

Just How Much of History Is Countable?

Daniel Carpenter*

Department of Government, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138, USA

*Corresponding author: Email: dcarpenter@gov.harvard.edu

Abstract

How can historical perspective be brought to the quantitative social sciences? The question has proven immensely popular, but answers that deal squarely with historical context and narrative remain elusive. An important recent book by Gregory Wawro and Ira Katznelson—*Time Counts*—provides an important new direction and a kit of useable tools. Using Wawro and Katznelson's approach and methods puts social scientists in a better position to appreciate the historicity of their data and to avoid common errors in statistical execution and inference. *Time Counts* also raises questions every bit as vital as those it answers, especially when it comes to the boundaries between narrative and quantitative work. An important concern is that inference from a particular historical setting (or what I call a "regime") cannot be reduced to a special case of inference from large-sample statistics. Historical judgment is at least partially incommensurable with the idea of probability, cases are often important precisely because they are not countable, and scientific rigor may demand avoiding quantification for part of the social scientist's approach.

Keywords: countability; causal analysis; historical social science

The explosive advance of quantitative political analysis marks one of the major transformations of knowledge in the twentieth century. To be sure, any number of students and scholars of human society had engaged in social and political quantification before 1900, and their activity in the revolutionary era of the late eighteenth century was especially generative: Jefferson's (1782) skillful demonstration of malapportionment in the Virginia House of Delegates, Zimmerman's *Survey on the Present State of Europe* (1788), Sieyès's (1789) analysis of representation in the French estates generals, especially of 1614, and Sir John Sinclair's *Statistical Accounts of Scotland* (1791).¹ Starting a rough century later, a congeries of developments changed the landscape permanently. They include the development of the correlation coefficient in the late nineteenth century, the evolution of quantitative

Daniel Carpenter is the Allie. S. Freed Professor of Government and Chair, Department of Government in the Faculty of Arts and Sciences at Harvard University. His most recent monograph—*Democracy by Petition: Popular Politics in Transformation, 1790–1870* (Harvard University Press, 2021)—establishes petitioning as a transformative force in American, Canadian, Caribbean, and Mexican political development, and it won the J. David Greenstone Prize of the American Political Science Association, the Seymour Martin Lipset Prize of the American Political Science Association, and the James Rawlins Prize of the New England Historical Association.

¹ Jefferson, *Notes on the State of Virginia*, Query XIII (1782); Zimmerman, *A Political Survey of the Present State of Europe, in Sixteen Tables; Illustrated With Observations on the Wealth and Commerce, the Government, Finances, Military State, and Religion of the Several Countries*. By E. A. W. Zimmermann, Professor Of Natural Philosophy At Brunswick, And Member Of Several Scientifical Societies (Dublin: Luke White, 1788); Emmanuel Sieyès, *Qu'est-ce que le Tiers État?* Ch. 3, S. 2 (Paris, 1789) and Sieyès, notes sur les états généraux (n.d., probably 1787 and 1788), AP 284 3, dossier 2, Fond Sieyès, Archives Nationales de France; Sinclair, *The Statistical Account of Scotland, Drawn Up from the Communications of the Ministers of the Different Parishes* (Edinburgh: William Creech, 1791).

hypothesis testing defined by reference of test statistics to probability distributions, the emergence of more sophisticated actuarial and life tables using notions of conditional probability, the explosion of census-based measurement of populations, and the transformative marriage of statistics and computing.

Today not a day or hour goes by in the social sciences when some scholar does not convert a set of observations from history into “data” and, using software that could be run from a phone, conduct some kind of statistical analysis upon it. The development of linear statistical models in historical research gave rise to historiographical cliometrics, economic history, the emergence of multidisciplinary communities of quantitative historical inquiry—the Social Science History Association is one—and the rise of quantitative historical political science and sociology, including the cross-disciplinary subfields of American political development (APD) and historical political economy (HPE).² More recently, much more expansive claims in the use of history as data have arisen and have begun to reshape political science. The first is the set of “legacy studies” in which a set of events in “the past” can be analyzed as predictive or causative of patterns later. In a second vein, studies in historical political economy (HPE) examine historical changes in a quantitative and rational choice context (usually dispensing with interpretive analysis or nonquantitative archival research) and cross several subfields of political science. Legacy studies often qualify as HPE studies, but the latter also include a range of situated cross-sectional designs and also formal-theoretic or rational choice exercises in political historical analysis.

In the view of critics from applied statistics and historiography, many such “regressions on history” exercises are mis-specified. Votes, wars, protests, and petitions do not statistical samples make. And the matter gets worse when we try to assume a linearity or stationarity of time such that actions and events in one period are assumed to be sufficiently similar to those of another as to readily constitute a sample for unified analysis. In *Time Counts: Quantitative Analysis for Historical Social Science*, the Columbia University scholars Gregory Wawro and Ira Katznelson marry historical social theory, institutional analysis, legislative studies, and applied statistics to generate both critique and solution. Wawro is a leading scholar of the U.S. Congress and its development who has also authored a range of impressive innovations in applied political methodology. Katznelson is a world-renowned historian and political scientist of American and European cities, race, class and politics, having also written in political philosophy. Each of the authors is something of a switch-hitter already. Putting their talents together makes for a tour-de-force combination. Think less Carlos Beltran or Chipper Jones, more Shohei Ohtani.

The terrain navigated by Wawro and Katznelson is, as the authors recognize, far more vast and more treacherous than any one book can handle. The authors focus their theoretical critiques upon political economy and political science, their empirical critiques on American political development, and their applied exercises upon U.S. congressional history. Their achievements are real and lasting. After this

² In their summary of the American political development field as of the dawn of the twenty-first century, Karen Orren and Stephen Skowronek explicitly acknowledged, included, and welcomed the contribution of quantitative modes of inquiry; Orren and Skowronek, *In Search of American Political Development* (Cambridge: Cambridge University Press, 2002). On empirical trends in HPE, consult Alexandra Cirone, “Data in Historical Political Economy,” in *Oxford Handbook of Historical Political Economy*, ed. Jeffrey Jenkins (Oxford: Oxford University Press, 2022).

book, it will be difficult if not indefensible for scholars to approach multiple periods of political history with a single statistical or regression model that casually or explicitly subsumes them. (Though I have little doubt that, until they have read *Time Counts*, many scholars will continue to do so.) Wawro and Katznelson employ their combination of vast expertise in legislative history and quantitative analysis to clarify if not resolve important empirical debates on party power in Congress, party identification and polarization in the American electorate, and Senate elections, among other issues. They develop, explain, and experiment with a suite of tools that specifically embed historical variation and context into statistical models. With a “thoughtful utilization of approaches that privilege parameter variation,” they bridge deep contextuality and causal inference.³

The landscape beyond *Time Counts* remains of interest, though, and Wawro and Katznelson do a sufficiently thorough job of problematizing statistical historiography that it is worth asking whether and how we ought to go further. By “further,” I mean more extensive in critique and more expansive in analysis. The authors leave open the possibility that what is problematic but solvable using parameter variation may in fact gesture to a harder problem, namely that social science “data” have incommensurability problems not merely across periods but also across space or, more generally, across *regimes* (to some, “structures,” to others, “cultures,” to still others, “equilibria”). The analytic strategies that could be deployed to meet these dilemmas also call for new tools, among which might be included the possibility that faced with certain mixtures of regimes, scholars should consider downplaying or even abandoning quantitative analysis and relying more heavily upon within-regime narrative.

Dilemmas of the Contemporary Political History Regression

The backdrop for *Time Counts* is the explosive growth of statistical correlation and multiple regression analysis in late twentieth century historical social science. The centrality of databases as a fundamental contribution of the historical social science dissertation became a central contribution of the *Annales* school of social research emanating from France. Economists like Robert William Fogel and his students subjected these data to “econometric” exercises such as estimating marginal associations and production functions, but economists were not the only group employing correlation-based techniques or linear models.⁴ A range of

³ Gregory Wawro and Ira Katznelson, *Time Counts: Quantitative Analysis for Historical Social Science* (Princeton, NJ: Princeton University Press, 2022), (hereafter TC), 2–3.

⁴ I recall a March 2013 conference at l’Université Marne-la-Vallée (*Pétitionner: L’Appel au Pouvoir*) at which every French doctorate in history or sociology had collected a *base de données* (database), even if their analysis turned more to the narrative and interpretive than the quantitative. Among the most influential works in history and historical sociology were Fernand Braudel, *La Méditerranée et le Monde Méditerranéen à l’Époque de Philippe II* (Paris: Librairie Armand Colin, 1949); Lee Benson, *The Concept of Jacksonian Democracy: New York as a Test Case* (Princeton, NJ: Princeton University Press, 1961); G. Kitson Clark, *The Making of Victorian England* (London, 1962); Robert William Fogel, *Railroads and American Economic Growth: Essays in Econometric History* (Baltimore, MD: Johns Hopkins University Press, 1964); Stephan Thernstrom, *The Other Bostonians: Poverty and Progress in the American Metropolis, 1880–1970* (Cambridge, MA: Harvard University Press, 1973); Stanley Engerman and Robert Fogel, *Time on the Cross: The Economics of American Negro Slavery*, Vols. 1 and 2. (New York: Little, Brown and Co., 1974); Richard McCormick, *From Realignment to Reform: Political Change in New York State, 1893–1910* (Ithaca, NY: Cornell University Press, 1981); Doug McAdam, *Political Process and the Development of Black Insurgency, 1930–1970* (Chicago: University of Chicago Press, 1982); Gary Cox, *The Efficient Secret: The Cabinet and the Development of Political Parties in Victorian England* (Cambridge: Cambridge University Press, 1987). Writers who began to question or problematize this movement include William O. Aydelotte, “Quantification in History,” *American Historical Review* 71, no. 3 (1966): 803–26; Allan G. Bogue, *Clio & the Bitch Goddess: Quantification in American Political History* (New York: SAGE Publications, 1983).

scholars in history and sociology began to question these moves, worrying that they would eclipse traditional or alternative models of historical inquiry. Yet the transformation continued apace. In American political history and American political development as it evolved before 2000, the quantitative turn was led by scholars such as Richard Bense, Charles Stewart III, John Aldrich, Lee Ann Banaszak, Keith Poole, Howard Rosenthal, and Elizabeth Sanders. Wawro and Katznelson are contributors to this development, as is this reviewer.⁵

Wawro and Katznelson survey this rich literature and train their fire on several recent developments. The first is the approach to historical and comparative analysis that premises the model of research upon a regression-based quasi-experiment. The paradigm for this is the King, Keohane, and Verba book *Designing Social Inquiry* (1994). Following others, Wawro and Katznelson rightly argue that the dialogue between statistical science and historical inquiry should not be reduced to the wholesale adoption by historians of positivist methodologies ranging from the Millian comparative method and causal inference. The second difference comes in the multimethod scholars Gary Goertz and James Mahoney in political science and sociology, who argue for the distinctiveness of historical methods that should, at least in part, be preserved as separable traditions of methodology even as they are developed. Between these paths, Wawro and Katznelson seek the “prospect of partial but deep mutual constitution where it makes the most sense” (12).⁶

Wawro and Katznelson “sally forth by showing how quantitative scholars highlight the limitations of much extant quantitative scholarship for serious historical work” (13). In so doing, they ably summarize a long thread of scholarship on what goes wrong when history—or more properly, the observable residua of historical processes—become data for a regression model. They rightly observe that much quantitative history is arguably less theoretically motivated or informative than narrative historiography, that much quantitative history sacrifices understanding of historical process to achieve fit with statistical theories, and

We wish to advance significant historical work in the social sciences that pays due attention not only to situation and context, but also to concerns regarding time and sequence that are central to the craft of historians. Qualitative social scientists have been taking such efforts forward in full awareness that history unfolds with an unstable combination of regularities, mechanisms, change and randomness that requires moving up and down a ladder of abstraction from the conceptual to proper-named people, places and events...⁷

⁵ Richard Bense, *Sectionalism and American Political Development* (Madison: University of Wisconsin Press, 1984); Charles Stewart, *Budget Reform Politics: The Design of the Appropriations Process in the House of Representatives, 1865–1921* (New York: Cambridge University Press, 1989); Lee Ann Banaszak, *Why Movements Succeed or Fail: Opportunity, Culture and the Struggle for Woman Suffrage* (Princeton, NJ: Princeton University Press, 1996); Keith Poole and Howard Rosenthal, *Congress: A Political-Economic History of Roll-Call Voting* (Oxford: Oxford University Press, 1997); Elisabeth Sanders, *Roots of Reform: Farmers, Workers and the American State* (Chicago: University of Chicago Press, 1999). Gregory J. Wawro and Eric Schickler, *Filibuster: Obstruction and Lawmaking in the US Senate* (Princeton, NJ: Princeton University Press, 2007); David Bateman, Ira Katznelson, and John Lapinski, *Southern Nation: Congress and White Supremacy After Reconstruction* (Princeton, NJ: Princeton University Press, 2019); Daniel Carpenter, *Democracy by Petition: Popular Politics in Transformation, 1790–1870* (Cambridge, MA: Harvard University Press, 2021).

⁶ Wawro and Katznelson, TC, 12. Gary King, Robert O. Keohane, and Sidney Verba, *Designing Social Inquiry: Scientific Inference in Qualitative Research* (Princeton, NJ: Princeton University Press, 1994); Gary Goertz and James Mahoney, *A Tale of Two Cultures: Qualitative and Quantitative Research in the Social Sciences* (Princeton, NJ: Princeton University Press, 2012).

⁷ Wawro and Katznelson, TC, 24.

that statisticians' understanding of time is often deeply ahistorical (sometimes linear, almost always monotonic). These problems they take as their starting point, and they wish to work at the interstices of fields without rejecting either quantification or narrative. A central lesson of *Time Counts* is that statistical samples taken from historical processes do not satisfy the *unit homogeneity assumption*. This assumption has been defined as the postulate "that cases have enough meaningful similarities to be comparable," but beyond comparability there is the added benefit of aggregation. In theory, the observations in a statistical sample are substitutable for one another in the aggregation that produces basic results like the central limit theorem. Wawro and Katznelson rightly counter that in making this assumption (often unconsciously or in ways that reflect little deliberation), statistical researchers commit themselves to a basic fallacy: "Observations that are governed by different data generating processes are shoe-horned into homogenous models that frequently are excessively straightforward." The key data generating processes here are historical, and Wawro and Katznelson focus on the pitfalls of including data from different periods in the same model or, when the data are included, failing to adjust the statistical model to account for fundamental differences among observations seen across historical periods. One of their primary contributions is a set of models that addresses this fact of unit heterogeneity, whether semiparametric methods such as local polynomial regression, change point models, or Markov switching models with time-varying transition probabilities.⁸

There are many payoffs from an exercise such as this, and two of the highlights are a comparison of two economic approaches—the more strictly analytic approach of Acemoglu-Robinson versus the more contextualized approach pioneered by Avner Greif—and a thoughtful replication and critique of Thomas Brunell and Bernard Grofman’s study of state delegation splits in the Senate.⁹ The comparison of Acemoglu-Robinson and Greif is one of the best parts of the book. Arguing that the theoretically motivated and meaningful social science study of history “requires moving up and down a ladder of abstraction from the conceptual to proper-named people, places, and events . . . ,” they side with Greif’s vision of analytic inquiry as opposed to that of Acemoglu and Robinson. In *Economic Origins of Dictatorship and Democracy* (2005), Acemoglu and Robinson fashion “a rigorous theory of institutional change and development that is not tested with thick historical evidence. Actors remain stylized. Proper names do not matter.” Greif, by contrast, has a narrative in which he argues that long-run trade and institutions co-evolve in ways that are contingent and highly contextually dependent, and his approach allows for multiple equilibria in ways that Acemoglu and Robinson’s do not. While one could point to several other articles that Acemoglu and Robinson have written with coauthors, articles in which they examine particular events and transformations more specifically, Wawro and

⁸ Wawro and Katznelson, *TC*, 14 (shoehorning), 43–45 (historical heterogeneity), 45–54 (local polynomial regression), 55–58 (change point models), 116–129 (Markov switching models). Michael C. Desch, “Culture Clash: Assessing the Importance of Ideas in Security Studies,” *International Security* 23, no. 1 (Summer 1998): 141–70 (definition of unit homogeneity assumption, 152); Elisabeth S. Clemens, “Toward a historicized sociology: Theorizing events, processes, and emergence,” *Annual Review of Sociology* 33 (2007): 527–49.

⁹ Daron Acemoglu and James A. Robinson, *Economic Origins of Dictatorship and Democracy* (Cambridge: Cambridge University Press, 2005); Daron Acemoglu, Simon Johnson, and James A. Robinson, "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review* 91, no. 5 (2001): 1369–1401; Avner Greif, *Institutions and the Path to the Modern Economy: Lessons from Medieval Trade* (Cambridge: Cambridge University Press, 2006).

Downloaded from <https://academic.oup.com/psq/advance-article/doi/10.1093/psqua/rqad117/7369559> by Harvard College Library, Cabot Science Library user on 07 November 2023

Turning to a 1998 study of Senate delegations by Thomas Brunell and Bernard Grofman, Wawro and Katznelson show how a study that compresses nearly two centuries of history into a single data set does not do so without serious costs.¹¹ Senate delegation splits are fascinating within a state because they point to a lack of partisan dominance in a constituency, a (possibly related) weakness of state parties, and the durability of a particular kind of representation or principal-agent problem in that the public anoints quite different agents who act in the name of one people. Brunell and Grofman showed that delegation splits were far more of a historical phenomenon than scholars had originally believed, and that the phenomenon observed a cyclicity associated with then-established patterns of realignment. Wawro and Katznelson step away from realignment theory to “employ a model that is more flexible about the locations of the change points that

¹¹ Thomas L. Brunell and Bernard Grofman, "Explaining Divided U.S. Senate Delegations, 1788–1996: A Realignment Approach," *American Political Science Review* 92 (1998): 391–400.

define shifts in partisan dominance.”¹² They shed doubt upon the Seventeenth Amendment as driving force in split delegations, a significant finding given the literature pointing to the Seventeenth Amendment as a decisive break in patterns of Senate representation. With the methods they develop, Wawro and Katznelson conclude that the properly identified historical “break” in split Senate delegations occurs in 1946 (*Time Counts*, Figure 4.4), consistent with the growing nationalization of political parties and changes in the structure of the post-World-War II legislative branch. Wawro and Katznelson take issue with Brunell and Grofman’s theoretical conclusions, as well. The idea of realignment-associated patterns in split Senate delegations may have worked in the antebellum period but does not afterwards.¹³

Narrative, Regime Multiplicity, and Incommensurability

So far, so good. Wawro and Katznelson’s framework permits a marriage of historical inquiry and a portfolio of models that permit the context of one historical period to differ systematically in its attributed causal dynamics from another. With exercises like those performed upon *longue durée* historical data sets on state Senate delegations or party power in House voting, Wawro and Katznelson create a space where the contextually insistent qualitative researcher and the applied statistician can sit down at the same table, have a mutually constructive dialogue, and simultaneously possess and consume their *gâteau*. Wawro and Katznelson have created a space in which, between these paths, we can imagine the “prospect of partial but deep mutual constitution where it makes the most sense.”¹⁴

This seems the question left begging. *Where does the mutual constitution of contextually insistent historical inquiry and flexible statistical methodology make the most sense? And where does it not?* What are the telltale signs of a history that it does *not* make the most sense, or *any* sense, to perform this *rapprochement*, this “partial but deep mutual constitution”? Which histories or historical processes are amenable to deep mutual constitution and which are not? By extension, which are subject to quantification but should not be quantified? Will quantitative measurements alone be useful for indicating which histories are subject to flexible parametric approaches that respect context, in which case (with the assistance of artificial intelligence) we might soon arrive at a point where humanistic historiographical inquiry no longer matters? Or if historiographical evidence and lenses are needed, which ones?

One alternative to the *rapprochement* is the idea that the simplified linear statistical model can be employed without the contextualization that Wawro and Katznelson recommend. (That has been happening for decades and continues at this writing.) But the other is more radical still: there might be a scientific rationale for refusing to consider seriously any quantification of a long historical process on the grounds that even partial, no less deep, mutual constitution does not “make sense.” Put differently, such quantification might be “senseless” or “meaningless” in the most literal terms. For some (not all) settings, the “scientific” approach to these historical processes may not include quantification.

¹² Wawro and Katznelson, *TC*, 76.

¹³ Wawro and Katznelson, *TC*, 78–79.

¹⁴ Wawro and Katznelson, *TC*, 68–74 (House and Senate empirical examples), 12 (“most sense”).

Wawro and Katznelson's lessons from state Senate delegations is a sobering one. A postulated logic and relationship—the coincidence of partisan realignments and split state delegations in the U.S. Senate—prevails for one part of a long-time series, but not for others. Even if the idea of realignment theory can survive its withering critiques at the hands of David Mayhew, there are other problems. The institutional study of legislatures has long been dominated, for many good reasons, by rational choice models, especially noncooperative game-theoretic models such as bargaining, principal agent, and signaling models. The possibility of a causal process fitting the antebellum era but not later periods raises the possibility that the very multiple equilibria that Wawro and Katznelson find so appealing in Avner Greif's work might be at play here, too.

Once we take multiple equilibria seriously, then the implications become more radical. It is a well-known fact of much game-theoretic and other economic modeling that it is not possible to conduct comparative statics across different equilibria of the same game. The very comparative statics that are generated by mathematical models can in general only be deducted from one equilibrium at a time. Economic theorists have developed ways of thinking about how comparative statics can be considered across equilibria of a game or economy, but the solution often involves the assumption that there is an index across which the equilibria can be compared and, beyond that, ordered. In cases such as aggregate production, aggregate wealth, economic growth and others, this ordering is a reasonable, maybe even natural, assumption. Some political scientists have used welfare considerations like these to order equilibria from games involving legislative institutions. Yet it is worth asking, from a rational choice sense, whether we would ever expect players to converge to the equilibrium that satisfies or maximizes some ordering criterion, especially in politics.¹⁵

The point extends beyond economics and comparative statics. One can imagine a world not of multiple equilibria but of *multiple regimes*, in the sense of spaces (worlds) governed by certain logics, cultures, and understandings. If historical spaces, institutions, and processes are considered like equilibria (or like combinations of equilibria), then a basic problem of incommensurability arises. There are, put differently, entire histories and cultures that scholars often regard as not fundamentally comparable with one another in the sense of constituting a “case” of a more general phenomenon. Consider the advice of historian of science Peter Galison.

Imagine a book entitled *A Case Study in European History: France*. This made-up title strikes me as immensely funny, not because it purports to be a detailed study of an individual country (there are many important national histories), but because it encourages the reader to imagine a homogeneous class of European countries of which France is an instance. The absurdity rests upon the

¹⁵ Regarding multiple equilibria, I include two meanings of the term: first, the more proper sense of multiple equilibria given that same parameter configuration (in the classic sense of the folk theorem) and second, multiplicity given different parameter configurations (as in Daniel Carpenter and Michael Ting, “Regulatory Errors with Endogenous Agendas,” *American Journal of Political Science* 51, no. 4, 2007). Some classic articles outlining and addressing the problem of comparative statics and multiple equilibria are Timothy J. Kehoe, “Multiplicity of Equilibria and Comparative Statics,” *Quarterly Journal of Economics* (February 1985): 119–47; and Paul Milgrom and John Roberts, “Comparing Equilibria,” *American Economic Review* 84, no. 3 (1994): 441–59. Milgrom and Roberts explain the basic problem quite cogently, as involving the applicability of the implicit function theorem to derive predicted changes in the endogenous variable $x^*(z)$ to the exogenous variable z (they use t , but I use z so as not to invoke “time”). The Milgrom-Roberts approach requires not a numerical measure but only a dimension on which an ordering is possible (“Comparing Equilibria,” 443, and their Theorem 2 and Corollary 5). For an example of using optimality criteria to choose among different equilibria, consult Christophe Crombez, Tim Groseclose, and Keith Krehbiel, “Gatekeeping,” *The Journal of Politics* 68, no. 2 (2006): 322–34.

discrepancy between the central and distinctive position we accord France in history and the generic position we must assume France occupies if we wish to treat it as a “case.”¹⁶

Galison's quip puts some pressure upon social scientists both quantitative and qualitative. Much qualitative and mixed-method research, before and after the King, Keohane, and Verba synthesis, approaches a range of countries, historical periods, cultures, institutions, and other settings as "cases." The possibility of multiple equilibria here means that the act of comparison, statistical or qualitative, needs a priori justification, limitation, and even the possibility of abandonment. The pitfalls of regarding certain regimes as cases does not mean that the regimes cannot be studied, but that they cannot be studied as "cases" to be subsumed in a comparative framework among others, at least not without thinking far more carefully about what we are comparing—less plausibly, cases of European history, more plausibly, cases of working-class formation—across national or geographic settings.

Two examples from historical political science might shed light on the difficulties of cross-regime comparison. In an innovative article, Arthur Spirling studied the text of treaties between the United States Government and various Native American nations or tribes between 1789 and 1912. He shows that fundamental shifts in the measured favorability of these treaties to Native interests occur at certain points in the long nineteenth century, especially in the late 1820s, the end of Reconstruction, and the end of the congressional treaty-making era in 1883. While Spirling's lessons come from a comparison of these treaties over a long time, one might also conclude that for other purposes, comparing Native–U.S. treaties before the War of 1812 and the Jacksonian political settlement to those after the Civil War is not really possible quantitatively. Once this possibility is raised, one might ask whether the same kind of breaks that Spirling observes among periods might also apply to different regions. (Spirling analyzes the treaties over time, not geographically.) Richard White's monumental *The Middle Ground* examines a 150-year period in the Great Lakes region (the mid-seventeenth century to the War of 1812, roughly) during which a power vacuum (more properly, a hegemony vacuum) reigned among a range of Native American villages and European empires and their settlers. The equilibrium was one of mutual containment but also stable misunderstandings (in the sense of which Indigenous people and settlers misunderstood each other even as they thought they knew one another well), such as the meaning of Jesus (God to French Jesuits, one manitou among others to many Algonquian Natives) and the implications of a murder. The history of the Great Basin at the same time reveals a much more violent regime, as does British North America east of the Appalachians. The *pays d'en haut* of the Great Lakes region from 1650 to 1812, the Great Basin before 1850, and the area now known as New England before 1750 are three regimes, in some sense distinct equilibria, which it may not be possible to compare as “countable cases” of the same general phenomenon, perhaps of any general phenomenon, in the sense that they could be aggregated to compute average differences or treatment effects.¹⁷

¹⁶ Peter Galison, *Image and Logic: A Material Logic of Microphysics* (Chicago: University of Chicago Press, 2005), 59.

¹⁷ Arthur Spirling, "US Treaty Making with American Indians: Institutional Change and Relative Power, 1784–1911," *American Journal of Political Science* 56, no. 1 (2012): 84–97; Richard White, *The Middle Ground: Indians, Empires and Republics in the Great Lakes Region, 1650–1815* (Cambridge: Cambridge University Press, 1992); Ned Blackhawk, *Violence Over the Land: Indians and Empires in the Early American West* (Cambridge, MA.: Harvard University Press, 2006); Lisa Brooks, *Our Beloved Kin: A New History of King Philip's War* (New Haven, CT: Yale University Press, 2017).

Some of the best recent research in historical social science makes the point quite well. Consider Tomila Lankina's path-breaking analysis of middle-class formation in Russia from the tsarist regime through communism. In *The Estate Origins of Democracy in Russia*—recently awarded the J. David Greenstone Prize of the American Political Science Association—Lankina upends common understandings in the study of Russia and in comparative politics alike. Far from leveling status and economic inequalities, the Soviet Union relied heavily upon, and even replicated, many of the tsarist estate class structures that preceded it. Lankina posits four pathways—education, professional incorporation, social closure, and time—through which the tsarist estates “colonized” the new regime’s “expertise

²¹ Robert Dahl, *Who Governs? Democracy and Power in an American City* (New Haven, CT: Yale University Press, 1961). As James Mahoney remarks in a review that came out before this one, *Time Counts* “does less to advance methods in the qualitative tradition” and that “many historians will see this book as undervaluing the importance of noncausal interpretation based on deep contextual knowledge, normative sensibilities, and narrative explanation.” Mahoney review of *Time Counts*, *Perspective on Politics* 20, no. 4 1475–76. For more general examination of cases that transcend countability, see Mahoney, “The Logic of Process Tracing Tests in the Social Sciences,” *Sociological Methods and Research* 41, no. 4 570–97, especially 571, 578–81 (on regime-specific causal analysis), and 590–93 (examples of regime-specific causal analysis in *longue durée* English political development).

With a combination of skill and sagacity that no two other scholars would likely have brought to the match, Wawro and Katznelson have productively identified an interaction at which quantitative historical research can be plausibly conducted. What is needed from the next generation of research is more on the boundary conditions that attach to this exercise. One might worry that readers of *Time Counts* (and to be fair, many other quantitative studies of history) might take flexible and historically contextualized parameter methods and run full sprint towards the practice of not needing to consult historiography and its methods at all. Assuming we have countability and some decent priors on the historicity of the process or a model that learns from the quantitative data alone, do we need historical methods anymore at all? What use is left for archives, narrative, or what is called historiography? The inquiry we most need is not merely a combination of narrative and quantitative approaches but a methodological treatment about how exactly narrative and quantification should be combined—or should not.

²² Tomila Lankina, *The Estate Origins of Democracy in Russia* (New York: Cambridge University Press, 2022), Chapter 1, especially 17, 22–31, 35–37, Chapters 3 and 4, e.g., 130–31 (Samara-based exemplar), Chapter 8, 299–309 (Samara).