turn from specific examples to broader theories, their discussions typically adopt fairly simplistic models of experimentalism or federalism — or both — and the federal government is often viewed as no more than a facilitator or coordinator of sub-federal experiments.

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,

²⁶ See Livermore, *supra* note 8, at 648 (“Within legal scholarship, experimentation is often understood through the lens of federalism.”).

²⁷ In work that comes closer to recognizing the role of the federal government in initiating policy innovation, Ann Carlson notes how the federal government and states sometimes trade off roles over time, with either the states or federal government pushing for innovative policy changes. See Ann E. Carlson, *Iterative Federalism and Climate Change*, 103 NW. U.L. REV. 1097, 1098-1100 (2009).

²⁸ See, e.g., WILLIAM R. LOWRY, *THE DIMENSIONS OF FEDERALISM: STATE GOVERNMENTS AND POLLUTION CONTROL POLICIES* 15-16 (1992) (noting that “[c]hange through dispersion and emulation of innovation is fostered by open lines of communication that can be maintained by vertical hierarchy,” with vertical hierarchy referring to federal intervention); William Boyd & Ann E. Carlson, *Accidents of Federalism: Ratemaking and Policy Innovation in Public Utility Law*, 63 UCLA L. REV. 810, 817 (2016) (noting federal “intentional policy nudges and subsidies to push states to innovate” toward low-carbon energy policy and exploring these “nudges” in depth); Susan Welch & Kay Thompson, *The Impact of Federal Incentives on State Policy Innovation*, 24 AM. J. POL. SCI. 715, 727 (1980) (concluding that “incentives provided by the federal government do stimulate the diffusion of [innovative] policies through the states”); *infra* sources cited note 57 (describing more sources that have focused on the government as an experiment facilitator); *infra* source cited note 60 and accompanying text.

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. In a forgettable opinion, a majority of the Court set the regulatory controls aside.²⁹ Justice Brandeis dissented. The nation then was deep in the Great Depression — “an emergency more serious than war,” in Brandeis’ words.³⁰ As he acknowledged, the path out of those dark times was far from clear.³¹ Brandeis was sure of one thing, however: “There 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.”³² 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.”³³

That sentence has become iconic.³⁴ The Court often quotes it, or, in more shorthand form, refers to “laboratories for experimentation.”³⁵ The references come in opinions by liberals and conservative justices alike and in cases addressing a wide variety of subject matter.³⁶ On an

²⁹ *New State Ice Co. v. Liebmann*, 285 U.S. 262, 278-80 (1932).

³⁰ *Id.* at 306 (Brandeis, J., dissenting).

³¹ *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.”).

³² *Id.* at 311.

³³ *Id.*

³⁴ 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”).

³⁵ *See Fisher v. Univ. of Tex.*, 136 S. Ct. 2198, 2214 (2016); *Ariz. State Legislature v. Ariz. Indep. Redistricting Comm’n*, 135 S. Ct. 2652, 2673 (2015); *see, e.g., Oregon v. Ice*, 555 U.S. 160, 171 (2009); *Arizona v. Evans*, 514 U.S. 1, 8 (1995); *Chandler v. Florida*, 449 U.S. 560, 579 (1981); *Reeves v. Stake*, 447 U.S. 429, 441 (1980); *Whalen v. Roe*, 429 U.S. 589, 597 & n.20 (1977); *San Antonio Indep. Sch. Dist. v. Rodriguez*, 411 U.S. 1, 49-50 (1973); *Fay v. New York*, 332 U.S. 261, 296 (1947).

³⁶ *See, e.g., Evenwel v. Abbott*, 136 S. Ct. 1120, 1141 (2016) (Thomas, J., concurring) (“[A]s the Court recently reminded us, States are free to serve as ‘laboratories’ of democracy.” (quoting *Ariz. State Legislature*, 135 S. Ct. at 2673)). The Court’s *Arizona State Legislature* opinion, which Justice Ginsburg authored, in turn

often-divided Court, the value of state laboratories is one thing everyone can agree on.³⁷ 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.³⁸ 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.³⁹ 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 powers facilitates experimentation,⁴⁰ and from "dynamic" or "interactive" federalists, who argue that overlap between federal and state powers similarly enables and strengthens policy experimentation.⁴¹ Indeed, the latter literatures, which address cooperative federalism and other variations from the traditional dual-federalist theme, often acknowledge the importance of federal involvement in experimentation and policy innovation, but typically as a facilitator or trigger, not initiator, of experiments.⁴² Work by both

quotes Justice Kennedy's concurring opinion in *United States v. Lopez*, 514 U.S. 549, 581 (1995), for the same proposition. Ariz. State Legislature, 135 S. Ct. at 2673.

³⁷ 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); *Gonzales v. Raich*, 545 U.S. 1, 42-43 (2005) (O'Connor, J., dissenting); *Boy Scouts of Am. v. Dale*, 530 U.S. 640, 664 (2000) (Stevens, J., dissenting).

³⁸ For a very rare example of the Court applying this reasoning to a local government, see *San Antonio Indep. Sch. Dist.*, 411 U.S. at 49-50 (noting "[a]n analogy to the Nation-State relationship in our federal system").

³⁹ See, e.g., Friedman, *supra* note 25, at 399 ("Common intuition suggests that the vast majority of techniques used today to govern were developed at the state and local level.").

⁴⁰ See Livermore, *supra* note 8, at 648-49 (summarizing and citing scholarship in this realm); see also *Gonzales*, 545 U.S. at 42-43 (O'Connor, J., dissenting) (arguing that protecting "historic spheres of state sovereignty from excessive federal encroachment" allows states to function as laboratories of democracy).

⁴¹ See Heather K. Gerken, *Federalism as the New Nationalism: An Overview*, 123 YALE L.J. 1889, 1902 (2014) (summarizing and citing multiple sources in this vein).

⁴² See, e.g., LOWRY, *supra* note 28, at 15 ("Vertical involvement [by the federal

schools of thought rarely attends to either the actual mechanics of experimentation or the experimental possibilities of federal governance.⁴³ 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 devotes 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.⁴⁴ Some articles at the edges of this

government through cooperative federalism] affects state leadership. State leadership involves cooperation in, competition over, and dissemination of innovations and effective practices. Federal intervention can facilitate coordination and communication of state efforts.”); Jessica Bulman-Pozen & Heather K. Gerken, *Uncooperative Federalism*, 118 YALE L.J. 1256, 1287 (2009) (arguing that cooperative federalism facilitates states’ proposed policy variations).

⁴³ See Dave Owen, *Regional Federal Administration*, 63 UCLA L. REV. 58, 113-14 (2016) [hereinafter *Regional Federal Administration*] (noting that conventional schools of federalist thought often assume effective communication among levels of government). For an exception, see Landau, *supra* note 20, 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. This work acknowledges the possibility of experiments beginning with the federal government. See, e.g., Boyd & Carlson, *supra* note 28; Kirsten H. Engel, *Democratic Environmental Experimentalism*, 35 UCLA J. ENVTL. L. & POL’Y 57, 70-71 (2017). 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.

⁴⁴ See generally Abramowicz et al., *supra* note 1 (explaining how statistical research should be conducted to influence policy); Daniel E. Ho, *Does Peer Review Work? An Experiment of Experimentalism*, 69 STAN. L. REV. 1 (2017). Kenneth Abbott and Duncan Snidal’s 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: Scientific Inquiry and Policy Learning* (unpublished draft on file with authors) (describing different types of experiments and briefly exploring the “experimentalist governance process”); see also ROBERT A.

vein just call for more governmental experimentation.⁴⁵ But the core studies argue that government should conduct rigorous, randomized policy experiments, and some work demonstrates, through real-world examples, how governments can experiment and the insights such experimentation can produce.⁴⁶ Similarly, evidence-based policy — a growing body of work that we place within this experimental design category — strives to incorporate important attributes of policy experimentation, such as controls and randomization, and to apply the lessons of these experiments in a range of contexts, from criminal procedure to health care and education policy.⁴⁷

SCHAPIRO, POLYPHONIC FEDERALISM: TOWARD THE PROTECTION OF FUNDAMENTAL RIGHTS 6-7 (2009) (arguing for a focus on how experimentation within federalism works); Doni Gewirtzman, *Complex Experimental Federalism*, 63 *BUFF. L. REV.* 241, 245 (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”).

⁴⁵ 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).

⁴⁶ See, e.g., Ho, *supra* note 44; Abramowicz et al., *supra* note 1, at 987-1004.

⁴⁷ See, e.g., JAMES A. RICCIO, *SUSTAINED EARNINGS GAINS FOR RESIDENTS IN A PUBLIC HOUSING JOBS PROGRAM: SEVEN-YEAR FINDINGS FROM THE JOBS-PLUS DEMONSTRATION* 3 n.3 (2010), https://www.mdrc.org/sites/default/files/policybrief_33.pdf [hereinafter *SUSTAINED EARNINGS GAINS*] (explaining the use of control and treatment groups in the Jobs-Plus program, which we describe in a case study below); Kathleen J. Sikkema, *HIV Prevention Among Women in Low-Income Housing Developments: Issues and Intervention Outcomes in a Place-Based Randomized Controlled Trial*, 599 *ANNALS AM. ACAD. POL. & SOC. SCI.* 52, 56-57 (2005) (describing policy trials involving “HIV prevention community-level intervention projects, focused on men patronizing gay bars in sixteen small U.S. cities” and a study that built from this to determine whether “HIV prevention intervention would have a similar effect with different and more disadvantaged populations,” thus applying the program to “impoverished and predominantly minority women who live in low-income, inner-city housing developments”); Roger K. Warren, *Evidence-Based Sentencing: The Application of Principles of Evidence-Based Practices to State Sentencing Practice and Policy*, 43 *U.S.F. L. REV.* 585, 596 (2009) (describing “rigorous evaluations of various types of corrections programs using non-treatment control groups well-matched to the treatment group”); David Weisburd, *Hot Spots Policing Experiments and Criminal Justice Research: Lessons from the Field*, 599 *ANNALS AM. ACAD. POL. & SOC. SCI.* 220, 221 (2005) (describing policy trials involving “[h]ot spots policing” — “the concentration of police resources in small discrete areas”). In contrast, some initiatives labeled as evidence-based policy merely require or encourage decisionmakers to collect data and incorporate those data into policy. See, e.g., Eloise Pasachoff, *Two Cheers for Evidence: Law, Research, and Values in Education Policymaking and Beyond*, 117 *COLUM. L. REV.* 1933, 1940-44 (2017) (critiquing the Every Student Succeeds Act of 2015, which requires only limited collection and implementation of data, and describing the generous definitions of “evidence” used in the Act); Will Rhee, *Evidence-Based Federal Civil Rulemaking: A New Contemporaneous Case Coding Rule*, 33

The second body of work considers “adaptive management.”⁴⁸ 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.⁴⁹ Proponents of adaptive management argue that in a world of uncertainty and limited knowledge, policy must and should evolve through learning.⁵⁰ That means treating policies as provisional experiments, monitoring their results, and continuously adjusting them.⁵¹

The adaptive management literature is enormous, and it explores a wide range of sub-issues.⁵² But for our purposes, just two points about both the adaptive management and experimental design literature are particularly important. Each typically focuses on experiments implemented by a single governing entity,⁵³ and neither has much to say about how its recommended governance approaches should be integrated with a federalist system.⁵⁴ In our view, this is not a failing.

PACE L. REV. 60, 86-95 (2013) (describing empirical studies of issues such as hung juries and sealed settlements and the Civil Rules Advisory Committee’s growing reliance on empirical evidence).

⁴⁸ See, e.g., Robin Kundis Craig & J.B. Ruhl, *Designing Administrative Law for Adaptive Management*, 67 VAND. L. REV. 1, 3-4, 9-10 (2014).

⁴⁹ E.g., KAI N. LEE, COMPASS AND GYROSCOPE: INTEGRATING SCIENCE AND POLITICS FOR THE ENVIRONMENT 9 (1993) (“[P]olicies are experiments; *learn from them.*”).

⁵⁰ See Craig R. Allen et al., *Adaptive Management for a Turbulent Future*, 92 J. ENVTL. MGMT. 1339, 1339 (2011).

⁵¹ 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).

⁵² The literature critiquing adaptive management is also extensive. See, e.g., Holly Doremus, *Adaptive Management, The Endangered Species Act, and the Institutional Challenges of “New Age” Environmental Management*, 41 WASHBURN L.J. 50, 51-66 (2001) (describing “barriers” to adaptive management as well as the “fuzziness” of the term); Dave Owen, *Probabilities, Planning Failures, and Environmental Law*, 84 TUL. L. REV. 265, 330-35 (2009) (offering a qualified critique of adaptive management while conceding its value in some circumstances).

⁵³ But see Welch & Thompson, *supra* note 28, at 717-28 (discussing federal grants-in-aid 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 & Jim Rossi, *Agency Coordination in Shared Regulatory 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”).

⁵⁴ See, e.g., Craig & Ruhl, *supra* note 48, at 63-87 (providing “The Model Adaptive

The lack of attention to federalism simply 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.⁵⁵ 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.⁵⁶ 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.⁵⁷

Management Procedure Act,” which envisions action by a single federal agency).

⁵⁵ See Michael C. Dorf & Charles F. Sabel, *A Constitution of Democratic Experimentalism*, 98 COLUM. L. REV. 267, 288, 345 (1998); Charles Sabel & William H. Simon, *Minimalism and Experimentalism in the Administrative State*, 100 GEO. L. J. 53, 55 (2011).

⁵⁶ Dorf & Sabel, *supra* note 55, at 288, 345.

⁵⁷ See, 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); Livermore, *supra* note 8 (exploring “managed experimentation”); 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 5, at 615-16 (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”); sources cited *supra* note 28 (introducing some of the sources that have focused on the federal government as a facilitator and coordinator of experiments). Brian Galle and Joseph Leahy observe that a federal facilitative role is not necessary in

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 three 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.⁵⁸ As Kenneth Abbott and Duncan Snidal note, “[a]ny situation in which actors ‘try different things’ is considered to be experimental, without engagement with the well-developed understandings of experimentation in the natural and social sciences.”⁵⁹

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,⁶⁰ but in their proposed governance model, “the state and local governments actually do the experimenting.”⁶¹ In contrast, the federal government, which they describe as highly centralized, facilitates and does little else.⁶² In reality, however, the federal government is itself decentralized in many ways, some of which can enable differentiation and experimentation.⁶³ 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, demonstrate,

some scenarios, such as when state or local governments are “highly heterogeneous” and will produce few positive externalities from which others would benefit. These entities might simply independently innovate. Galle & Leahy, *supra* note 5, at 1361.

⁵⁸ See, e.g., Sabel & Simon, *supra* note 55, 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.

⁵⁹ Abbott & Snidal, *supra* note 44, at 2.

⁶⁰ See Dorf & Sabel, *supra* note 55, at 332-36, 382-88 (describing innovative federal policies).

⁶¹ *Id.* at 428.

⁶² See, e.g., Dorf & Sabel, *supra* note 55, at 428.

⁶³ See Owen, *Regional Federal Administration*, *supra* note 43, 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).

and diffuse better agricultural conservation policy⁶⁴ and fire management⁶⁵ 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.⁶⁶ 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.⁶⁷

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.⁶⁸ Real-world 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 emerged naturally and 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

⁶⁴ See *infra* note 159 and accompanying text.

⁶⁵ See *infra* notes 246–52 and accompanying text.

⁶⁶ Dave Owen, *Cooperative Subfederalism* 1-4 (UC Hastings College of Law, Research Paper No. 258, 2017), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3071907 [hereinafter *Cooperative Subfederalism*].

⁶⁷ U.S. DEP'T OF EDUC., *FUNDAMENTAL CHANGE: INNOVATION IN AMERICA'S SCHOOLS UNDER RACE TO THE TOP* xvi, 32 (2015), <https://www2.ed.gov/programs/racetothetop/rttfinalrpt1115.pdf> [hereinafter *FUNDAMENTAL CHANGE*] (“States are using performance management approaches to help districts support effective interventions in their lowest-performing schools. . . . States like Tennessee, North Carolina and Massachusetts created networks of their lowest-performing schools that improved supports for teachers and school and district leaders. . . . Ohio districts hired former principals with track records of improving student achievement to coach principals in struggling schools.”); see DIST. REFORM SUPPORT NETWORK, *TRANSFORMING THE CULTURE OF TEACHING AND LEARNING: FOUR RACE TO THE TOP-DISTRICT GRANTEES' IMPLEMENTATION OF PERSONALIZED LEARNING* 4, <https://rttd.grads360.org/services/PDCService.svc/GetPDCDocumentFile?fileId=21503>.

⁶⁸ See, e.g., Owen, *Cooperative Subfederalism*, *supra* note 66, at 10-12 (describing air quality management districts and land use planning agencies with territories encompassing multiple cities and counties).

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.⁶⁹ 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 test.
- 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.⁷⁰ But in either case, the differentiation should allow a comparison that will put the experimental hypothesis to the test.⁷¹
- Third, experimentation requires control of confounding variables. In a controlled experiment, experimenters can randomize the distribution of subjects into groups with different treatments, and they can control variables by focusing differentiation on a single key attribute.⁷² For natural experiments, such control is much more difficult.⁷³

⁶⁹ See Michael J. Saks, *Scientific Method: The Logic of Drawing Inferences from Empirical Evidence*, in 1 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY 305, 312 (David Faigman et al. eds., 2017).

⁷⁰ See, e.g., Mark R. Rozenssweig & Kenneth I. Wolpin, *Natural “Natural Experiments” in Economics*, 38 J. ECON. LITERATURE 827, 828 (2000).

⁷¹ See Saks, *supra* note 69, at 304 (explaining that an experiment should try to falsify its hypothesis).

⁷² See Abramowicz et al., *supra* note 1, at 934-37 (explaining how randomization works and the advantages it provides).

⁷³ See *id.* at 939-43 (describing advantages of randomization relative to regression analyses of natural experiments).

Nevertheless, techniques like regression analysis can sometimes allow for enough control of enough confounding variables to facilitate qualified confidence in experimental conclusions.⁷⁴

- Fourth, experimentation requires observation and data collection, and requires analysis of those observations and data.⁷⁵
- 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.⁷⁶ Typically, that paper will be peer-reviewed prior to publication.⁷⁷ 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.⁷⁸
- 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.⁷⁹ They instead

⁷⁴ See, e.g., Jeffrey J. Rachlinski, *Noah by the Numbers: An Empirical Evaluation of the Endangered Species Act*, 82 CORNELL L. REV. 356, 373-88 (1997) (reviewing CHARLES C. MANN & MARK L. PLUMMER, *NOAH'S CHOICE: THE FUTURE OF ENDANGERED SPECIES* (1995)) (using statistical analysis to evaluate the effectiveness of various policy strategies deployed under the Endangered Species Act, while controlling for variables, to undercut claims that the Endangered Species Act is ineffective).

⁷⁵ See Saks, *supra* note 69, 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).

⁷⁶ 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 researchers 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 1, at 943 (describing this problem).

⁷⁷ See Jerome P. Kassirer & Edward W. Champion, *Peer Review: Crude and Understudied, but Indispensable*, 272 JAMA 96, 96 (1994) (describing and critiquing peer review practices).

⁷⁸ 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.”).

⁷⁹ See Saks, *supra* note 69, at 305 (noting that researchers become confident that a hypothesis is probably correct only after it has survived repeated attempts at

try to reproduce experiments, both to test the validity of the original results and to discern the sensitivity of results to different interventions.⁸⁰

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,⁸¹ 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,⁸² 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.⁸³ 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.⁸⁴ These tin-standard experiments 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.⁸⁵

falsification).

⁸⁰ See Arturo Casadevall & Ferric C. Fang, *Reproducible Science*, 78 INFECTION & IMMUNITY 4972, 4972 (2010) (describing reproducibility as “a bedrock principle in the conduct and validation of experimental science”).

⁸¹ To be fair, the pots are calling the kettle black. We both have used the term “experiment” somewhat indiscriminately in our own past work.

⁸² 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.

⁸³ 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 2* (Nat'l Bureau of Econ. Research, Working Paper No. 22595, 2017), <http://www.nber.org/papers/w22595.pdf>.

⁸⁴ Much of the literature on evidence-based policy and governmental innovation, while open to more rigorous experimental approaches, encourages and recognizes the value of these alternative approaches. See sources cited *supra* note 47.

⁸⁵ See *infra* Part II.B (discussing reasons, some of them perfectly valid, why governments might not want to experiment).

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

Policy differentiation is the reason why a connection between federalism and experimental governance seems obvious.⁸⁶ Indeed, a federalist system seems designed, above all else, to allow different subnational jurisdictions to adopt differing policies.⁸⁷ 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.⁸⁸ 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.⁸⁹ Imitation, after all, is usually easier than invention.⁹⁰ 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.⁹¹ There are countervailing

⁸⁶ 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”).

⁸⁷ 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.”).

⁸⁸ See Rose-Ackerman, *supra* note 5, at 614-15 (arguing that federalism is not nearly as likely to promote experimentation as its proponents suggest); Galle & Leahy, *supra* note 5, at 1339 (largely agreeing with Rose-Ackerman’s conclusions).

⁸⁹ See Rose-Ackerman, *supra* note 5, at 610-11.

⁹⁰ See Scott Dodson, *The Gravitational Force of Federal Law*, 164 U. PA. L. REV. 703, 707-29 (2016) (documenting state’s reluctance to differentiate state law from federal even in realms where states have clear authority to chart their own courses).

⁹¹ See Paul Frymer & Albert Yoon, *Political Parties, Representation, and Federal*

incentives, of course. Leaders can make their names as policy entrepreneurs,⁹² and sometimes a subnational political climate will reward leaders for differentiating their jurisdiction from its neighbors or from the national government.⁹³ 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. For example, two local governments might adopt different responses to the same policy problem. 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.⁹⁴ 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

Safeguards, 96 NW. U. L. REV. 977, 980 (2002) (arguing that the agendas of national political parties dominate state policy selection).

⁹² See, e.g., Mintrom, *supra* note 4, at 739, 765 (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”).

⁹³ See, e.g., 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); Hiroko Tabuchi, *U.S. Climate Change Policy: Made in California*, N.Y. TIMES (Sept. 27, 2017), <https://www.nytimes.com/2017/09/27/climate/california-climate-change.html>.

⁹⁴ See Abbott & Snidal, *supra* note 44, at 21 (noting the tendency of experimenters to “bunch” around promising interventions, avoiding others that might in fact prove successful”); Gary King et al., *A “Politically Robust” Experimental Design for Public Policy Evaluation, with Application to the Mexican Universal Health Insurance Program*, 26 *J. POLY 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).

differentiation, even if sometimes overstated, has led to a history of differentiation. 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.⁹⁵ Yet the subnational jurisdictions of the United States are more akin to the former menagerie than the latter controlled environment. 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 variables is through randomization.⁹⁶ But randomization requires a high level of control from the experimenter, who randomly assigns subjects to different treatments.⁹⁷ And a core purpose of federalism is to limit this type of centralized control, and thus to protect the partial sovereignty of state governments.⁹⁸ 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.⁹⁹ They will likely have adopted

⁹⁵ Cf. FEELEY & RUBIN, *FEDERALISM*, *supra* note 5, at 28 (noting that in medical research “it would be unusual for the researcher to authorize the subjects to follow whatever course of treatment they desire” and observing that a centralized authority would have to be involved in designing an experiment to avoid this type of result).

⁹⁶ See Abramowicz et al., *supra* note 1, at 933 (arguing that randomization should be used for more policy analyses).

⁹⁷ 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.

⁹⁸ See *Gregory v. Ashcroft*, 501 U.S. 452, 458 (1991) (extolling the benefits of divided government).

⁹⁹ 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).

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.¹⁰⁰ Or, somewhat similarly, the only data that interests them may be information about the political marketability of their initiative, not its actual success in achieving policy goals.¹⁰¹ Additionally, data collection is often expensive.¹⁰² If overall program funds 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 and already well-designed, and 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.¹⁰³ Or, it might even think that because its neighbors will collect data on some policy experiment, it can skimp on data collection and 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 addresses uniform metrics and were collected in consistent ways. Otherwise, any attempt at 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.¹⁰⁴ That differentiation will not

¹⁰⁰ 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.”).

¹⁰¹ See Livermore, *supra* note 8, at 639 (noting that politicians may pursue experiments for reasons other than helping the public).

¹⁰² See Biber, *supra* note 99, at 31.

¹⁰³ See Hannah J. Wiseman, *Regulatory Islands*, 89 N.Y.U. L. REV. 1661, 1713-14 (2014) [hereinafter *Regulatory Islands*] (noting disincentives to gather and share information).

¹⁰⁴ See, e.g., NAT’L ACAD. OF PUB. ADMIN., *EVALUATING ENVIRONMENTAL PROGRESS: HOW EPA AND THE STATES CAN IMPROVE THE QUALITY OF ENFORCEMENT AND COMPLIANCE*

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.¹⁰⁵

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.¹⁰⁶ With 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.¹⁰⁷ 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 and local governments to obtain information about what other jurisdictions' policies even are.¹⁰⁸ 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. This makes obtaining information about those policies a time-consuming exercise with quickly-outdated results.¹⁰⁹ To gather information about the consequences of those policies will be even more difficult if such

INFORMATION 24 (2001) (describing state-to-state differences in data collection practices).

¹⁰⁵ See, e.g., *id.*

¹⁰⁶ 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); Katerina Linos, Note, *When Do Policy Innovations Spread? Lessons for Advocates of Lesson-Drawing*, 119 *HARV. L. REV.* 1467, 1468-70 (2006); sources cited *supra* note 6 (describing some of the political science literature on diffusion).

¹⁰⁷ See Abbott & Snidal, *supra* note 44, at 10 (describing how political boundaries can facilitate comparisons); Berry & Berry, *supra* note 4, at 403 (noting “research that has found that there are states to which the other states in a region look most frequently for innovative ideas”).

¹⁰⁸ See Wiseman, *Regulatory Islands*, *supra* note 103, at 1699-702.

¹⁰⁹ See, e.g., *id.* at 1694-704 (describing how hard it can be for states to obtain information about other states' fracking policies).

information even exists.¹¹⁰ 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.¹¹¹ Consequently, in the political real world, policy propagation is likely to be haphazard and uninformed. And the ideological narratives surrounding a policy may drive the spread of policies more effectively than 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 offers the possibility of such management. While the political and, sometimes, judicial and 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.¹¹² 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.¹¹³ And in state-local relationships, the states could play that same centralized coordinating role.¹¹⁴

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.

¹¹⁰ See *supra* notes 101–09 and accompanying text (describing disincentives to gather and share monitoring data).

¹¹¹ See Wiseman, *Regulatory Islands*, *supra* note 103, at 1715–16 (noting disincentives to gather and share information).

¹¹² 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).

¹¹³ See *supra* notes 56–68 and accompanying text.

¹¹⁴ See Owen, *Cooperative Subfederalism*, *supra* note 66, at 3–5.

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. The first is the degree of federal involvement in the experiment or, conversely, state or local leadership. The second is 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, education, and welfare 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 experimental rigor is optimal. The answers to those questions will likely be highly contextual, and they 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.¹¹⁵ Through happenstance, this type of haphazard

¹¹⁵ This type of experiment has evolved in the context of regulating the environmental impacts of oil and gas development. See generally Wiseman, *Regulatory Islands*, *supra* note 103 (describing varied state policies that lack a common goal or approach and that seem unmoored from any sort of broader experimental approach);

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.¹¹⁶ Alternatively, the federal government may act alone.¹¹⁷ We characterize these extremes, and the many gray areas in between, in two primary ways. We note, first, that policy experiments differ in terms of which levels of government design and implement them (the federalism aspect of policy experimentation). Second, experiments differ in the extent to which they incorporate key features, such as differentiation and control of confounding variables (the design aspect of policy experimentation). The framework does not capture every important variable, of course. The balance of congressional and agency involvement is another potential way of differentiating federal experimental programs,¹¹⁸ as is the degree to which participation in the experimental program is mandatory or voluntary.¹¹⁹ But the degree

Hannah J. Wiseman, *Risk and Response in Fracturing Policy*, 84 U. COLO. L. REV. 729 (2014) [hereinafter *Risk and Response*] (describing substantial variation in state oil and gas policy).

¹¹⁶ See *infra* notes 268–91 and accompanying text (describing the Race to the Top program).

¹¹⁷ See *infra* notes 234–69 and accompanying text (describing federal experiments with wildfire).

¹¹⁸ Our examples primarily describe agency experimentation, but Congress funded the Soil Experiment Stations described in Part III.B.1 below.

¹¹⁹ Voluntary or informal governance, in which actors exercise influence through ongoing relations, norms, non-contractual mechanisms, and other modes of influence beyond direct mandates or agreements is recognized in a variety of other contexts, including internationally and domestically within particular legal fields like contracts, but not extensively in the federalism context. See generally INTERNATIONAL HANDBOOK ON INFORMAL GOVERNANCE (Thomas Christiansen & Christine Neuhold eds., 2012); AVINASH K. DIXIT, LAWLESSNESS AND ECONOMICS: ALTERNATIVE MODES OF GOVERNANCE (2004). Our examples of experimentation primarily involve voluntary programs, although some involve so much funding that they fall more toward the mandatory side of the spectrum, in that farmers likely would not participate absent the huge infusion of cash offered to them. See, e.g., NAT. RESOURCES CONSERVATION SERV., ILLINOIS RIVER SUB-BASIN AND EUCHA-SPAVINAW LAKE WATERSHED INITIATIVE PAYMENT RATES (2014), https://www.nrcs.usda.gov/wps/PA_NRCSCconsumption/download?cid=stelprdb1246424&ext=pdf (showing payment rates of more than \$19,000 per acre for certain soil conservation practices). Experimentation with mandatory policies can be more difficult. Abramowicz et al. provide a helpful analysis of the potential ethical and

of experimental rigor and the balance of federal and state involvement are particularly important distinguishing criteria.

Figure 1 visually depicts this framework.¹²⁰ The letters A through I represent nine approaches to policy experimentation that could fall under this framework.¹²¹

Figure 1. Policy Experiments and their Governance Locus

		Experimental rigor		
		Minimal	Moderate	Rigorous
Federal involvement	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.¹²² The courts seem to assume that a hands-off federal

legal issues of this type of experimentation, such as ensuring that experimentation does not take away liberties already enjoyed by individuals (a difficult-to-meet standard due to the challenges associated with properly defining the regulatory “baseline”). They also explore the most analogous legal cases, which have upheld lotteries in which only some individuals obtain the benefits of a policy of limited availability. See Abramowicz et al., *supra* note 1, at 963-73.

¹²⁰ 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.

¹²¹ 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.

¹²² See *supra* notes 30–44 and accompanying text.

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.¹²³ Experimental rigor has never been a focus of judicial discussion, and thus it is hard to discern whether judges envision action in 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. It thus focuses on boxes C, F, and I. But unlike the federalism camps, it often ignores the levels of government at which the experiment plays out.¹²⁴ It therefore 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.¹²⁵ 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.¹²⁶ 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.¹²⁷

¹²³ See, e.g., *Oregon v. Ice*, 555 U.S. 160, 170 (2009).

¹²⁴ See *supra* notes 45–55 and accompanying text.

¹²⁵ 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).

¹²⁶ See Eloise Pasachoff, *Agency Enforcement of Spending Clause Statutes: A Defense of the Funding Cut-Off*, 124 *YALE L.J.* 248, 260–66 (2014) [hereinafter *Agency Enforcement*] (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.

¹²⁷ See, e.g., *FACT SHEET: Data by the People, for the People — Eight Years of Progress Opening Government Data to Spur Innovation, Opportunity, & Economic Growth*, WHITE HOUSE PRESIDENT BARACK OBAMA (Sept. 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).

Meanwhile, real-world practices do sometimes fall within Box A, with neither experimental rigor nor meaningful federal involvement.¹²⁸

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 in Part II. The federal government can design and implement experiments without much state or local assistance and sometimes has done so.¹²⁹ It also has played a coordinating and cooperating role in experimental programs partially staffed by local governments and states.¹³⁰

The following case studies — borrowed from the somewhat disparate fields of agricultural, natural resources, education, and welfare policy — are centrally designed to support our first and primary point, from which further analysis will (we hope) flow. They show that real-world policy experiments, both old and new, have involved the federal government in instigating, implementing, and coordinating experiments. The agricultural example occupies the bulk of our discussion because it best exemplifies the degree and extent to which the federal government can be centrally involved in implementing and designing policy experiments. Our shorter forestry, education, and welfare case studies further highlight some points made in the agricultural context while in other ways providing contrasting examples. There are, of course, numerous other examples of experimentation — including federal involvement in experimentation — from these and other fields,¹³¹ which could merit exploration in further work. But these case studies, we believe, provide a solid foundation for applying the framework we crafted in Part II.A, for they highlight the many governance levels involved in single policy experiments and varying degrees of rigor in experimentation.

Skeptics might argue that our agricultural and forestry case studies are simply unusual subject areas — ones that have more to do with uncontroversial physical science than with policy, and where federal experimenting therefore would come abnormally easily. These skeptics might nod to the Department of Education and Jobs-Plus

¹²⁸ 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 generally Wiseman, *Risk and Response*, *supra* note 115.

¹²⁹ See, e.g., *infra* notes 231–66 and accompanying text (discussing wildfire policy).

¹³⁰ See, e.g., *infra* notes 265–88 and accompanying text (discussing Race to the Top).

¹³¹ See sources cited *supra* note 47.

examples as true “policy” experiments because they involve testing of social policy and are truly “human centric”; they do not involve a physical medium, like soil or trees, that falls more traditionally within the realm of scientific research. But characterizing only these latter examples as true policy experimentation would be wrong. The case studies below show several areas in which physical science and policy happen to be closely linked, but this does not detract from their qualities as *policy* experiments. In our agricultural example, the ultimate goals were not only to teach farmers how to keep soil on their property. Instead, the goals also included experimenting with various methods for ensuring maximum farmer buy-in (which was more of a political experiment¹³²) and informing legal decisions on the criteria for allocating massive sums of federal money.¹³³ 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 and health policy,¹³⁴ financial regulation,¹³⁵ and endangered species protection,¹³⁶ policy decisions grounded in technical research, and often in science, are routine, and pure policy decisions are rather hard to find. Even when experiments use human subjects, and thus might be viewed more as policy-oriented than, say, efforts to choose ideal wildfire policy, they still have a “scientific” or technical bent. Indeed, in proposing truly randomized policy experimentation, Michael Abramowicz, Ian Ayres, and Yair Listokin point to policy experiments that some might argue are difficult to separate from technical economic research — for example, varying policy with respect to short sales of stock and measuring impacts on stock liquidity.¹³⁷ In short, a sharp distinction

¹³² See *supra* note 15 and accompanying text.

¹³³ See *infra* note 219 and accompanying text.

¹³⁴ See, e.g., *Center for Education Policy Research*, HARV. U., <https://cepr.harvard.edu/> (last visited Nov. 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.*, 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).

¹³⁵ See, e.g., *About the Division of Economic and Risk Analysis*, U.S. SEC. & EXCH. COMM’N, <https://www.sec.gov/dera/about> (last visited Nov. 14, 2017); *Economic Research*, BOARD GOVERNORS FED. RES. SYS., <https://www.federalreserve.gov/econres.htm> (last visited Nov. 14, 2017) (providing links to policy-relevant economic research and describing the Fed’s research support efforts).

¹³⁶ 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).

¹³⁷ See Abramowicz et al., *supra* note 1, at 989.

between technical and policy experimentation does not apply to the case studies here — or to the rest of policymaking.

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.¹³⁸ 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."¹³⁹ 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.¹⁴⁰ Agricultural interests and their elected supporters have aggressively deployed federalist arguments in their opposition to anything that verges on land use regulation.¹⁴¹ Yet a rigorous, long-lasting, and far-reaching set of policy experiments resides in the area of agricultural policy. Specifically, federal policy was designed to reduce the loss of valuable agricultural topsoil through erosion on millions of acres of private farmland and to persuade farmers to implement this policy on millions of individual farms and ranches.¹⁴² These experiments began in the 1920s, grew in

¹³⁸ Indeed, when President Roosevelt first devoted some Civilian Conservation Corps funds to controlling soil erosion on private lands this was a controversial move due to "[c]oncern about the public's objections to expenditures of federal funds on private lands." Helms, *The Civilian Conservation*, *supra* note 15, at 184.

¹³⁹ *Solid Waste Agency v. U.S. Army Corps of Eng'rs*, 531 U.S. 159, 174 (2001); *Fed. Energy Regulatory Comm'n v. Mississippi*, 456 U.S. 742, 768 n.30 (1982). The latter statement is inaccurate; most land use regulation is done by local governments.

¹⁴⁰ See, e.g., *Rapanos v. United States*, 547 U.S. 715, 738 (2006) (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*, 531 U.S. at 174; *Fed. Power Comm'n v. E. Ohio Gas Co.*, 338 U.S. 464, 489 (1950) (noting states' utility policies). 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 3, at 1752-75.

¹⁴¹ See, e.g., Dave Owen, *Urbanization, Water Quality, and the Regulated Landscape*, 82 U. COLO. L. REV. 431, 477 & nn.288-89 (2011) (compiling quotes from legislative hearings).

¹⁴² 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. And not all measures are voluntary. Farmers

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

The federal government was concerned about the loss of soil through erosion well before the notorious dust storms of the 1930s. As early as 1917, the United States Department of Agriculture (“USDA”) sent an agricultural engineer to study methods of preventing soil loss in the southeastern United States.¹⁴³ After visiting farms throughout five states, this engineer found one farm, in particular, that seemed to produce promising results, and he suggested that the techniques deployed on this farm be applied with needed modifications, elsewhere.¹⁴⁴ But as historians tell it, the USDA’s true 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.”¹⁴⁵ This document alerted the public, as well as congressmen

who do not implement USDA conservation practices are ineligible for a variety of federal funds. See, e.g., *Environmental Quality Incentives Program*, USDA NAT. RESOURCES CONSERVATION SERV., <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/programs/financial/eqip/> (last visited Sept. 9, 2018) (describing farmers selected for funding based specifically on the conservation practices they implement).

¹⁴³ See G. E. MARTIN, TERRACING IN OKLAHOMA 6-7 (U.S. Dep’t of Agric., Circular No. 218, 1930).

¹⁴⁴ See *id.*

¹⁴⁵ H.H. BENNETT & W.R. CHAPLINE, SOIL EROSION A NATIONAL MENACE (U.S. Dep’t of Agric., Circular No. 33, 1928); see U.S. DEP’T OF AGRIC., 1980 APPRAISAL PART I: SOIL, WATER, AND RELATED RESOURCES IN THE UNITED STATES STATUS, CONDITION, AND TRENDS 9 (1981) [hereinafter 1980 APPRAISAL]; Helms, *Two Centuries*, *supra* note 15, at 25.

and federal agencies, to the problem of soil loss caused by improperly tilled and managed soils that blew or washed away.¹⁴⁶ The authors estimated that 1.5 billion tons of soil¹⁴⁷ and at least 126 billion “pounds of plant-food material” were lost annually.¹⁴⁸ 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,¹⁴⁹ and even counties and individual farms, such as “an apple orchard near Lookout Mountain in northeastern Kansas,”¹⁵⁰ and “one place a few miles south of Troy, Kans.”¹⁵¹ 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.”¹⁵²

Sadly, they were right. During the 1930s, a combination of drought, wind, and ill-advised policies (many of them federal¹⁵³) 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.¹⁵⁴ As historian Donald Worster has written:

The story of the southern plains in the 1930s 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,

¹⁴⁶ The report focused on water erosion but noted in footnote 1 that “[m]uch damage is also done by wind erosion.” BENNETT & CHAPLINE, *supra* note 145, at 1 n.1.

¹⁴⁷ *Id.* at 5. 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.* at 4. See also U.S. DEP’T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 9 (describing this estimate).

¹⁴⁸ BENNETT & CHAPLINE, *supra* note 145, at 2.

¹⁴⁹ See *id.* at 11.

¹⁵⁰ *Id.* at 10.

¹⁵¹ *Id.*

¹⁵² *Id.* at 22.

¹⁵³ See U.S. DEP’T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 7-9 (describing the acts that encouraged farmers who had recently emigrated from Europe and were unfamiliar with U.S. landscapes, or how to properly till them, to acquire property at low or even no cost).

¹⁵⁴ See DONALD WORSTER, DUST BOWL: THE SOUTHERN PLAINS IN THE 1930S 13 (1979).

of fine powder caking one's lips, of springtime turned to despair, of poverty eating into self-confidence.¹⁵⁵

While real-life Tom Joads fled west, and while John Steinbeck and other writers penned their laments,¹⁵⁶ 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 one billion [privately-owned] acres” scattered around the United States.¹⁵⁷ Such a dispersed and localized problem also might seem rather ill-suited for a federal response. Nevertheless, a broad federal program emerged. The program had two goals: to identify the best soil conservation practices, and to determine the best ways to persuade millions of farmers to adopt the practices.

The USDA began its implementation efforts by funding and establishing “soil experiment stations,” which it deployed around the United States.¹⁵⁸ 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.¹⁵⁹ 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.¹⁶⁰

Congress and the executive branch created new administrative structures to advance the soil management program. First, in 1933,

¹⁵⁵ *Id.*

¹⁵⁶ See generally JOHN STEINBECK, *THE GRAPES OF WRATH* (1939).

¹⁵⁷ U.S. DEP'T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 9 (providing statistics from the 1930s). As of 1980, there were approximately “1.5 billion acres of nonfederally owned” U.S. land, twenty-seven percent of which was devoted to rangeland for livestock, an equal percentage of which was crop land, and nine percent of which was pasture. *Id.* at 2.

¹⁵⁸ See DOUGLAS HELMS, HUGH HAMMOND BENNETT AND THE CREATION OF THE SOIL EROSION SERVICE 2 (2008), https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcs143_021210.pdf [hereinafter CREATION]. These were later renamed soil conservation experiment stations. *Id.*

¹⁵⁹ See Helms, *Two Centuries*, *supra* note 15, at 26 (“In the new Soil Erosion Service, Bennett located soil conservation projects in the watersheds near erosion experiment stations so that the directors of the stations could utilize the research information.”); Bennett, *Conservation Work*, *supra* note 15, at 10-11 (noting “120 field stations scattered throughout the country” in 1940).

¹⁶⁰ See Bennett, *Conservation Work*, *supra* note 15, at 14-15.

Congress funded and created a new agency within the Department of the Interior called the Soil Erosion Service, with Bennett at its head.¹⁶¹ 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 USDA. The goal of the service was to "provide permanently for the control and prevention of soil erosion."¹⁶² Diverse research, demonstration, and other projects would accomplish these objectives. Specifically, the act directed the Secretary of Agriculture to "conduct surveys, investigations, and research relating to the character of soil erosion and the preventive measures needed;" 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.¹⁶³ 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,¹⁶⁴ state forestry boards,¹⁶⁵ 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 universities¹⁶⁶) and funded state agricultural experiment stations, "having due regard to the varying conditions and needs of the respective States or Territories."¹⁶⁷ Based on work at these experiment

¹⁶¹ See HELMS, CREATION, *supra* note 158, at 11-12. In 1935, Congress established the Soil Conservation Service ("SCS") within the USDA, and thus the SES became SCS. Soil Conservation and Domestic Allotment Act of 1935, ch. 85, § 5, 49 Stat. 163, 164.

¹⁶² Soil Conservation and Domestic Allotment Act § 1; see also Douglas Helms, SCS: 50 Years Young, FARMER, Mar. 16, 1985, at 48, 48 [hereinafter SCS].

¹⁶³ Soil Conservation and Domestic Allotment Act §1.

¹⁶⁴ See, e.g., MARTIN, *supra* note 143 (This source is a publication by the Oklahoma Agricultural and Mechanical College and United States Department of Agriculture, Cooperating, and the Extension Service County Agent in Stillwater, Oklahoma).

¹⁶⁵ See, e.g., CAL. STATE BD. OF FORESTRY, REPORT TO THE LEGISLATURE ON SENATE CONCURRENT RESOLUTION NO. 27 (LEGISLATURE OF 1921), at 4 (1923).

¹⁶⁶ A land grant university is a university built on land donated by the federal government to a state. See U.S. DEP'T OF AGRIC. & NAT'L INST. OF FOOD & AGRIC., NIFA LAND-GRANT COLLEGES AND UNIVERSITIES (2014), https://nifa.usda.gov/sites/default/files/resource/lgu_map_6_25_2014.pdf.

¹⁶⁷ Hatch Act of 1887, ch. 314, §§ 1-2, 24 Stat. 440, 440-41.

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 examples from these state experiment stations. For example, it contrasted two approaches at the "Spur substation of the Texas Agricultural Experiment Station,"¹⁶⁸ and also drew upon reports from similar experiment stations in Missouri and North Carolina.¹⁶⁹ 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,¹⁷⁰ in part because the federal government recognized that securing effort and commitments from individual farmers would be easier with state and local support.¹⁷¹

The effort to involve state and local governments in soil conservation practices demonstrated a key focus of the policy experiment — determining how to persuade farmers to adopt soil conservation practices on their individual properties. While Hugh Hammond Bennett pushed for soil experiment stations and federal provision of materials and labor to individual farmers, the Assistant Secretary of Agriculture championed more federal collaboration with state and local government units, so that farmers would feel that they had a more direct stake in the project.¹⁷²

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

¹⁶⁸ See BENNET & CHAPLINE, *supra* note 145, at 6.

¹⁶⁹ See *id.*

¹⁷⁰ See Soil Conservation and Domestic Allotment Act of 1935, ch. 83, § 1(3), 49 Stat. 163, 163 (allowing the USDA to "cooperate or enter into agreements with, or to furnish financial or other aid to, any agency, governmental or otherwise, or any person . . . for the purposes of this Act").

¹⁷¹ As one source explains, "Widespread local leadership was required to motivate and guide" the thousands of private landowners operating farms and ranches. U.S. DEP'T OF AGRIC., 1980 APPRAISAL PART I: SOIL, WATER, AND RELATED RESOURCES IN THE UNITED STATES: STATUS, CONDITION, AND TRENDS 14 (1981); see also Douglas Helms, *Soil and Soil Conservation*, in *ENCYCLOPEDIA OF SOUTHERN CULTURE* 361, 362 (Charles Reagan Wilson & William Ferris eds., 1989).

¹⁷² See Helms, *Two Centuries*, *supra* note 15, at 26 (describing M.L. Wilson's push for soil conservation districts and arguing that Wilson believed that the soil conservation districts would allow conservation practices to "spread nationwide," even after the CCC no longer operated, and that farmers who were part of conservation districts "would be more interested and involved in planning and carrying out of the work").

conservation districts.¹⁷³ This model law, adopted by all states in some form,¹⁷⁴ provided for federal involvement in these districts and enabled close collaboration between the districts and the federal government.¹⁷⁵ Conservation districts continue to operate and exercise

¹⁷³ Under the model act, each state formed a soil conservation district as a “governmental subdivision” of the state. U.S. DEP’T OF AGRIC. SOIL CONSERVATION SERV., A STANDARD STATE SOIL CONSERVATION DISTRICTS LAW 3 (1936) [hereinafter STANDARD CONSERVATION LAW]. The act provided that the state should first establish a soil conservation committee of which a federal USDA representative could also serve as a member. *See 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. *Id.* at 7. 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, thus creating a five-supervisor governing body of the district. *Id.* at 10, 15.

¹⁷⁴ *See* U.S. DEP’T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 14.

¹⁷⁵ 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, and 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. U.S. DEP’T OF AGRIC., STANDARD CONSERVATION LAW, *supra* note 173, at 15-17. For a description of the local, grassroots nature of these districts and their connection to national goals, see *Hearing on S. 2549*, 73d Cong. 3 (1933) (statement of Hugh Hammond Bennett) (transcript available in Iowa State University Special Collections and University Archives Dept.) (explaining that the districts “are essentially large groups of farmers and ranchers working together in neighborly fashion and helping one another under a formalized type of organization” and observing that the districts have “been effectively utilized in advancing the national program of soil conservation”). For a description of federal involvement in these districts, see, e.g., Hugh Hammond Bennett, Chief, Soil Conservation Serv., Address Before the National Association of County Agricultural Agents: The Cooperative Approach to Land Problems 2 (Dec. 6, 1939) (transcript available at Iowa State University Special Collections and University Archives Dept.) (noting that “most of the districts so far organized have already called on both the Extension Service and the Soil Conservation Service for help”); Bennett, Conservation Work, *supra* note 15, at 9 (noting in 1940 that the federal “Soil Conservation Service . . . has adopted a definite policy of helping the districts wherever possible, and right now, we are cooperating with 263 districts comprising 159 million acres in 30 States,” including assistance with surveys and conservation projects).

these powers. Three thousand districts¹⁷⁶ now emphasize three soil conservation strategies¹⁷⁷ initially championed by Hugh Hammond Bennett, applying and differentiating these practices across “nearly all private rural land” in the United States.¹⁷⁸

In addition to this reliance on state and local structures, the new federal program continues to take advantage of its own geographic decentralization. The Natural Resources Conservation Service (“NRCS”) — the same agency originally headed by Bennett¹⁷⁹ — has built from the original strategies emphasized by Bennett and has developed more formal, detailed federal standards for soil conservation (and other conservation practices).¹⁸⁰ But these are not uniform standards. The NRCS has field offices in all fifty states¹⁸¹ and operates local service centers,¹⁸² and these offices tailor federal conservation standards to local conditions.¹⁸³ Farmers who meet the local conservation standards receive federal funding for implementing soil conservation practices.¹⁸⁴

¹⁷⁶ *About NACD*, NAT’L ASS’N CONSERVATION DISTRICTS, <http://www.nacdnet.org/about-nacd/> (last visited Sept. 10, 2018).

¹⁷⁷ *Soil*, NAT’L ASS’N CONSERVATION DISTRICTS, <http://www.nacdnet.org/about-nacd/what-we-do/soil/> (last visited Sept. 10, 2018) (noting that the districts focus on “crop rotation, cover crops and no- or minimum tillage systems”).

¹⁷⁸ U.S. DEP’T OF AGRIC., 1980 Appraisal, *supra* note 145, at 14.

¹⁷⁹ 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 NRCS*, USDA NAT. RESOURCES CONSERVATION SERV., <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/about/history/> (last visited Sept. 10, 2018).

¹⁸⁰ *See Conservation Practices*, USDA NAT. RESOURCES CONSERVATION SERV., https://www.nrcs.usda.gov/wps/portal/nrcs/detailfull/national/technical/cp/ncps/?cid=nrcs143_026849 (last visited Sept. 10, 2018).

¹⁸¹ *See State Offices Directory*, USDA NAT. RESOURCES CONSERVATION SERV., <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/contact/states/> (last visited Sept. 10, 2018) (showing that employees of the field offices have “USDA.gov” e-mail addresses).

¹⁸² *Id.*

¹⁸³ 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*, *supra* note 180.

¹⁸⁴ *See National Conservation Practice Standards*, USDA NAT. RESOURCES CONSERVATION SERV., <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/technical/cp/ncps/> (last visited Sept. 10, 2018) (showing that farmers must follow state, not national, conservation practice standards). Farmers wishing to receive EQIP funds must prepare and have approved a Conservation Activity Plan, which is “developed for producers to identify conservation practices needed to address a specific natural resource need.” *FY 2016 EQIP Conservation Activity Plan (CAP)*, USDA NAT. RESOURCES CONSERVATION SERV.,

Federal soil conservation policy also utilizes another national program called the Cooperative Extension Service, which has been active since 1914.¹⁸⁵ 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.¹⁸⁶ These extensions, through federal-state partnerships, conduct localized research and then disseminate it in order to encourage farmers to adopt practices with demonstrated beneficial results.¹⁸⁷ Land-grant universities in each state house primary state extension offices staffed with USDA employees, and each state also has a “network of local or regional offices.”¹⁸⁸ The Extension Service has “an office in or near most of the nation’s approximately 3,000 counties.”¹⁸⁹

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.¹⁹⁰ And the program’s reach is striking. The data from just the first year of the Civilian Conservation Corps’ work shows 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,

<https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/programs/financial/equip/?cid=nrcseprd401472> (last visited Sept. 10, 2018). As noted above, the USDA defines national conservation practices that are then localized through field specific manuals. *Conservation Practices*, *supra* note 180 (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”).

¹⁸⁵ *Cooperative Extension History*, USDA NAT’L INST. FOOD & AGRIC., <https://nifa.usda.gov/cooperative-extension-history> (last visited Sept. 10, 2018).

¹⁸⁶ See NAT’L INST. OF FOOD & AGRIC., U.S. DEP’T OF AGRIC., 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); *Extension*, USDA: NAT’L INST. FOOD & AGRIC., <https://nifa.usda.gov/extension> (last visited Sept. 10, 2018).

¹⁸⁷ See *How We Work*, USDA NAT’L INST. FOOD & AGRIC., <https://nifa.usda.gov/how-we-work> (last visited Sept. 10, 2018).

¹⁸⁸ *USDA Local Offices*, USDA, <https://www.outreach.usda.gov/USDALocalOffices.htm> (last visited Sept. 10, 2018).

¹⁸⁹ *Cooperative Extension History*, *supra* note 185.

¹⁹⁰ See *USDA Local Offices*, *supra* note 188.

and building “420,000 erosion control check-dams” and 4,000 miles of fence.¹⁹¹ By 1940, the federal Soil Conservation Service had deployed demonstration projects on “about 70,600 farms comprising roughly 20,000,000 acres” in more than forty states.¹⁹² 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 the 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 funding.¹⁹³ The agency also held “about 9,000 public participation meetings, attended by over 164,000 persons.”¹⁹⁴ 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 different hypotheses about how to handle soil conservation policy. 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.¹⁹⁵ 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

¹⁹¹ U.S. DEP’T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 11.

¹⁹² See Bennett, Conservation Work, *supra* note 15, at 5, 8.

¹⁹³ U.S. DEP’T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 1, 35 (describing the Iowa State-led National Resource Inventories).

¹⁹⁴ *Id.* at 5.

¹⁹⁵ See HELMS, CREATION, *supra* note 158, at 9-10.

similar conditions.¹⁹⁶ 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 the unique soils, rainfall levels, and other conditions around the country, and with farmers' differing preferences for and abilities to implement control techniques.¹⁹⁷ 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.¹⁹⁸

Federal agencies did this testing through a deliberate program of differentiation. In part, this differentiation flowed from the geography of the new experiment locations. 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 climate and would help to stabilize grazing areas.¹⁹⁹ And the Soil Conservation Service demonstration projects in the Pacific Northwest promoted winter cover crops because of the heavy winter rainfall in this area.²⁰⁰ 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."²⁰¹ Another central purpose in locating the experiment stations around the country was to demonstrate and persuade: the federal government aimed to convince farmers of the efficacy of successful conservation practices.

In addition to varying soil conservation approaches among localities, USDA also varied its approaches within localities, testing

¹⁹⁶ See *infra* notes 215-23 and accompanying text.

¹⁹⁷ See Helms, *The Civilian Conservation*, *supra* note 15, at 185 ("To Bennett's thinking, erosion had to be reduced through a coordinated effort that allowed farmers to continue farming without reducing income. . . . Bennett's years of observation had taught him to be wary of single-method approaches that could create new problems while mitigating existing ones."); *id.* at 186 (noting that farmers were amenable to practices such as planting alfalfa in strip crops, which helped to reduce erosion, but that they struggled to implement these types of practices due to the cost of seed and fertilizer).

¹⁹⁸ See, e.g., Bennett, *Conservation Work*, *supra* note 15, at 5-8, 10 (describing the demonstration projects in place as of 1940, which covered approximately "70,600 farms comprising roughly 20,000,000 acres," as well as "120 field stations").

¹⁹⁹ See Helms, *Two Centuries*, *supra* note 15, at 26.

²⁰⁰ See Helms, *The Civilian Conservation*, *supra* note 15, at 188.

²⁰¹ Helms, *SCS*, *supra* note 162, at 48. In other circumstances, the Corps applied national strategies deemed to be effective in most regions, such as fencing cattle out of areas prone to erosion. See U.S. DEP'T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 11.

alternative hypotheses and then emphasizing successful approaches. For example, at the San Dimas watershed experiment station in California, Civilian Conservation Corps (“CCC”) 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.²⁰² 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.²⁰³ 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.”²⁰⁴ 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).²⁰⁵

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 deliberately created “control plots”²⁰⁶ to account for variables other than soil conservation practices that affect erosion.²⁰⁷ Similarly, Lowdermilk pioneered techniques to

²⁰² See Douglas Helms, *Walter Lowdermilk’s Journey: Forester to Land Conservationist*, 8 ENVTL. REV. 132, 138 (1984) [hereinafter *Forester to Land Conservationist*].

²⁰³ *Id.*

²⁰⁴ *Id.* (quoting Walter C. Lowdermilk, *Further Studies of Factors Affecting Surficial Run-Off and Erosion*, 1929 PROC. INT’L CONGRESS FORESTRY EXPERIMENT STATIONS 625 (1929)).

²⁰⁵ *Id.* at 139.

²⁰⁶ U.S. DEP’T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 10.

²⁰⁷ Lowdermilk was inspired in part by 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 INST. OF EXPERIMENTAL FORESTRY, GUIDE TO THE EXCURSIONS OF THE INTERNATIONAL CONGRESS OF FORESTRY EXPERIMENTAL STATIONS 1-19 (1929); see Helms, *Forester to Land Conservationist*, *supra* note 202, at 138 (noting Lowdermilk’s participation in the Stockholm meeting).

control the many factors that confound efforts to measure the causes of soil erosion at watershed scales.²⁰⁸ 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.²⁰⁹ 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. Other experimentation stations deployed similar control techniques to isolate variables and identify the most effective erosion control practices.²¹⁰

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, they required the USDA to collect and analyze data on the effectiveness of soil conservation practices.²¹¹ Staff at 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 agencies²¹² and in peer-reviewed science journals.²¹³ By 1940, Bennett noted that

²⁰⁸ 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 Helms, *Forester to Land Conservationist*, *supra* note 202, at 136 (“In an open [watershed] setting there were too many variables.”).

²⁰⁹ Walter C. Lowdermilk, *Studies in the Role of Forest Vegetation in Surficial Run-Off and Soil Erosion*, 12 AGRIC. ENGINEERING 107, 108 (1931).

²¹⁰ See HELMS, CREATION, *supra* note 158, at 6.

²¹¹ 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. For a discussion of the expansion of USDA with special attention to conservation issues beyond the preservation of soil, see C.W. Richardson et al., *The Conservation Effects Assessment Project Benchmark Watersheds: Synthesis of Preliminary Findings*, 63 J. SOIL & WATER CONSERVATION 590, 590 (2008).

²¹² See, e.g., BUREAU OF PUB. RDS., U.S. DEP’T OF AGRIC., SECOND PROGRESS REPORT ON EROSION AND RUN-OFF EXPERIMENTS IN PIEDMONT, NORTH CAROLINA (1929) (written “in cooperation with the North Carolina Department of Agriculture”).

²¹³ See, e.g., Lowdermilk, *supra* note 209. Publication of experimental data and results has continued through the present. See generally 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 CONSERVATION 339 (2008); J. Cho et

the Soil Conservation Service “had made hundreds of thousands of quantitative measurements of soil and water losses from rains at different intensities and durations,” with 300,000 measurements taken at just one experiment station over an eleven-year period.²¹⁴

Widespread data collection and analysis in this area continue to the present day. The 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 to 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.²¹⁵ Independently and in collaboration with state and local governments, the Soil Conservation Service has completed hundreds of detailed studies of reduced soil erosion.²¹⁶ The results from these studies provide quantitative support for effective soil conservation practices²¹⁷ and

al., *Water Quality Effects of Simulated Conservation Practice Scenarios in the Little River Experimental Watershed*, 65 J. SOIL & WATER CONSERVATION 463 (2010); G.W. Feyereisen et al., *Long-term Stream Chemistry Trends in the Southern Georgia Little River Experimental Watershed*, 63 J. SOIL & WATER CONSERVATION 475 (2008); R.D. Harmel et al., *Conservation Effects Assessment Project Research in the Leon River and Riesel Watersheds*, 63 J. SOIL & WATER CONSERVATION 453 (2008); Douglas L. Karlen et al., *Is No-Tillage Enough? A Field-Scale Watershed Assessment of Conservation Effects*, 7 ELECTRONIC J. INTEGRATIVE BIOSCIENCES 1 (2009); R.A. Kuhnle et al., *Conservation Practice Effects on Sediment Load in the Goodwin Creek Experimental Watershed*, 63 J. SOIL & WATER CONSERVATION 496 (2008); R.N. Lerch et al., *Overview of the Mark Twain Lake/Salt River Basin Conservation Effects Assessment Project*, 63 J. SOIL & WATER CONSERVATION 345 (2008); G.W. McCarty et al., *Water Quality and Conservation Practice Effects in the Choptank River Watershed*, 63 J. SOIL & WATER CONSERVATION 461 (2008); A. Simon & L. Klimetz, *Relative Magnitudes and Sources of Sediment in Benchmark Watersheds of the Conservation Effects Assessment Project*, 63 J. SOIL & WATER CONSERVATION 504 (2008).

²¹⁴ *Hearing on S. 2549*, *supra* note 175, at 8.

²¹⁵ Maurice J. Mausbach & Allen R. Dedrick, *The Length We Go: Measuring Environmental Benefits of Conservation Practices*, 59 J. SOIL & WATER CONSERVATION 96A, 100A (2004).

²¹⁶ See, e.g., R.F. Cullum et al., *Effects of Conservation Reserve Program on Runoff and Lake Water Quality in an Oxbow Lake Watershed*, 5 J. INT'L ENVTL. APPLICATION & SCI. 318 (2010); Karlen et al., *supra* note 213, at 1; R.A. Kuhnle et al., *supra* note 213, at 496.

²¹⁷ 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 211, at 590. And in 2002, when Congress significantly increased funding for agricultural conservation practices, Congress also mandated a study with similar quantification goals, leading to the creation of the USDA’s Conservation Effects Assessment Project. *Id.* at 590-91. The CEAP, implemented through the USDA’s Agricultural Research Service and cooperative extensions, *id.*, will provide “in-depth quantification of water quality and soil quality impacts of conservation practices at the

inform and improve detailed and frequently-updated national models. Those models then provide broader predictions about the effectiveness of soil conservation practices where a particular watershed has not been studied.²¹⁸

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.²¹⁹ 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.²²⁰ They additionally conduct farmer surveys and “other one-time special studies” to supplement the sample information.²²¹ Rarely are agencies 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 Agricultural Resource Service, which employs 2000 scientists,²²² collects detailed data on the impacts of various erosion control practices and other conservation practices

local level,” such as reduced nutrients, pesticides, and sediments entering water bodies. Lisa F. Duriancik et al., *The First Five Years of the Conservation Effects Assessment Project*, 63 J. SOIL & WATER CONSERVATION 185A, 185A, 187A (2008); see also Mausbach & Dedrick, *supra* note 215, 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 211, 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.*

²¹⁸ 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 AGRIC. & BIOLOGICAL ENGINEERS 1211, 1211-12 (2007) (describing the models).

²¹⁹ USDA, 2012 NATURAL RESOURCES INVENTORY SUMMARY REPORT 7-2 (2015), https://www.nrcs.usda.gov/Internet/FSE_DOCUMENTS/nrcseprd396218.pdf [hereinafter 2012 RESOURCES INVENTORY]. 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: NAT. RESOURCES CONSERVATION SERV., https://www.nrcs.usda.gov/wps/portal/nrcs/detail/national/technical/nra/ceap/na/?cid=nrcs143_014144 (last visited Oct. 21, 2018).

²²⁰ USDA, 2012 RESOURCES INVENTORY, *supra* note 219, at 7-2 to 7-3.

²²¹ *Id.* at 1-2.

²²² *About ARS*, USDA, <https://www.ars.usda.gov/about-ars/> (last visited Oct. 17, 2018).

incentivized through federal programs, and it has done so for quite a long time.²²³

The data also do not just sit on shelves. Many of the scientific reports describing the 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.²²⁴ 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.²²⁵

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.²²⁶ It has followed this practice since the formation of the

²²³ 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 218, 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. *Id.* 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 1217-24. 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.

²²⁴ See, e.g., Cullum et al., *supra* note 216, 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.”); Kuhnle et al., *supra* note 213, 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).

²²⁵ See, e.g., Richardson et al., *supra* note 211, 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).

²²⁶ See, e.g., Karlen et al., *supra* note 213, at 1-2 (noting that “although previous studies had suggested that reduced tillage and extended cropping systems would be more sustainable than the continuous corn grown on the site since the early 1960s, quantitative evidence was lacking”).

first experiment stations, when government scientists tested results on farms near the research stations.²²⁷ In a 1930 cooperative extension publication from Oklahoma, for example, the Extension Engineer noted that for the terracing of soil, the USDA sent an agricultural engineer to study effective practices in the Southeast. The engineer published his findings in 1917, and, from studies of “the Carolinas, Georgia, Alabama, and Mississippi,” “it was found that . . . [terracing efforts] on the farm of Mr. P. H. Mangum near Wake Forest, North Carolina, most nearly gave the desired results and were therefore to be recommended to other farmers.”²²⁸ State extension services then applied and modified this technique elsewhere, “carefully” noting when the terraces failed and why.²²⁹ In addition to retesting hypotheses, the USDA uses experiments throughout the United States to validate and refine its models.²³⁰

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.

C. “Playing with Fire”²³¹

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.²³² 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 Part 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.

²²⁷ See Gassman et al., *supra* note 218, at 1215.

²²⁸ MARTIN, *supra* note 143, at 6-7.

²²⁹ *Id.* at 7.

²³⁰ Richardson et al., *supra* note 211, at 591.

²³¹ John McLaughlin, *Restoring Fire to the Environment in Sequoia and Kings Canyon National Parks*, 12 PROC. OF THE TALL TIMBERS FIRE ECOLOGY CONF. 391, 394 (1973) (“I suppose one could say we are playing with fire . . .”).

²³² See generally Stephen J. Pyne, *Between Two Fires: The Past and Future of Fire in America*, 18 PA. ST. ENVTL. L. REV. 129 (2010) (providing an overview of this history).

For most of the twentieth century, American wildfire policy was dominated by a single, clear idea: put it out, and quickly.²³³ 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,²³⁴ 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.²³⁵ By the 1930s, that ambition hardened into a simple rule: every fire should be out by 10:00 AM the day after it was spotted.²³⁶ The United States Forest Service led the crusade against wildfire, but other federal agencies followed suit, as did many states.²³⁷ Only in the southeast did widespread tolerance for both wildfire and prescribed burning remain.²³⁸

From the outset, however, some forest managers expressed doubt about this policy, and by the 1960s, the doubts were beginning to swell.²³⁹ 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.²⁴⁰ It could not be suppressed forever, they argued, for fuel buildup would only increase, and the fires that eventually did burn would be

²³³ See Jan W. van Wagtenonk, *The Evolution of National Park Service Fire Policy*, 52 FIRE MGMT. NOTES 10, 10-14 (1991).

²³⁴ This desire was not shared by the National Park Service, but it was central to Forest Service culture.

²³⁵ 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 RES. 145, 148-52 (2004) (describing this evolution).

²³⁶ Stephens & Sugihara, *supra* note 235, at 434.

²³⁷ See Jim Brenner & Dale Wade, *Florida's Revised Prescribed Fire Law: Protection for Responsible Burners*, in PROCEEDINGS OF FIRE CONF. 2000: THE FIRST NATIONAL CONGRESS ON FIRE ECOLOGY, PREVENTION, AND MANAGEMENT 132, 132 (Krista E.M. Galley et al. eds., 2003) ("During the early part of the 20th century, the use of fire as a management tool by state and federal agencies was seen by the forestry community as anathema.").

²³⁸ See *id.* at 133 ("Florida has led the nation in acreage treated with prescribed fire every year since records have been kept.").

²³⁹ See 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."); Stephens & Sugihara, *supra* note 235, at 433 (describing opposition in the 1880s); Jan W. van Wagtenonk, *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").

²⁴⁰ See van Wagtenonk, *supra* note 233, at 11-12.

enormous, dangerous, and ecologically catastrophic.²⁴¹ 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.²⁴² 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.²⁴³ In Sequoia, Kings Canyon, and Yosemite National Parks, managers began experimenting with controlled burns.²⁴⁴ 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.²⁴⁵ 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 hypotheses from it, and tested those hypotheses through carefully observed trials.²⁴⁶ Rather than just running their experiments once, they repeated them at other sites.²⁴⁷ And because adjacent forests 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

²⁴¹ 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).

²⁴² *Id.* at 533; see also A. STARKER LEOPOLD ET AL., *WILDLIFE MANAGEMENT IN THE NATIONAL PARKS* 12 (1963).

²⁴³ 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 CONF.* 45, 46 (1976). But until the early 1970s, the Everglades experiment remained exceptional. *Id.*

²⁴⁴ 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 *PROC. TALL TIMBERS FIRE ECOLOGY CONF.* 377, 380 (1973); see also Jerry Lester & Eleanore Browning, *UC Foresters Aid Fire Ecology Program at Yosemite Park*, 25 *CAL. AGRIC.* 1, 3 (1971).

²⁴⁵ Stephens & Ruth, *supra* note 241, at 533.

²⁴⁶ See Kilgore, *supra* note 243, at 50-51 (describing the importance of "carefully controlled laboratory and field studies" carried out in conjunction with university and Forest Service scientists).

²⁴⁷ See *id.* at 47 (describing burning experiments in a variety of different ecological settings).

scientists used to generate a huge outpouring of written research studies.²⁴⁸ Land managers also began hosting annual conferences to discuss their findings.²⁴⁹ 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.²⁵⁰

That transition was not always smooth, partly because this was also an experiment in public acceptance of modified fire practices. 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 the agencies' embrace of fire.²⁵¹ Today, many fire scientists agree that the embrace — particularly by the Forest Service — still is not nearly enthusiastic enough.²⁵² 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.²⁵³ 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 testing, observation and data collection, and analysis and dissemination of outcomes; 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.²⁵⁴ Federal lands

²⁴⁸ For a compilation of papers, see *Sequoia and Kings Canyon National Parks: Fire Research Papers*, NAT'L PARK SERV., https://www.nps.gov/seki/learn/nature/fic_fireres.htm (last updated July, 22 2016) [hereinafter *Sequoia and Kings Canyon*].

²⁴⁹ Many of the papers cited here come from the proceedings of the Tall Timbers Fire Ecology Conference.

²⁵⁰ See Kilgore, *supra* note 243, at 47-48 (describing the spreading of the Park Service's let-it-burn policy); Stephens & Ruth, *supra* note 241, at 534 (describing changes to federal fire policies).

²⁵¹ See van Wagtenonk, *supra* note 233, at 8, 11.

²⁵² 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>.

²⁵³ See *id.* at 6-7.

²⁵⁴ See, e.g., *Sequoia and Kings Canyon*, *supra* note 248. 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 Lauren Wishnie, *Fire and Federalism*, 17 *N.Y.U. ENVTL. L.J.* 1006,

also were the initial focus of policy reform.²⁵⁵ 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.²⁵⁶ 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.²⁵⁷ Some of the scientists who helped instigate the process also worked at state universities.²⁵⁸ But many state policies, particularly in the West, still reflect the old suppress-everything ethos, as do private land management strategies.²⁵⁹

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.²⁶⁰ But those explanations, though accurate,

1015 (2008) (noting “the preeminent position of the federal agencies” in wildfire policy, though also discussing state involvement).

²⁵⁵ See Stephens & Ruth, *supra* note 241, at 536. States also have become increasingly involved in fire planning, particularly for “urban-wildland interface” areas where dispersed rural settlement brings people into fire-prone areas.

²⁵⁶ 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.”).

²⁵⁷ W. James Barry & R. Wayne Harrison, *Prescribed Burning in the California State Park System*, 1997 PROC. SYMP. ON FIRE CAL. ECOSYSTEMS: INTEGRATING ECOLOGY, PREVENTION & MGMT. 203, 204-06. 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.

²⁵⁸ See Kilgore, *supra* note 243, at 46.

²⁵⁹ See Karen Bradshaw, *A Modern Overview of Wildfire Law*, 21 FORDHAM ENVTL. L. REV. 445, 452-53 & n.30 (2010) (describing differences between federal and state policies); Michelle M. 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, no. 3, 2017, at 7 (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 accumulation” and describing how Oregon state law emphasizes complete fire suppression and slowed the pace of federal reform).

²⁶⁰ See Busenberg, *supra* note 235, at 149-50 (describing this federal role and its

are incomplete. Private forests (which states regulate) and state-owned forests are also abundant and sometimes are more fire-prone than nearby federal lands.²⁶¹ State and private owners therefore had incentives to undertake similar reforms, even though most did not do so.²⁶² Similarly, Congress, which funds a shockingly expensive fire suppression program, had ample reason to pursue innovation, but did not.²⁶³ 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, “[National Park Service] 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.”²⁶⁴

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.

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

consequences).

²⁶¹ See, e.g., Steen-Adams et al., *supra* note 259 (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”).

²⁶² 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 costly settlements of litigation over forest fires that began on private land).

²⁶³ See Bradshaw, *supra* note 259, 449-50 (providing distressing statistics on the costs of firefighting).

²⁶⁴ van Wagendonk, *supra* note 233, at 14.

to the Top.²⁶⁵ Like the federal experiments with soil and fire, Race to the Top was a federally-instigated program designed to spur policy innovation.²⁶⁶ But unlike the soil and fire programs, the field staff implementing the program were state and local teachers and administrators, not federal employees.²⁶⁷ Race to the Top thus exploited, rather than worked around, the political structures of federalism.²⁶⁸

The Race to the Top program sought to improve the quality of primary and secondary school education across the United States.²⁶⁹ Its creators had ambitious ideas about improving teacher training, school administration, transparency, and accountability.²⁷⁰ But unlike the Park Service or the Forest Service, which could experiment on federal lands, the Department of Education (“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.²⁷¹ Indeed, the most recent federal expansion into the field of education reform — the No Child Left Behind Act — had already become deeply unpopular.²⁷²

Consequently, DOE turned to a program of competitive and conditional grants to influence policy.²⁷³ On its own, that approach is nothing new; using the federal spending power to influence state and local action is now commonplace.²⁷⁴ But this particular grant program

²⁶⁵ See U.S. DEP’T OF EDUC., *FUNDAMENTAL CHANGE*, *supra* note 67.

²⁶⁶ 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) [hereinafter *Race to the Top Fund*] (describing the program purposes).

²⁶⁷ See *generally id.* (calling for state and local implementation).

²⁶⁸ See *generally* Martin Kurzweil, *Disciplined Devolution and the New Education Federalism*, 103 CALIF. L. REV. 565, 570-71 (2015) (analyzing Race to the Top as part of an emerging new federalism model).

²⁶⁹ *Id.* at 571.

²⁷⁰ William G. Howell, *Results of President Obama’s Race to the Top*, 15 EDUC. NEXT 58, 60 (2015) (noting the degree of control exercised by the Department of Education).

²⁷¹ See CASS R. SUNSTEIN, *AFTER THE RIGHTS REVOLUTION: RECONCEIVING THE REGULATORY STATE* 22-24 (1990).

²⁷² See Kurzweil, *supra* note 268, at 601-02 (describing opposition).

²⁷³ See *Race to the Top Fund*, *supra* note 266, at 59,836; DEP’T OF EDUC., *SETTING THE PACE: EXPANDING OPPORTUNITY FOR AMERICA’S STUDENTS UNDER RACE TO THE TOP* 1-2 (2014).

²⁷⁴ See Pasachoff, *Agency Enforcement*, *supra* note 126, at 260-66 (describing the scope of federal grantmaking programs).

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.²⁷⁵ DOE did not explain exactly how it expected state and local governments to achieve all of those goals,²⁷⁶ 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.²⁷⁷ 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.²⁷⁸

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

[Grantee states] 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.²⁷⁹

²⁷⁵ Race to the Top Fund, *supra* note 266, 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.

²⁷⁶ Some of the criteria were so specific that they left little room for flexibility. See Kurzweil, *supra* note 268, at 603-04 (noting examples).

²⁷⁷ See U.S. DEP'T OF EDUC., FUNDAMENTAL CHANGE, *supra* note 67, at vii.

²⁷⁸ See Race to the Top Fund, *supra* 266, at 59,842-45 (defining criteria); Howell, *supra* note 270, at 60 (noting the degree of control exercised by the Department of Education).

²⁷⁹ Race to the Top Fund, *supra* note 266, 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.²⁸⁰

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.²⁸¹ 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.²⁸² But the commitment to data collection and results dissemination, the maintenance of some control over variables (through the menu of options that states could select), and 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 a reflection of their ideas.²⁸³ 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.²⁸⁴ 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.

²⁸⁰ *Id.*

²⁸¹ *See id.* at 59,842 (demanding plans for “statewide” improvement). The Department of Education also favored applications from states that were adopting the Common Core curricular standards, and thus encouraged some uniformity among as well as within states. *See id.* at 59,843.

²⁸² *See, e.g., id.* at 59,836 (“The purpose of the Race to the Top Fund . . . is to encourage and reward States that are creating the conditions for education innovation and reform”); U.S. DEP’T OF EDUC., *FUNDAMENTAL CHANGE*, *supra* note 67, at 49 (“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.”).

²⁸³ *See* Sabel & Simon, *supra* note 55, at 55-56 (asserting that “the Race to the Top education program can only be understood in experimentalist terms”) (footnote omitted).

²⁸⁴ *See supra* notes 55–62 and accompanying text.

Whether this program has improved educational outcomes is hard to say.²⁸⁵ 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.²⁸⁶ And while that question is crucially important to the teachers and students 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.

E. Jobs-Plus

A final example of federally-led experimentation — in the area of welfare policy — is similar to Race to the Top in that it involved the federal government relying on sub-federal entities to implement an experiment. But it was also an explicit and quite rigorous experiment, thus falling within the box labeled “F” in Figure 1 above.

The Jobs-Plus Community Revitalization Initiative for Public Housing Families (“Jobs-Plus”), which took place from 1998 to 2003, addressed unemployment and poverty in federally-subsidized housing projects.²⁸⁷ It was initiated, designed, and funded by the U.S. Department of Housing and Urban Development (“HUD”) and nonprofit partners, with the nonprofit group MDRC taking the lead on implementation.²⁸⁸ It used money from a public-private partnership,

²⁸⁵ Many education policy analysts have criticized the program. See, e.g., Natalie Gomez-Velez, *Urban Public Education Reform: Governance, Accountability, Outsourcing*, 45 URB. LAW. 51, 67 (2013) (summarizing criticisms and explaining that “the focus of the Race to the Top initiative on charter schools, the use of consultants for innovation, and the use of standardized testing to reform teacher evaluation has been viewed by some critics as directly related to reports about the increased interest in private investment in public education as way of accessing a huge pool of public money without sufficient evidence of effectiveness and without adequate regulatory oversight”).

²⁸⁶ See LISA DRAGOSET ET AL., RACE TO THE TOP: IMPLEMENTATION AND RELATIONSHIP TO STUDENT OUTCOMES XIV-XV (2016), <https://files.eric.ed.gov/fulltext/ED569959.pdf>; Howell, *supra* note 270, at 58.

²⁸⁷ RICCIO, SUSTAINED EARNINGS GAINS, *supra* note 47, at 1; see also DAVID M. GREENBERG ET AL., THE SECOND GENERATION OF JOBS-PLUS PROGRAMS iii (2015), https://www.mdrc.org/sites/default/files/CEO-SIF_Jobs-Plus_2015_FR.pdf (labeling the initiative as the “Jobs-Plus Public Housing Revitalization Initiative”).

²⁸⁸ See HOWARD S. BLOOM & JAMES A. RICCIO, USING PLACE-BASED RANDOM ASSIGNMENT AND COMPARATIVE INTERRUPTED TIME-SERIES ANALYSIS TO EVALUATE THE JOBS-PLUS EMPLOYMENT PROGRAM FOR PUBLIC HOUSING RESIDENTS 1, 11

the Social Innovation Fund, which is an evidence-based policy experimentation branch of the federal Corporation for National and Community Service.²⁸⁹ HUD and its partners relied on local entities to implement the program — specifically, at each project site (city) selected, the “local housing authority, resident representatives, the welfare department, and the workforce development system” within the community.²⁹⁰ But MDRC also provided extensive technical support from the top down.²⁹¹ The program’s goal was to improve employment for residents of public housing developments, with a specific hypothesis to be addressed: “Can a multicomponent employment initiative that is located in public housing developments help residents work, earn more money, and improve their quality of life?”²⁹² The federal and nonprofit project leaders hypothesized “that, by substantially increasing residents’ rates and stability of employment, other improvements in residents’ quality of life would follow”²⁹³ They further broke down this hypothesis into three sets of outcomes to be measured, including impacts on the extent to which residents were prepared for work or for advancement within their workplace, impacts on actual work outcomes, such as increased “employment and earnings,” less reliance on welfare, and “[i]mprovements in [p]ersonal [w]ell-[b]eing, [c]ommunity [c]onditions, and [q]uality-of-[l]ife [o]utcomes,” such as safety and improved conditions within the housing development.²⁹⁴

The “multicomponent initiative” — or, in experimental terms, the treatment — designed to test this hypothesis included “(1) employment services offered at on-site job centers, (2) changes in rent rules that provided financial incentives to work, and (3) community support for work through neighbor-to-neighbor conversations.”²⁹⁵ Employment services included offerings such as helping residents

https://www.mdrc.org/sites/default/files/full_600.pdf [hereinafter PLACE-BASED RANDOM ASSIGNMENT] (noting that these organizations designed the program and selected the cities to participate but that MRDC staff and consultants did the work such as visiting the cities and providing Jobs Plus workshops and training programs).

²⁸⁹ GREENBERG ET AL., *supra* note 287, at iii; *Social Innovation Fund*, CORP. FOR NAT’L & COMMUNITY SERV., <https://www.nationalservice.gov/programs/social-innovation-fund> (last visited Sept. 8, 2018).

²⁹⁰ *Social Innovation Fund*, *supra* note 289.

²⁹¹ See HOWARD S. BLOOM ET AL., PROMOTING WORK IN PUBLIC HOUSING: THE EFFECTIVENESS OF JOBS-PLUS, at ES-2 (2005) [hereinafter PROMOTING WORK].

²⁹² *Id.* at iii.

²⁹³ *Id.* at 5-6.

²⁹⁴ *Id.* at 15.

²⁹⁵ GREENBERG ET AL., *supra* note 287, at iii.

search for jobs, providing vocational training, and supplying the types of support that allow people to work, such as transportation and child care.²⁹⁶ Where the program was implemented, HUD changed rent policies such that when residents' earnings increased, they did not have to pay a higher proportion of their rent.²⁹⁷ And with respect to neighbor-to-neighbor communications, the program endeavored to foster discussions about "job opportunities or employment services."²⁹⁸

Through a competition, the federal government and its nonprofit partners selected cities with large public housing developments as sites for the experiment.²⁹⁹ They also established criteria to be demonstrated in an application, including prioritizing a "diverse set of sites where joblessness in public housing was a serious problem" and where there appeared to be a good opportunity to implement a program to address the problem, such as demonstrated "quality of local housing authority management."³⁰⁰ Further, the applicants had to commit to cooperate "with substantial data collection efforts."³⁰¹

The policy experiment was designed to test the hypotheses noted above and to attempt to control confounding variables. HUD and its partners "randomly assign[ed] entire housing developments (rather than individual residents) to either a program group or a comparison group within each site,"³⁰² explaining that the method of randomly allocating groups "was used in order to avoid . . . selecting for Jobs-Plus the best managed of the available developments or those that enjoyed the most favorable conditions for achieving employment outcomes."³⁰³ Specifically, the local housing authorities selected to participate had to "have at least two — and preferably three — developments" within their city "that would qualify for Jobs-Plus."³⁰⁴ The program implementers then used computerized random generation to pick developments that would receive the Jobs-Plus

²⁹⁶ BLOOM ET AL., PROMOTING WORK, *supra* note 291, at ES-2.

²⁹⁷ *Id.*

²⁹⁸ *Id.*

²⁹⁹ BLOOM & RICCIO, PLACE-BASED RANDOM ASSIGNMENT, *supra* note 288, at 25-26 (noting that 442 housing developments met the criteria, invitations to submit statements of interest were sent to fifty cities, and forty-one cities responded).

³⁰⁰ BLOOM ET AL., PROMOTING WORK, *supra* note 291, at 9.

³⁰¹ *Id.* at 10.

³⁰² *Id.* at 12.

³⁰³ *Id.* at 12.

³⁰⁴ *Id.*

program within the city and those that would serve as comparison groups within the same city.³⁰⁵

Beyond a concerted effort to establish control groups to account for confounding variables, the experiment involved extensive data collection and analysis, with project leaders examining pre- and post-project earnings trends, among numerous other data points.³⁰⁶ They collected data from public housing authorities on “residents’ background characteristics and their tenure in public housing” and also accessed state unemployment records and welfare payment records.³⁰⁷ Further, researchers conducted baseline and follow-up surveys of residents within the developments.³⁰⁸ The goal of the data collection was to measure the success of the program along three metrics: average increased earnings of residents and percentage of residents employed; the extent to which the effects varied among the sites and different residents and resident cohorts; and the extent to which Jobs-Plus lowered welfare reliance within the developments.³⁰⁹

Another classically experimental aspect of Jobs-Plus was the extent to which the program drew on lessons learned in similar experiments. The program was designed to build from a set of “comprehensive community initiatives” that involved entire communities and a broad set of services, rather than just individuals and individual services, in efforts to lift individuals out of poverty.³¹⁰ Jobs-Plus incorporated similar approaches but deliberately added a “third component” of a community program that was specifically focused on work support.³¹¹ And the program designers looked to similar, recent tests in Minnesota, Canada, and Wisconsin when forming the Jobs-Plus experiment.³¹² Further, researchers later replicated Jobs-Plus in other communities, examined the results from these programs, and compared them with the original Jobs-Plus results. The Social Innovation Fund of the Corporation for National and Community Service funded a “Jobs-Plus scale-up and replication” in the Bronx and San Antonio, and HUD is also implementing the program more broadly under the “Jobs Plus Pilot Program,” which commenced in

³⁰⁵ *Id.* at ES-3.

³⁰⁶ *Id.* at 67-68.

³⁰⁷ *Id.* at 14.

³⁰⁸ *Id.*

³⁰⁹ *Id.* at 67.

³¹⁰ BLOOM & RICCIO, PLACE-BASED RANDOM ASSIGNMENT, *supra* note 288, at 22.

³¹¹ *Id.* at 23.

³¹² *Id.* at 21-22.

2015.³¹³ The results of these replications are being documented and studied.³¹⁴

Retrospective review suggests substantial program success. Although there appears to have been surprisingly little external review of the program (and seemingly no meaningful attention in the legal literature), the limited external assessments are positive. An Urban Institute report concludes that Jobs-Plus was successful in boosting residents' earnings,³¹⁵ although it notes that the increased earnings did not spill over in terms of "improvements in residents' quality of life or community well-being" in the three communities tested under this sub-hypothesis.³¹⁶ More broadly, however, the report notes that Jobs-Plus, along with two other HUD programs, "represent serious investments in rigorous research" and "paid off" in terms of producing "significant new insights on strategies for tackling concentrated poverty and isolation."³¹⁷ The authors also observe that the lessons "should enable policymakers and practitioners to move forward more intelligently" in terms of steering low-income families toward neighborhoods that might present more opportunities for moving out of poverty and designing and providing employment programs and rent rules in low-income communities.³¹⁸

Internal program review revealed that at the four sites where the most complete implementation of Jobs-Plus took place, the program "markedly increased the *earnings* of public housing residents" (emphasis in original) across residential cohorts and subgroups, and effects on employment rates were positive but not as clear.³¹⁹ At three of the sites, the "program impacts were sustained for at least four years and showed no signs of diminishing."³²⁰ Later analysis noted that in the four sites with full Jobs-Plus implementation, "Jobs-Plus boosted residents' annual earnings by 16 percent, or \$1,300 per year, an effect

³¹³ GREENBERG ET AL., *supra* note 287, at ES-1.

³¹⁴ See *id.* at 111-13 (noting that the report investigates the early results of the scale-up and describing further nationwide efforts to implement and study similar Jobs-Plus programs).

³¹⁵ MARGERY AUSTIN TURNER & LYNETTE A. RAWLINGS, URBAN INST., OVERCOMING CONCENTRATED POVERTY AND ISOLATION: LESSONS FROM THREE HUD DEMONSTRATION INITIATIVES 2, 18 (2005), <https://www.urban.org/sites/default/files/publication/51636/311205-Overcoming-Concentrated-Poverty-and-Isolation-Executive-Summary-.PDF>.

³¹⁶ *Id.* at 19.

³¹⁷ *Id.* at 2.

³¹⁸ *Id.* at 3.

³¹⁹ BLOOM ET AL., PROMOTING WORK, *supra* note 291, at 67.

³²⁰ *Id.* at 68.

that endured seven years without abating.”³²¹ Program residents’ welfare reliance also “dropped precipitously,” but because the researchers had used careful controls and were able to tease out potential causal factors unrelated to the Jobs-Plus program, they concluded that the welfare effects were likely the result of changes in the national economy, among other national variables.³²²

The Jobs-Plus case study is another example of a rigorous experiment — similar to the coordinated and carefully-designed soil conservation policy experiments in the agricultural context — in which the federal government was centrally involved. HUD was instrumental in designing and funding the policy initiative and implementing portions of it, such as helping select the communities to which the initiative would apply and changing rent policies within these communities.³²³ And HUD’s primary national non-profit partner in the experiment, MDRC, was centrally involved in the experiment, sending “site representatives’ and other experts to provide ongoing operations-related technical assistance” in each community.³²⁴

* * *

Combined, these four case studies show varying degrees of federal involvement and, to some extent, rigor. As we note above, the federal government did not even describe Race to the Top as an experiment when it initiated the program. But all of the studies demonstrate that the federal government often is a central actor within policy “experimentation” of various types — not just in the role of facilitator or coordinator but often as a designer and implementer of experiments.

IV. EXPLAINING FEDERAL EXPERIMENTATION

In American politics and law, the balance of federal and state authority is a subject of never-ending debate.³²⁵ 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 reserve within the boundaries of federal

³²¹ GREENBERG ET AL., *supra* note 287, at iii.

³²² BLOOM ET AL., PROMOTING WORK, *supra* note 291, at iii, 106.

³²³ BLOOM & RICCIO, PLACE-BASED RANDOM ASSIGNMENT, *supra* note 288, at 10.

³²⁴ *Id.*

³²⁵ See *New York v. United States*, 505 U.S. 144, 149 (1992) (describing federalism as “our oldest question of constitutional law”).

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.³²⁶ Administrative agencies repeatedly confront the same questions, and in a variety of different contexts. So do the courts. Beneath preemption,³²⁷ Commerce Clause,³²⁸ and Tenth Amendment cases,³²⁹ among others, and beneath statutory interpretation cases involving the shadows of constitutional law,³³⁰ 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.

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.

³²⁶ 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 POL., POL'Y & L. 461, 463-66 (2011).

³²⁷ See Roderick M. Hills, Jr., *Against Preemption: How Federalism Can Improve the National Legislative Process*, 82 N.Y.U. L. REV. 1, 3 (2007).

³²⁸ E.g., *Nat'l Fed'n of Indep. Bus. v. Sebelius*, 567 U.S. 519, 533-37 (2012).

³²⁹ E.g., *Printz v. United States*, 521 U.S. 898, 918-22 (1997).

³³⁰ E.g., *Solid Waste Agency 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, 457-64 (1991) (using avoidance principles to narrowly interpret a statute that would have invalidated state laws mandating age-based retirement for judges).

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, Jobs-Plus, and fire management programs all 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 “Soil Erosion, A *National Menace*.”³³¹

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.³³² Even with federal incentives, such as grants or prizes, the federal government might not be able to inspire systematically differentiated efforts by sub-federal governments. In many circumstances, therefore, the federal government itself will be the best experimenter. In other words, 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.³³³ 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

³³¹ See U.S. DEP'T OF AGRIC., 1980 APPRAISAL, *supra* note 145, at 9 (emphasis added).

³³² See *supra* notes 88–93 and accompanying text.

³³³ See Owen, *Regional Federal Administration*, *supra* note 43, at 109. Parallel relationships are also possible between state governments, which often have field or regional offices within their states, and local governments.

reduce the transaction and monitoring costs that accompany reliance on subfederal 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.³³⁴ States and local governments have little incentive to collectively agree upon a uniform system or to share results — particularly negative ones.³³⁵ 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 and Jobs-Plus, 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.³³⁶

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. In fact, as the soil and fire examples illustrate, federal agency staff are just as close to many problems, if not closer, than their state counterparts.³³⁷

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.³³⁸ Usually we value treating like cases

³³⁴ 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 TUL. L. REV. 325, 326 (2016).

³³⁵ See *supra* notes 104–07 and accompanying text.

³³⁶ See *supra* notes 265–82 and accompanying text (describing the impact of the Race to the Top and Jobs-Plus programs and lessons learned with respect to future policy efforts).

³³⁷ See *supra* notes 158–96 and accompanying text (describing the localized staffing of federal soil conservation initiatives).

³³⁸ See Owen, *Regional Federal Administration*, *supra* note 43, at 76–77 (noting recurring concerns about inconsistent federal policies).

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.³³⁹ Our summary response is simply that we are persuaded by the many arguments in favor of administrative legitimacy.³⁴⁰ 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.

A final concern associated with federal experimentation — and one that we take quite seriously — was raised when the policy experimentation term was first coined in 1932. Justice Brandeis emphasized that a single, courageous state could experiment “*without risk to the rest of the country.*”³⁴¹ Some localized or state experimentation can in fact have significantly negative spillover effects, particularly in areas where experimentation occurs on resources that cross state lines, for example.³⁴² Indeed, much of the justification for federal intervention

³³⁹ Compare, e.g., *E.P.A. v. EME Homer City Generation, LP*, 134 S. Ct. 1584, 1610 (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, 1542-62 (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 generally the majority and dissenting opinions in *Free Enter. Fund v. Pub. Co. Accounting Oversight Bd.*, 561 U.S. 477 (2010).

³⁴⁰ See, e.g., Emily Hammond & David L. Markell, *Administrative Proxies for Judicial Review: Building Legitimacy from the Inside-Out*, 37 HARV. ENVTL. L. REV. 313, 359-64 (2013) (describing how court’s formal requirements for agency legitimacy can be translated into broader metrics to independently assess an agency’s legitimacy, such as avoiding arbitrary action and protecting democratic norms); Seidenfeld, *supra* note 339 (arguing that “[a]dministrative agencies . . . fall between the extremes of the politically over-responsive Congress and the over-insulated courts” and arguing that agencies are legitimate in their ability to best engage in deliberative decision-making bounded by judicial review and review by “politically accountable” Congress, among other factors such as agency policymaking processes).

³⁴¹ *New State Ice Co. v. Liebmann*, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) (emphasis added).

³⁴² See, e.g., Livermore, *supra* note 8, at 676-89 (noting potentially negative

arises where there are concerns about races to the bottom.³⁴³ But there is a greater risk that if the federal government experiments — and even when it differentiates policies among regions — the effects of experimental policy will be allocated along politicized lines, with favored states becoming the experimental sites for policies deemed likely to succeed, or disfavored states serving as guinea pigs for long-shot innovations. Though it does not arise from a policy experimentation context, Florida's recent special treatment in offshore drilling policy readily illustrates the potential political problems.³⁴⁴ Where that threat arises, and where traditional politics are not sufficient to check it, requiring some degree of state and local involvement in experimental policymaking can protect against federal politicization of the spoils of experimentation.

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 experiment stations.³⁴⁵ 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; disseminating the proven approaches, such as tree planting; and

consequences of state experimentation with water quality).

³⁴³ See Richard B. Stewart, *Pyramids of Sacrifice? Problems of Federalism in Mandating State Implementation of National Environmental Policy*, 86 *YALE L.J.* 1196, 1212 (1977).

³⁴⁴ See Matt Pearce & Gray Rohrer, *Democrats on both Coasts Cry Foul after Trump Exempts Florida from Oil-Drilling Plan*, *CAP. GAZETTE* (Jan. 9, 2018, 9:10 PM), <http://www.capitalgazette.com/la-na-trump-oil-drilling-florida-20180109-story.html> (detailing states' objections after the Trump administration "exempted Florida from expanded offshore drilling . . . without offering similar exemptions to other coastal states").

³⁴⁵ The government did, however, typically rely on donation of state lands, and later partnered with the states to conduct the experiments.

providing seeds and labor, on thousands of individual farms.³⁴⁶ 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, continuing the effort to demonstrate conservation practices within communities and encourage farmers to adopt them.³⁴⁷ 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. And in the Jobs-Plus example, the federal government used federally-funded housing projects as the experimental medium, although it relied heavily on the local actors who run these projects to implement the experiment and help collect data.

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. As scholars of the policy experimentalism field note, the medium of experimentation, or the “units of randomization,” vary substantially; while some experiments have occurred in federally-funded housing developments, others have played out in hospitals,³⁴⁸ doctors' family practices,³⁴⁹ and local public schools.³⁵⁰ Often, this will simply be an accident of federalism and path dependence —

³⁴⁶ See *supra* notes 159–62 and accompanying text.

³⁴⁷ See *Local Service Centers Directory*, USDA NAT. RESOURCES CONSERVATION SERV., <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/contact/local/> (last visited Sept. 9, 2018) (providing the location and contact information of USDA Service Centers).

³⁴⁸ See, e.g., Laura C. Leviton & Jeffrey D. Horbar, *Cluster Randomized Trials for the Evaluation of Strategies Designed to Promote Evidence Based Practice in Perinatal and Neonatal Medicine*, 599 ANNALS AM. ACAD. POL. & SOC. SCI. 94, 96-97 (2005) (describing a policy experiment to improve caregivers' implementation of more effective treatment practices — altering “medical practice to be consistent with the evidence” — in which the units of study were hospitals).

³⁴⁹ Jeremy Grimshaw et al., *Cluster Randomized Trials of Professional and Organizational Behavior Change Interventions in Health Care Settings*, 599 ANNALS AM. ACAD. POL. & SOC. SCI. 71, 78 (2005) (describing a policy trial examining the extent to which medical research results are incorporated in to practice and the use of the “family practice” as the level at which to randomize the trial).

³⁵⁰ See, e.g., Brian R. Flay et al., *Effects of 2 Prevention Programs on High-Risk Behaviors Among African-American Youth*, 158 ARCH. PEDIATR. ADOLESC. MED. 377 (2004) (describing a randomized policy trial that involved public schools and the communities in which they were located).

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.³⁵¹ Yet by granting funds based on a checklist of factors that states and local governments could experiment with, and enlisting experts to periodically check in with states and school districts and making them talk to each other on conference calls, the federal government helped to differentiate the experimental approach, collect important data, and spread lessons learned.³⁵²

* * *

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.³⁵³ 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.

³⁵¹ *The Federal Role in Education*, U.S. DEPT. EDUC., <https://www2.ed.gov/about/overview/fed/role.html> (last modified May 25, 2017) (explaining that “the Federal contribution to elementary and secondary education is about [eight] percent” of total budgets); see also *Brown v. Bd. of Educ.*, 347 U.S. 483, 493 (1954) (describing education as “perhaps the most important function of state and local governments”).

³⁵² See *supra* notes 279, 286 and accompanying text.

³⁵³ See Owen, *Cooperative Subfederalism*, *supra* note 66.

It is also worth noting that there are political, economic, and other reasons that explain why relatively rigorous, federally-led experimentation seems to dominate some policy areas — such as agriculture — but not others, such as water pollution control.³⁵⁴ And perhaps some areas are more conducive to federal experimentation than others, as Michael Livermore has noted.³⁵⁵ There is a large literature on the drivers of and hurdles to decentralized experimentation,³⁵⁶ but the motivations for federal experimentation remain somewhat of a mystery. As we note in the agricultural case study, Congress requires evidence of the results of agricultural conservation programs and incorporation of this evidence into improved policy.³⁵⁷ This might have resulted from a relatively straightforward desire to ensure that federal funds are spent in the most efficient and effective way possible, or more politically-motivated purposes, such as demonstrating positive results for powerful agricultural interest groups within House and Senate members' home jurisdictions. Congress similarly at least paid lip-service to experimentation in certain recent educational initiatives. But why would Congress not require the same type of experimentation in the many other areas in which it doles out funds? One might expect that the federal government would feel the most freedom to risk experimentation in areas where there is less risk of backlash from powerful stakeholders. The case studies above suggest that federal experimentation occurs both in areas of relatively strong stakeholder influence, such as agriculture, and somewhat more neutral areas, such as forest fire policy.³⁵⁸ But the agricultural experimentation appears to occur primarily within the area of conservation funding, and not, for example, the more politically-charged, stakeholder-dominated aspects

³⁵⁴ See, e.g., Livermore, *supra* note 8, at 681-83 (describing limited experimentation, as well as certain limits on experimental opportunities, in this area and noting that “[m]ore constrained federal jurisdiction would reduce the ability of EPA to develop complementary policies to encourage innovation on the part of the states”).

³⁵⁵ See *id.* at 676-88 (describing why reduced federal jurisdiction would not likely lead to beneficial decentralized experimentation in the water pollution context).

³⁵⁶ See *supra* notes 5-6.

³⁵⁷ See *supra* note 217.

³⁵⁸ There are, of course, numerous stakeholders concerned about forest fires, such as landowners who live on the urban-suburban edge, ranchers, state emergency response officials, and others. See Bradshaw, *supra* note 259, 451-66 (2010) (describing the stakeholders). But these stakeholders do not seem as organized or well-funded as the powerful agricultural lobby.

of agricultural policy such as crop insurance or fertilizer subsidies.³⁵⁹ Thus, freedom from traditional “capture” might be part of the story.

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.³⁶⁰ 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 is ample room, in other words, for additional use of, and inquiry into, our federal laboratories of democracy.

³⁵⁹ See, e.g., Linda Breggin & D. Bruce Myers, Jr., *Subsidies with Responsibilities: Placing Stewardship and Disclosure Conditions on Government Payments to Large-Scale Commodity Crop Operations*, 37 HARV. ENVTL. L. REV. 487, 515-21 (2013) (describing the subsidy programs).

³⁶⁰ See generally Livermore, *supra* note 8 (discussing downsides of experimental policy). In a recent example of a dangerous and deeply problematic application of federal laboratories, the Trump Administration ran what it called a “pilot program” for separating immigrant children from their families at the federal government’s immigration center in El Paso, Texas, in 2017, before the official “zero tolerance” program was implemented. See Lisa Riordan Seville & Hannah Rapple, *Trump Admin Ran ‘Pilot Program’ for Separating Migrant Families in 2017*, NBC NEWS (June 29, 2018, 1:30 AM), <https://www.nbcnews.com/storyline/immigration-border-crisis/trump-admin-ran-pilot-program-separating-migrant-families-2017-n887616>.