Making Law Modern:
Discovering the Restructuring of Law through Legal Citations in American Treatises

Kellen Funk
Columbia Law School
krf2138@columbia.edu

Lincoln Mullen
Roy Rosenzweig Center for History and New Media
lincoln@lincolnmullen.com

From the founding of the United States to the year 1920, private entrepreneurs and publicly subsidized reporters published around 1.4 million judicial decisions from America's vast array of local, state, and federal courts. To get a handle on the ever-burgeoning wave of precedents, practitioners in turn published over 10,000 digests, commentaries, and treatises over the same period. Our project seeks to understand, document, and visualize the interconnections between these two corpora. What follows is a historical introduction to the problem and a detailed discussion of our method for detecting and compiling case citations, followed by a few preliminary examples of the kinds of analysis made possible by our historically informed method. We aim to show how American law was restructured from common law categories of writs and tenures to the modern categories now familiar to the law school curriculum.

We have derived our corpora from two databases of scanned nineteenth-century volumes. Gale Group's Making of Modern Law: Legal Treatises database (MOML) is a comprehensive collection of legal commentaries, digests, and treatises (as well as biographies and law reform tracts) published in England and America from the year 1800 to 1920, a period commonly regarded as the golden age of legal treatise literature in the United States. Harvard Law Library's Caselaw Access Project (CAP) offers a comprehensive collection of published decisional law from virtually all American state, federal, and territorial courts of record from the founding of the nation to the present. Using the methods described in Part II, we have detected across the MOML treatises 8.2 million citations to approximately 368,000 cases whose text and metadata are available in CAP. Part III presents a preliminary analysis of the citational data. In future work, we plan to compile a database of citations from the cases themselves—allowing us to explore case-to-case citations in addition to treatise-to-case citations—and to incorporate corpora of case law from the United Kingdom, allowing us to track transatlantic citations.

I. Background: The Extensive Writtenness of Unwritten Law

The famously acerbic reformer Jeremy Bentham once complained in a letter to the U.S. President James Madison that the usual description of the common law as "unwritten law" was mistaken, "for of *writing* there is beyond comparison."[1] Since at least the time of Lord Coke's judicial career (1606 to 1616), enterprising lawyers and publishers had begun to systematically

[1] Jeremy Bentham to James Madison, October 30, 1811, in *The Papers of James Madison, Presidential Series, vol. 3, 3 November 1810–4 November 1811*, ed. J. C. A. Stagg et al. (UVA 1996), 505–35. Bentham wrote with an offer to codify the common law into a single volume of statutes. Madison, after five years delay, politely declined Bentham's offer, assuring the philosopher that the Revolution had sufficiently "lopped off" so much of England's common law that the crisis of unwritten law was less severe in the United States than Bentham imagined. See *id.*, James Madison to Jeremy Bentham, May 8, 1816.

"report" judicial decisions from the courts at Westminster, significantly expanding the amount of written decisional law across the eighteenth century.[2]

As it would turn out, the English situation was not "beyond comparison" as Bentham had it. Around the time of Bentham's letter to Madison, case reporting in the United States began to dramatically outstrip the production of the English courts. The English Reports, a comprehensive collection of most extant case reports from the English royal courts, boasted 124,882 case reports from the thirteenth century to 1873.[3] By that same year, *twice* as many American case decisions (269,635) had been published in a tiny fraction of the time.

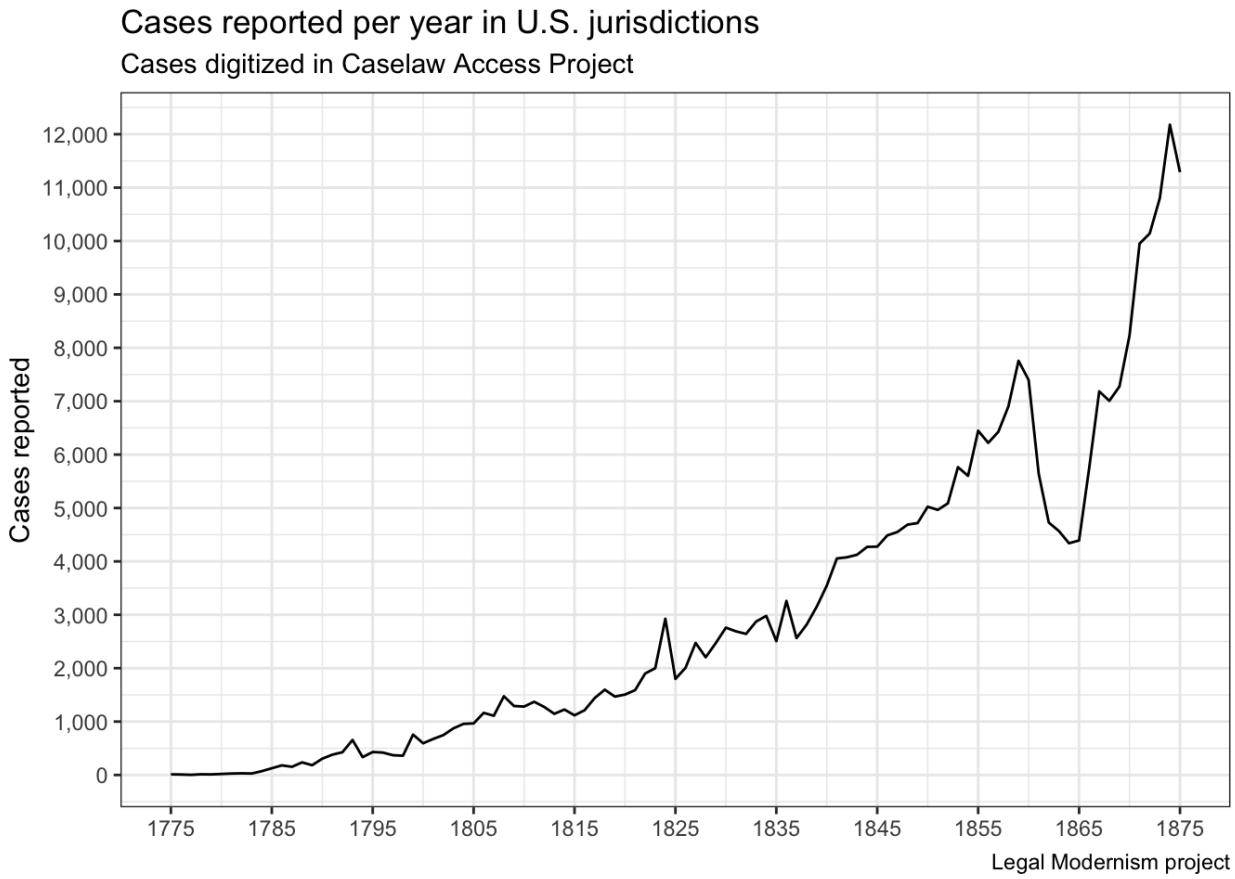


Figure 1. The "unwritten" common law was written down extensively in American case law reports which grew rapidly throughout the nineteenth century.

[2] William S. Holdsworth, *A History of English Law* (1966), 5:373–74, 12:111, 13:425–27, 15:248–50; W. T. S. Daniel, *History of the Origin of the Law Reports* (London, 1884); T. F. T. Plucknett, "The Genesis of Coke's Reports," 27 Cornell Law Quarterly 190 (1942).

[3] A full text database of the English Reports is available at http://www.commonlii.org/uk/cases/EngR/. The English Reports are explored in Peter Grajzl & Peter Murrell, "Using Topic-Modeling in Legal History, with an Application to Pre-Industrial English Case Law on Finance," 40 *Law & History Review* 189 (2022); Peter Grajzl & Peter Murrell, "A Machine-Learning History of English Caselaw and Legal Ideas Prior to the Industrial Revolution I: Generating and Interpreting the Estimates," 17 *Journal of Institutional Economics* 1 (2021); Peter Grajzl & Peter Murrell, "A Machine-Learning History of English Caselaw and Legal Ideas Prior to the Industrial Revolution II: Applications," 17 *Journal of Institutional Economics* 201 (2021).

For a legal system that became so dependent on judicially announced precedent, the surprising feature of early American legal history is just how little published precedent there was at the beginning. Most courts treated reports of English decisional law as authoritative or else maintained an oral culture of precedent that relied on the memories of judges and lawyers to a living tradition of consistent practice. By 1800, only a few slender volumes of American decisions had been published from Connecticut to Virginia, with not a single published decision appearing further south. In 1804, both New York and Massachusetts initiated public subsidies for case reporters, and private enterprise began filling out volumes both north and south.[4] The explosion in U.S. case reporting was nothing short of a technological and intellectual revolution. While England had gradually transitioned from an oral to a written system of precedent over the course of three centuries, the United States took less than three decades.

By the 1840s, lawyers were complaining not that they had too little published precedent, but far too much. As the prominent legal ethicist David Hoffman lamented, “The increase of this portion of our legal literature within the last thirty years, has no parallel in the juridical history of any other country. More than four hundred and fifty volumes of American law reports now load our shelves!” Another attorney estimated that a minimally adequate law library should contain 800 to 1,000 volumes. As counsel drew on all these disparate sources for their arguments, one lawyer despondently wrote, “Who can guess what he may have to meet in a law suit, as no lawyer can afford to buy or read all the books in the world?”[5]

It appears the lawyers were right to worry, for even New York—far and away the most influential jurisdiction in the antebellum period—refused to confine itself to its own decisional law. By mid-century New York courts were citing in thousands of instances the judicial opinions of every other state in the Union as persuasive authority on what the law was or should be.[6]

[4] See John H. Langbein, “Chancellor Kent and the History of Legal Literature,” 93 *Columbia Law Review* 547 (1993).
[5] David Hoffman, *A Course of Legal Study* (2d ed. 1846), 657; Michael Hoffman, “Letter on Reforms Necessary in the Body of Law, in the Written Pleadings, and in the Practice of the Courts” (Mar. 21, 1846), in Thomas Prentice Kettle, ed., *Constitutional Reform in a Series of Articles Contributed to the Democratic Review* (1846), 63–64; Hiram P. Hastings, *An Essay on Constitutional Reform* (1846), 27.
[6] We do not undertake the analysis here, but these patterns might be productively compared to European courts, which likewise rely on but are not bound by one another’s decisions. See Erik Voeten, “Borrowing and Nonborrowing among International Courts,” 39 *Journal of Legal Studies* 547 (2010); Martin Gelter & Mathias M. Siems, “Citations to Foreign Courts—Illegitimate and Superfluous, or Unavoidable? Evidence from Europe,” 62 *American Journal of Comparative Law* 62 (2014).

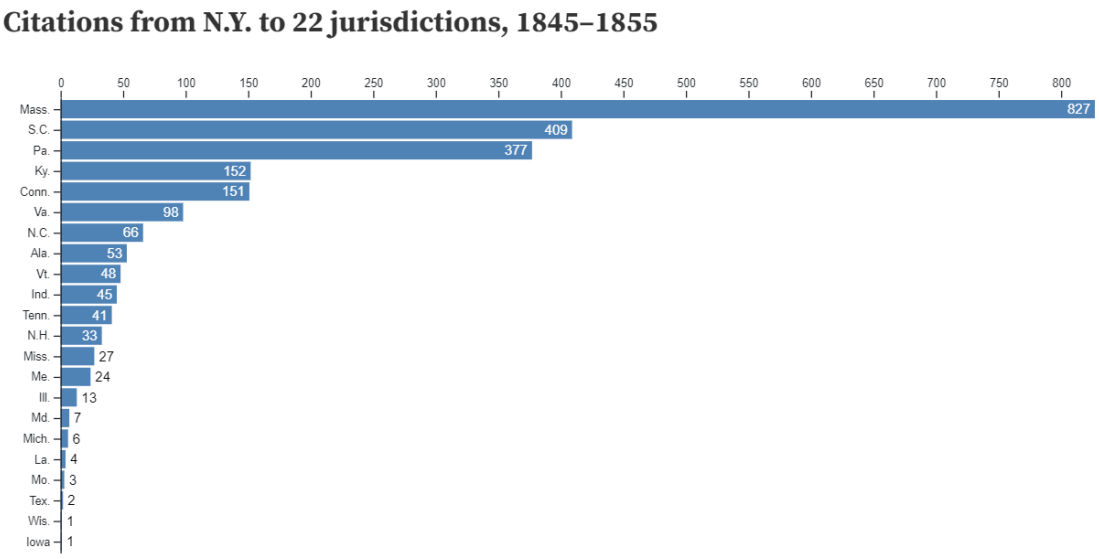
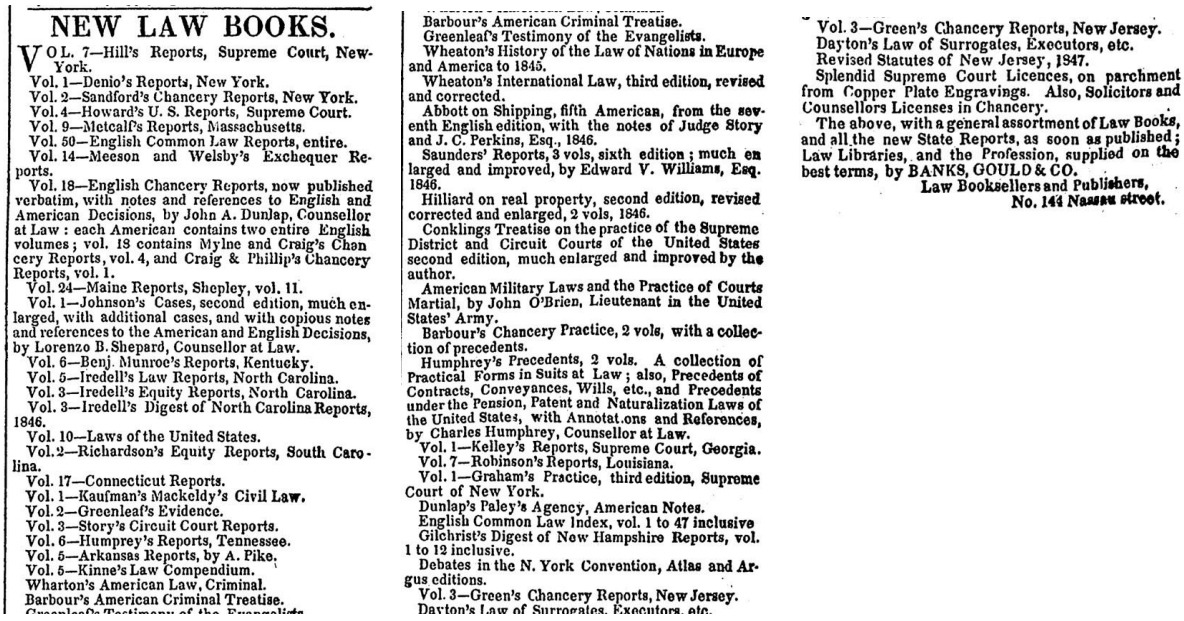


Figure 2. In the decade from 1845 to 1855, New York cited 22 other jurisdictions in its case law, with Massachusetts by far the largest jurisdiction from which it imported law. Data is drawn from the Caselaw Access Project's reports of citations between cases. For an interactive figure, see <https://observablehq.com/@lmullen/legal-citations-across-state-jurisdictions>.

No wonder the lawyers worried. Indeed, who could guess but that a recent decision out of far-flung Texas might be dredged up by one's opponent and persuade the judge in the next case? Advertisements like the following, from 1847 New York, filled practitioners with dread: decisional law only ever multiplied as the young nation grew.



NEW LAW BOOKS.

VOL. 7—Hill's Reports, Supreme Court, New-York.
Vol. 1—Denio's Reports, New York.
Vol. 2—Sandford's Chancery Reports, New York.
Vol. 4—Howard's U. S. Reports, Supreme Court.
Vol. 9—Metcalf's Reports, Massachusetts.
Vol. 50—English Common Law Reports, entire.
Vol. 14—Meeson and Welsby's Exchequer Reports.
Vol. 18—English Chancery Reports, now published verbatim, with notes and references to English and American Decisions, by John A. Dunlap, Counsellor at Law: each American contains two entire English volumes; vol. 18 contains Mylne and Craig's Chancery Reports, vol. 4, and Craig & Phillip's Chancery Reports, vol. 1.
Vol. 24—Maine Reports, Shepley, vol. 11.
Vol. 1—Johnson's Cases, second edition, much enlarged, with additional cases, and with copious notes and references to the American and English Decisions, by Lorenzo B. Shepard, Counsellor at Law.
Vol. 6—Benj. Munroe's Reports, Kentucky.
Vol. 5—Iredell's Law Reports, North Carolina.
Vol. 3—Iredell's Equity Reports, North Carolina.
Vol. 3—Iredell's Digest of North Carolina Reports, 1846.
Vol. 10—Laws of the United States.
Vol. 2—Richardson's Equity Reports, South Carolina.
Vol. 17—Connecticut Reports.
Vol. 1—Kaufman's Mackeldy's Civil Law.
Vol. 2—Greenleaf's Evidence.
Vol. 3—Story's Circuit Court Reports.
Vol. 6—Humprey's Reports, Tennessee.
Vol. 5—Arkansas Reports, by A. Pike.
Vol. 5—Kinne's Law Compendium.
Wharton's American Law, Criminal.
Barbour's American Criminal Treatise.

Barbour's American Criminal Treatise.
Greenleaf's Testimony of the Evangelists.
Wheaton's History of the Law of Nations in Europe and America to 1845.
Wheaton's International Law, third edition, revised and corrected.
Abbott on Shipping, fifth American, from the seventh English edition, with the notes of Judge Story and J. C. Perkins, Esq., 1846.
Saunders' Reports, 3 vols, sixth edition; much enlarged and improved, by Edward V. Williams, Esq. 1846.
Hilliard on real property, second edition, revised corrected and enlarged, 2 vols, 1846.
Conklings Treatise on the practice of the Supreme District and Circuit Courts of the United States second edition, much enlarged and improved by the author.
American Military Laws and the Practice of Courts Martial, by John O'Brien, Lieutenant in the United States' Army.
Barbour's Chancery Practice, 2 vols, with a collection of precedents.
Humphrey's Precedents, 2 vols. A collection of Practical Forms in Suits at Law; also, Precedents of Contracts, Conveyances, Wills, etc., and Precedents under the Pension, Patent and Naturalization Laws of the United States, with Annotations and References, by Charles Humphrey, Counsellor at Law.
Vol. 1—Kelley's Reports, Supreme Court, Georgia.
Vol. 7—Robinson's Reports, Louisiana.
Vol. 1—Graham's Practice, third edition, Supreme Court of New York.
Dunlap's Paley's Agency, American Notes.
English Common Law Index, vol. 1 to 47 inclusive
Gilchrist's Digest of New Hampshire Reports, vol. 1 to 12 inclusive.
Debates in the N. York Convention, Atlas and Argus editions.
Vol. 3—Green's Chancery Reports, New Jersey.
Dayton's Law of Surrogates, Executors, etc.

Vol. 3—Green's Chancery Reports, New Jersey.
Dayton's Law of Surrogates, Executors, etc.
Revised Statutes of New Jersey, 1847.
Splendid Supreme Court Licences, on parchment from Copper Plate Engravings. Also, Solicitors and Counsellors Licenses in Chancery.
The above, with a general assortment of Law Books, and all the new State Reports, as soon as published; Law Libraries, and the Profession, supplied on the best terms, by BANKS, GOULD & CO.
Law Booksellers and Publishers,
No. 144 Nassau street.

Figure 3. National Police Gazette *(New York, N.Y.), June 5, 1847.*

One resource for collecting what one lawyer called the "overwhelming showers" of precedents raining down on the bar was the legal treatise.[7] Treatises had a long history in theology, medicine, and law as a device for organizing knowledge, but the legal treatise especially entered a golden age in the nineteenth century.[8] Around ten thousand treatises were published in the United States from 1800 to 1920, with another ten thousand appearing in England over the same timespan.[9] Some treatises, like James Kent's *Commentaries on American Law* spanned multiple volumes and dozens of editions and attempted to synthesize all legal knowledge at a fairly high level of generality. Others could be far more specialized, detailing for instance the law of water rights in the West, the legal regulation of municipal corporations, or the routine practices of bankruptcy courts.[10]

Indeed, legal treatises took such a variety of approaches to their framing, content, and extent of coverage that contemporary scholarship has struggled to define treatise literature with precision. Any description of a treatise's scope and aim soon calls up a host of counter-examples. Many treatises aimed to provide a quick and ready digest or index of recently decided cases, while many others aimed to synthesize areas of the law into general principles that skimmed over the top of the noisy case law.[11] Many treatises provided an introduction to their subject to neophytes entering the profession, while many others sought to persuade the master jurists on arcane points of possible legal development.[12] Joseph Story—a prolific treatise author—believed treatises could serve as stepping stones to codification, while Herbert Wechsler—architect of

[7] William Sampson, "Anniversary Discourse on the Origin, Progress, Antiquities, Curiosities, and Nature of the Common Law, Delivered Before the New York Historical Society, December 6, 1823," in Pishey Thompson, ed., *Sampson's Discourse and Correspondence with Various Learned Jurists Upon the History of the Law* (1826), 38.

[8] Historiographical interest in legal treatises has recently revived thanks in part to Angela Fernandez & Markus D. Dubber, *Law Books in Action: Essays on the Anglo-American Legal Treatise* (2012). For responsive essays, see Richard A. Danner, "Foreword: Oh, the Treatise!," 111 *Michigan Law Review* 821 (2013); Steven Wilf, "The Legal Treatise," in Simon Stern, Maksymillian del Mar, and Bernadette Meyler, eds., *The Oxford Handbook of Law and the Humanities* (2020). For the classic trilogy of historiographical treatment of treatises as such, see G. Edward White, *The Marshall Court and Cultural Change, 1815–1835* (1988), 81–104; A. W. B. Simpson, "The Rise and Fall of the Legal Treatise: Legal Principles and the Forms of Legal Literature," 48 *University of Chicago Law Review* 632 (1981); Morton J. Horwitz, "Part III—Treatise Literature," 69 *Law Library Journal* 460 (1976).

[9] The Making of Modern Law corpus includes 9,749 treatises labeled as U.S. and published between those dates, not counting miscellaneous other documents included in the corpus such as biographies, addresses, and reports of single cases.

[10] E.g., John Norton Pomeroy & Henry Campbell Black, *A Treatise on the Law of Water Rights as the Same is Formulated and Applied in the Pacific States* (rev. ed. 1893); John F. Dillon, *Treatise on the Law of Municipal Corporations* (1872); Thorndike Saunders, *Bankruptcy Practice under the Law of the United States of 1867* (1868).

[11] See Angela Fernandez & Markus D. Dubber, "Introduction: Putting the Legal Treatise in Its Place," in *Law Books in Action*, 1–21.

[12] See Lindsay Farmer, "Of Treatises and Textbooks: The Literature of the Criminal Law in Nineteenth-Century Britain," in *Law Books in Action*, 145–64; Jim Phillips, "A Low Law Counter Treatise? "Absentees" to "Wreck" in British North America's First Justice of the Peace Manual," in *id.*, 202–19.

the Model Penal Code—believed codes served as stepping stones to great treatises.[13] And so on.[14]

Case reports were somewhat more standardized in form and purpose, but could still vary in important regards over time. John Langbein dates the canonical form of case reporting to Sir James Burrow's English court reports from 1756 to 1772. Burrows "omitted all cases where the Question turned upon Facts and Evidence" instead of contestable points of law. The cases Burrow did include began with Burrow's own construction of the facts (drawn from pleadings and in-court statements), a reconstruction of the lawyers' arguments (including lists of cited authorities), and the opinion of the court. American publishers largely followed Burrows's system. Early case reports indulged in more paraphrasing than did later ones, and over time reliance on judges to script and edit their own opinions increased. Lawyers' arguments gradually disappeared, while statements of facts were given by the court opinion rather than by the reporter. Concurring and dissenting opinions grew in length and sophistication, the more to give the impression that they were indeed rivalrous alternatives to the opinion of the majority.[15]

Both treatises and case reports thus ultimately arrived at a similar aim. Both sought to suppress the particular facts of decided cases while extracting rules of law that could be applied in future cases and scientifically arranged to guide practitioners along. In a work subtitled *Explaining the Nature, Sources, and Practical Applications of Legal Science*, Joel Prentiss Bishop, a prolific author of American treatises, defined the legal treatise as "an orderly statement of those principles in which law consists, whether drawn from the reports of law cases, from natural reason, or from any other source."[16] It remains unclear just how much of "any other source" really made it into the legal treatises. Only the first item on Bishop's list, the rules drawn from the reports of law cases, is unmistakably present in every volume and on nearly every page. Case reports, that is, were the building blocks of treatises, which were in turn the building blocks of legal knowledge and the emerging "science of law" in the nineteenth century.

While we can and will run our detection algorithm to find citations in the case reports themselves, we believe starting with and keeping the focus on treatise literature offers a novel opportunity to analyze genre and network formations in a relatively unexplored field. There are countless computational studies of judicial citation practices, some produced as early as the 1930s.[17] By contrast, other than a few close readings by cultural historians, citational practices

---

[13] See Barry Wright, "Renovate or Rebuild? Treatises, Digests and Criminal Law Codification," in *Law Books in Action*, 181–201.
[14] For further explorations of the binary qualities of legal treatise literature, see Wilf, *The Legal Treatise*; Christopher Tomlins, "Effects of Scale: Toward a History of the Literature of Law," in *Law Books in Action*, 220–42.
[15] See John H. Langbein, Renee Lettow Lerner, & Bruce P. Smith, *History of the Common Law: The Development of Anglo-American Institutions* (Aspen 2009), 822–37.
[16] Joel Prentiss Bishop, *The First Book of the Law: Explaining the Nature, Sources, and Practical Applications of Legal Science, and the Methods of Study and Practice* (1868), 137.
[17] See, e.g., Rodney L. Mott, "Judicial Influence," 30 *American Political Science Review* 295 (1936); John H. Merryman, "Toward a Theory of Citations: An Empirical Study of the Citation Practice of the California Supreme Court in 1950, 1960, and 1970," 50 *Southern California Law Review* 381 (1977); Lawrence M. Friedman, Robert A. Kagan, Bliss Cartwright & Stanton Wheeler, "State Supreme Courts: A Century of

in treatise literature has been largely overlooked or impractical to study before the mass digitization efforts of the last decade.[18] Facing incentives of profit and prestige different from judges sitting within an institutional hierarchy, treatise writers offer a different set of network dynamics to investigate questions of "influence" or "bias."[19]

Overall, the unparalleled production and organization of written law distinguished nineteenth-century jurisprudence from the common law tradition that preceded it, a revolution in legal epistemology commonly labeled "modernism," though scholars have disputed which exact characteristics mark the divide between modern and pre-modern law.[20] Our aim in this project is to better define and understand legal modernism through its textual production and features. We seek to understand how the authors of legal treatises created the categories of modern law from the building blocks of cases. This aim is not only technologically feasible with contemporary methods of computational text analysis but also, we would argue, historically justified because it heeds the instruction of the very architects and theorists of legal modernization themselves.

Start, for instance, with William Blackstone's *Commentaries on the Laws of England*, widely recognized as the most influential legal treatise ever written in English. In a pivotal passage in the first volume, Blackstone justified the entire common law system and the authority of judge-decreed law by reference to its technology of knowledge. Judges, according to Blackstone, "are the depositaries of the laws; the living oracles, who must decide in all cases of doubt, and who are bound by an oath to decide according to the law of the land. The knowledge

---

Style and Citation," 33 *Stanford Law Review* 773 (1981); David J. Walsh, "On the Meaning and Pattern of Legal Citations: Evidence from State Wrongful Discharge Precedent Cases," 31 *Law & Society Review* 337 (1997); Rachel K. Hinkle & Michael J. Nelson, "The Transmission of Legal Precedent among State Supreme Courts in the Twenty-First Century," 16 *State Politics & Policy Quarterly* 391 (2016). For a comprehensive review, see Jens Frankenreiter, "Promises and Challenges of Studying Citations in Judicial Opinions," in *Oxford Handbook of Comparative Judicial Behavior* (forthcoming).

[18] For close reading analysis of citational practices, see especially Barbara Young Welke, *Recasting American Liberty: Gender, Race, Law and the Railroad Revolution, 1865–1920* (2001); Hendrik Hartog, *Public Property and Private Power: The Corporation of the City of New York in American Law, 1730–1870* (1983); Janet Halley, "What Is Family Law?: A Genealogy," 23 *Yale J. of L. & the Humanities* 1 (2011).

[19] For salient studies of these phenomena in the judicial context, see John Szmer, Robert K. Christensen & Samuel Grubbs, "What Influences the Influence of U.S. Courts of Appeals Decisions?" 49 *European Journal of Law & Economics* 55 (2020); Montgomery N. Kosma, "Measuring the Influence of Supreme Court Justices," 27 *Journal of Legal Studies* 333 (1998); Frank B. Cross, "The Ideology of Supreme Court Opinions and Citations," 97 *Iowa Law Review* 693 (2012); Stephan J. Choi & G. Mitu Gulati, "Bias in Judicial Citations: A Window into the Behavior of Judges?" 37 *Journal of Legal Studies* 87 (2008). On institutional dynamics, see Kirk A. Randazzo, "Strategic Anticipation and the Hierarchy of Justice in U.S. District Courts," 36 *American Politics Research* 669 (2008); Ali S. Masood, Benjamin J. Kassow & Donald R. Songer, "Supreme Court Precedent in a Judicial Hierarchy," 45 *American Politics Research* 403 (2017); Rachel K. Hinkle, "Legal Constraint in the U.S. Courts of Appeals," 77 *Journal of Politics* 721 (2015).

[20] See Steven Wilf, "The Invention of Legal Primitivism," 10 *Theoretical Inquiries in Law* 485 (2009); James Q. Whitman, "The Transition to Modernity," in Markus D. Dubber & Tatjana Hörnle, eds., *The Oxford Handbook of Criminal Law* (Oxford 2014), 91–94; Kunal M. Parker, *Common Law, History, and Democracy in America, 1790–1900: Legal Thought before Modernism* (Cambridge 2011); Roger Berkowitz, *The Gift of Science: Leibniz and the Modern Legal Tradition* (Fordham 2010).

of that law is derived from experience and study; from the [nighttime labors of many years]. . . and from being long personally accustomed to the judicial decisions of their predecessors."[21]

The *Commentaries* (themselves an oral production before an enterprising publisher reduced them to writing) linked quasi-religious authority to a technological one: common law judges have a kind of access to decisional law that others do not. Through their craft disciplines and their learning they keep alive within themselves the received traditions of the law, a law that like the *Commentaries* themselves, are primarily oral and not textual. It was exactly this quasi-mystical premise that provoked Jeremy Bentham and his American acolytes to rail against the "spurious" unwritten common law and its "superstitious votaries."[22]

The links between religious authority and technologies of the book would be a major theme in an essay by Samuel Taylor Coleridge a half century later. Coleridge wrote, "In times of old books were as religious oracles; as literature advanced, they next became venerable preceptors; [but] as their *numbers* increased, they sank still lower [until] at present they seem degraded into culprits to hold up their hands at the bar of every self-elected, yet not the less peremptory, judge."[23] Coleridge's complaint sounds remarkably familiar, reminiscent of the lawyers' laments over the ever-expanding publishing of case reports. Indeed, it seems that case reports are intended to do exactly what Coleridge charged, to break the oracular authority of the judges by taking their special access to decisional law and publicizing it to every willing purchaser.

While men of letters like Coleridge bewailed the bastardization of their craft, American judges embraced the technological revolution. If the precedents were taken from the minds of jurists and reduced to printed texts, the jurists could at least maintain their status by demonstrating their continued mastery of the literature. As New York's Chancellor James Kent explained, his own *Commentaries* were designed to solve "the evils resulting from an indigestible heap of laws, and legal authorities [which] destroy the certainty of the law, and promote litigation, delay and subtilty."[24] And in his own reported decisions, Kent showed lawyers how to cite and argue from written precedent rather than oral practices, drawing upon a whole library of legal literature for his reasoned decisions.[25]

As historians of the book have long been aware, methods of organizing knowledge can imperceptibly shift into the production of knowledge, setting implicit boundaries on the world of "thinkable thoughts."[26] In perhaps the quintessential essay on modernity in art, Walter Benjamin

---

[21] William Blackstone, *Commentaries on the Law of England* (1765), 1:68–70.
[22] See David Lieberman, "Bentham's Jurisprudence and Democratic Theory: An Alternative to Hart's Approach," in Xiaobo Zhai, ed., *Bentham's Theory of Law and Public Opinion* (Cambridge 2014), 136–39.
[23] Samuel Taylor Coleridge, *Biographia Literaria* (1817), 1:57.
[24] James Kent, *Commentaries on American Law* (1826), 1:441–42.
[25] See Langbein, "Chancellor Kent and the History of Legal Literature,"; Daniel J. Hulsebosch, *Constituting Empire: New York and the Transformation of Constitutionalism in the Atlantic World, 1664–1830* (UNC 2005), 283–87.
[26] Robert C. Berring, "Legal Research and the World of Thinkable Thoughts," 2 *J. App. Prac. & Process* 305 (2000). See also Maurice S. Lee, *Overwhelmed: Literature, Aesthetics, and the Nineteenth-Century*

argued that the mechanical reproducibility of a work of art in an industrial age made public and massively available what had been private and usually confined to religious, sacred spaces. Benjamin referred to this as the "aura" of an original work of art, and he argued that mass reproduction of an original work of art not only lacked the aura of the original, but in some sense it destroyed the aura of the original by profaning it. Technological revolutions not only changed how art was produced after the revolution, but what art from before the revolution even meant or could mean.[27]

Overall, our project aims to explore the work of law in the age of its mechanical reproduction. It seeks to understand how legal knowledge was created, transmitted, and organized into the categories we now recognize as fundamental, the proverbial ways of "thinking like a lawyer." To get at these broad questions, we start where the treatises start, with appeals to authority in the form of citations to decided cases and their published opinions.

II. Method of Citation Detection

For this project, we draw on two large corpora of legal texts. Our primary corpus is Gale's Making of Modern Law: Legal Treatises database (hereafter MOML). This corpus contains some 10,498 distinct editions of legal treatises published in the United States from 1800 to 1926.[28] The other corpus we draw on is Harvard's Caselaw Access Project (hereafter CAP), an open access library of reported cases in the United States.[29] Limiting the corpus to the same years covered by MOML (1800 to 1926) yields just over 1.6 million reported U.S. cases. Our work studying these corpora is available in a GitHub repository.[30]

As noted above, what went into a case report or a treatise or commentary could vary significantly over time and depending on purpose. Fortunately for a computational project, there was one thing in both the treatises and case reports that remained absolutely constant across our period and across the two corpora: the citational form. Whether or not a citation to a case decision included the party names, identity of the rendering court, or the date (all three became more common over time), from the eighteenth century to the twentieth, every citation at the very least contained what both lawyers and programmers could call a "regular expression":

Information Revolution (2019); Andrew Piper, *Enumerations: Data and Literary Study* (2018); Robert Darnton, *The Case for Books: Past, Present, and Future* (2010).

[27] Walter Benjamin, *The Work of Art in the Age of Mechanical Reproduction* (1935); see also Miriam Bratu Hansen, "Benjamin's Aura," 34 *Crit. Inquiry* 336 (2008).

[28] For details on the compilation of MOML, see Bennett Graff, "The Making of Gale's Making of Modern Law," 21 *Legal Information Management* 28 (2021). MOML contains some 25,017 volumes, including both British and U.S. treatises. But it also contains a fair number of addresses, papers of prominent legal figures, and biographies. And of course legal treatises were sometimes multivolume works. For the purposes of this paper, we have removed everything which is obviously not a legal treatise from the United States, and we have treated multi-volume treatises as a single work. We have treated multiple editions of the same work as if they were distinct publications, because the editions were often separated by many years and made reference to the case law which had been reported in that ensuing period.

[29] See https://case.law/about/.

[30] https://github.com/lmullen/legal-modernism.

`[volume number`[31]`] [reporter name`[32]`] [initial page of report`[33]`]`

Occasionally two or more very short case reports might begin on the same page,[34] but setting that problem aside for now, one could think of every case as having a single address formed by the unique combination of those three components. Given a perfectly accurate OCR transcription of a legal text, one could simply search for each address as a regular expression to detect and compile a table of citations.

That indeed is the approach of the Free Law Project's "eyecite citation extractor," a tool used by CAP itself to automate hyperlinking between cases cited throughout its corpus. "eyecite" uses a regular expression database keyed to known official reporters to parse clean text and extract citations, citations which in turn can be easily associated with the reported case in CAP.[35]

There are two major limitations to employing "eyecite" in our project, one technological and one historical. The technological problem is one common to many large historical corpora: the OCR transcriptions available in MOML and CAP are fairly messy, while "eyecite" assumes relatively clean text. Its regular expressions database does not employ wildcards, stemming, or fuzziness features but rather relies on precise citation of known reporters. Although "eyecite" includes packages for pre-processing and cleaning text, these tools are so far not well calibrated to the variety of errors that crop up when attempting to transcribe scans of nineteenth-century texts.

Part of our method involves dealing with such variations in the OCR. We do so by creating a list of OCR variations (prioritized by how frequently they appear) and mapping them to the standard reporter name. For example, the Atlantic Reporter has a standard citation "A." We list 52 variations of that, both intentional (e.g., "Atl." and "Atl. Rep") and as the result of OCR errors (e.g., "Atl. Itep."). We have painstakingly compiled this list alongside two law student research assistants. The virtue of this approach is that it allows us to take very messy citations and turn them into known good citations, which gives us much more confidence at the analytical phase.

Even if the OCR output were cleaner or the pre-processing tools more robust, the major weakness of detecting citations through "eyecite" is a problem of history. "eyecite" is very

[31] Nearly always 1 to 3 digits.
[32] Reporter names, usually abbreviated, are highly but not perfectly standardized. While not a trivial wrinkle, the set of alternatives is manageable. For instance, Kelly's Georgia Reports would usually be cited as `Kelly`, but might appear as `Kelly's`, `Kelly R.`, `Kelly Rep.`, or any of the preceding with the specifying parenthetical `(Ga.)`, `(Georgia)`, or `(Georgia Reports)`.
[33] Virtually always 1 to 4 digits. Legal writers familiar with the practice of "pinciting" know that the precise page of a quotation or direct authority for a proposition may be supplied after the initial page (following a comma), or in place of it in the "short form" citation. In standard usage, the short form citation is only employed after a citation has first appeared in full (following the regular expression given here). Since our aim for now is to track *the* citation of *a* case in a single source, and not multiple citations to the same case, short form citations can be disregarded for the present.
[34] In our experience, this phenomenon seems much more common in the English case reports than in American reports, which tend to occupy at least one or two pages per case.
[35] Jack Cushman, Matthew Dahl, and Michael Lissner, "eyecite: A tool for parsing legal citations," (August 2021), available at https://joss.theoj.org/papers/10.21105/joss.03617.

well-calibrated to extracting citations that follow modern standardized forms. “eyecite” is much less successful at finding historically accurate but now obsolete citations. But more than that, the tool’s extraction method actually obscures historical citation practices.

We can best understand the problem by working through an actual historical example. Before states publicly subsidized official reports, case reports were usually known by the name of the enterprising lawyer or publisher who compiled them. In some instances, these reports—appropriately styled “nominate” reports—continue to be referenced even today by the name of their original publisher. A common nominate reporter still widely cited today is Oliver Lorenzo Barbour’s reports of common law decisions in the State of New York, a series of reports that ultimately ran to 62 volumes.[36] Invariably, a case from Barbour’s Reports would be cited today as `[vol.] Barb. [page]`, “Barb.” being the universally accepted abbreviation for this nominate reporter. But today’s consensus is not yesterday’s, and over 2,000 times the MOML treatises cite `Barb. R.` instead, probably to make clear they are citing to Barbour’s law reports and not his comparatively less influential three volumes of New York chancery reports.[37]

That’s a fairly trivial variation. A more significant one arises from the fact that many nominate reporters *became* official reporters over time, as states adopted volumes that had been produced by private enterprise and transformed them into the official word of the state courts. That’s what happened, for instance, to James M. Kelly’s reports of cases decided in Georgia. Kelly produced three volumes that, for most of the nineteenth century, were cited with some variation as `[vol.] Kelly [page]`. Not long after Kelly completed his series, Georgia adopted it as the first three official volumes of the Georgia Reports. Today, one would cite a case in Kelly’s Reports as `[1/2/3] Ga. [page]`, a different address to arrive at the same case.[38] Indeed, from the beginning of the Georgia reports to the end of our period, U.S. treatises cited to the “official” version of the three volumes of the Georgia reports 4,185 times, but to the “nominate” version 2,361 times. To go by “eyecite” alone would miss nearly one third of the relevant citations to Georgia law.

[36] Available at https://cite.case.law/barb/.
[37] Available at https://cite.case.law/barb-ch/.
[38] Or at most one might include a parallel citation either of the form `[1/2/3] Ga. (Kelly) [page]` or `[1/2/3] Ga. [page], [1/2/3] Kelly [page].`

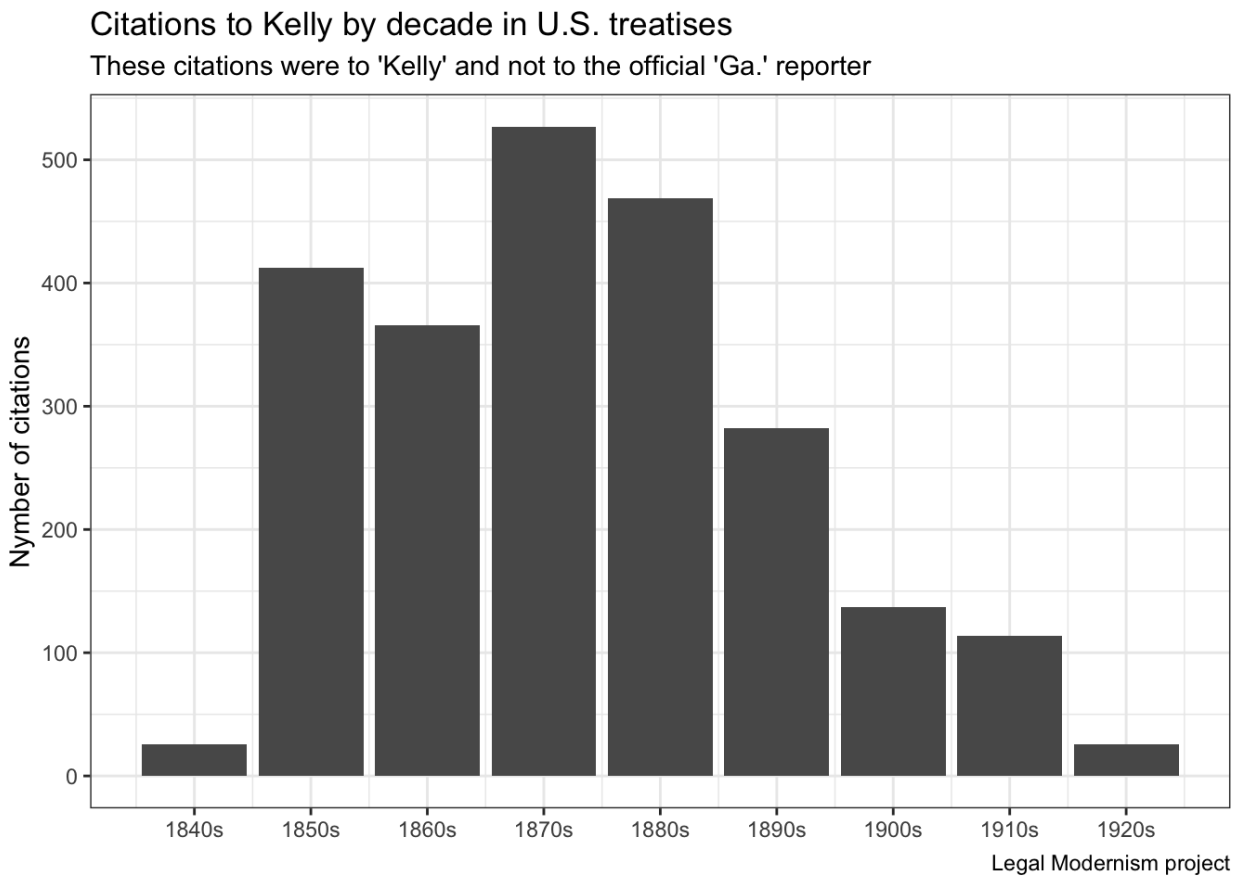


Figure 4. Legal treatises continued to cite Kelly's Reports instead of the official Georgia Reports throughout the nineteenth-century. The rate of citation to Kelly presumably declined as more treatise writers had the official Georgia volumes on their shelves, rather than the original Kelly volumes from which the first three Georgia volumes were derived.

In theory, one could just keep adding to the "eyecite" regular expressions database to account for all the variations while simultaneously employing a more calibrated OCR cleaner to make the text legible to the parsing in "eyecite." Both CAP and "eyecite" to some extent account for a small library of parallel cites (the same case cited in different reporters, as in "Kelly" and "Ga."). But our aim is not just to arrive at the right address—to correctly detect the citation of a uniquely identifiable case—but also to track historically how others arrived at that address, that is, how cases were actually cited in practice, allowing us to document, for instance, how long it took for nominate citations to trail off after states had officially incorporated the reports as their own.

So instead of compiling thousands of regular expression queries targeting precise citations, we have taken the opposite approach and began with the most capacious and historically justified regular expression extractor we could devise.

There are over 55 million occurrences of this expression in MOML, but of course many of those occurrences can be discarded because they are detecting pieces of a table of contents, or a citation of a treatise or statute (independently interesting, but not the target of the current

project). To clean up our data, we extracted the middle term of the expression and ranked the occurrences from most frequent to least. With the help of law student researchers, we then verified the genuineness of the most frequently detected terms to ensure they were indeed case reporters (or variations or textual distortions of known case reporters).

Occasionally a further historical wrinkle presented itself. When Georgia adopted the three volumes of Kelly's reports, those three volumes became the first three volumes of the Georgia reports. In other words, a citation to "1 Kelly" was equivalent to a citation to "1 Georgia." But other states that adopted nominate reporters as their official reports retroactively did not always use such a straightforward sequence. For example, the fourteen volumes of "Allen's Reports" were adopted as part of the official Massachusetts Reports. But "1 Allen" became "83 Mass." and "14 Allen" became "96. Mass." As another example, the eighteen volumes of Stewart's Equity Reports became the official New Jersey Equity Reports, beginning at volume 28. By relying on library catalogs of nominate reporters, including especially the one compiled by HeinOnline, we were able to compile tables of nominate reporters incorporated into different volumes of a jurisdiction's official reports.

Through this process we created a whitelist of candidate reporters, designating a standardized term for each reporter, tagging reporters from the UK (for future analysis), and indicating, where available, the corresponding standardized official reporter in CAP.

Once this process was complete, we were able to connect citations from specific pages in the Making of Modern Law treatises, with accompanying metadata for those treatises, to specific cases in the Caselaw Access Project, with the metadata for those cases. Altogether we have detected 8,239,691 treatise citations of cases cataloged in CAP.[39] This derivative dataset offers many possibilities for understanding how treatise authors created the categories of modern law out of the cases so voluminously reported and cited in the nineteenth-century United States.

III. Preliminary Analysis

In this paper we present preliminary results of four avenues of analysis. In the first, we show how the the age of the authorities changes over time in the treatises cited. In the second, we show a network of the relationship between the treatises, created by using the cases they cite as a bibliographic coupling. The third approach refines the second, narrowing down and focusing in on a network of cases cited in the civil procedure curricula of leading law schools. Our final section takes individual cases and tracks their citation patterns throughout the MOML corpus.

[39] Since we are indifferent to how many times the same treatise cites the same case, we are counting only one association between each treatise and cited case. We detected 12,145,000 citations matching CAP cases across all treatise pages in MOML, which reduces to the 8.1 million edges noted above once duplicate citations within the same treatise volume are ignored. The duplication of citations within a treatise, either because of extensive discussion or indexing of cited cases, gives us confidence that we are detecting genuine citations despite the messiness of the OCR. For comparison, "eyecite" detects 4,699,024 citations to CAP cases in the treatises, 43% fewer than those detected by our method.

A. Age of authorities cited

Until the twentieth century, cases were typically cited without any reference to the date of the decision in the citation or in the text of the commentary. It is impossible to discern just from the text of a treatise how old the precedential law is upon which it relies. But by linking case citations in MOML to the cases themselves in CAP, we can associate each citation with the date of its original publication and compare that to the date of the treatise commentary. The resulting picture is necessarily skewed by the absence of citations to English case law (for now). Still, we can trace the relative age of American legal rules sorted by the library catalog designations in MOML.

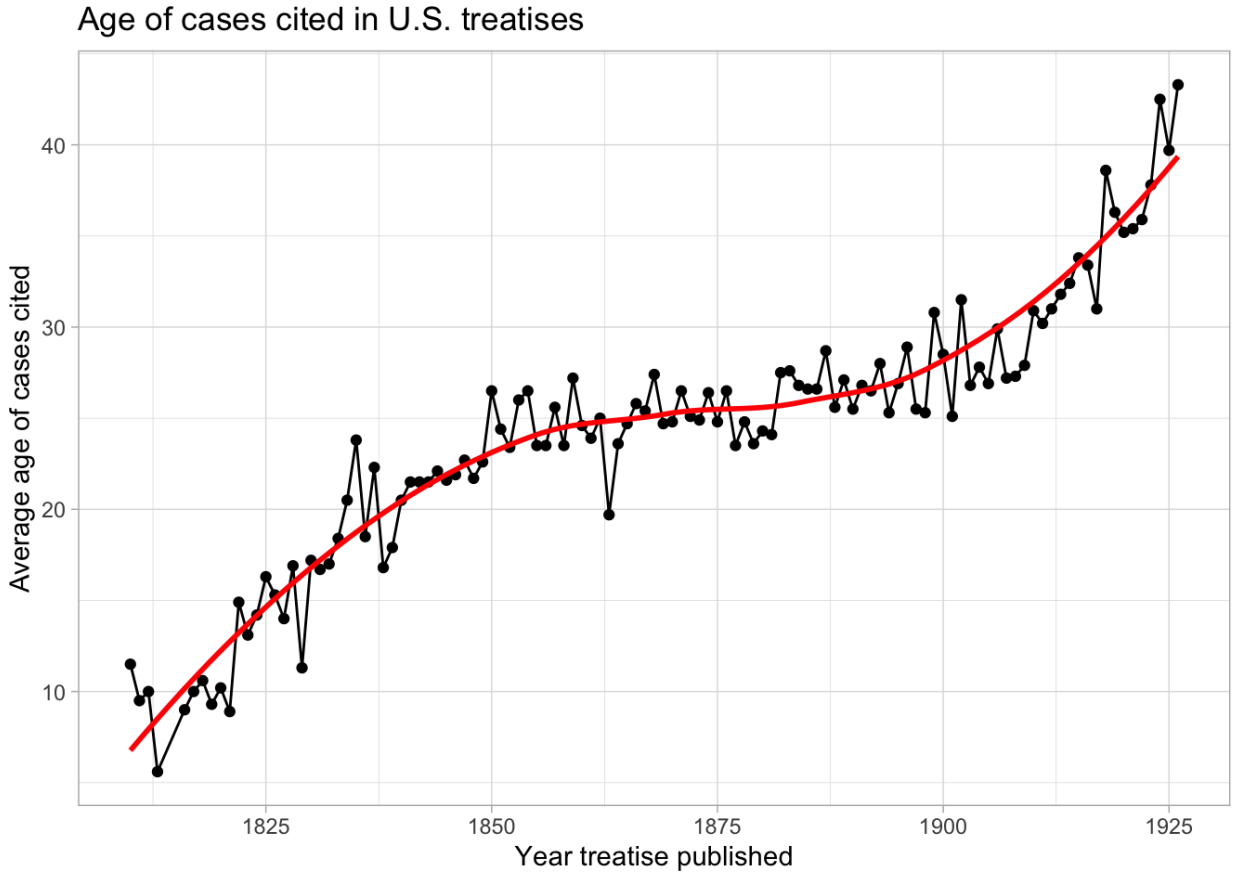


Figure 5. The average age of U.S. cases cited by American treatises climbed over time. But this growth was not uniform, demonstrating that at certain periods the development of the law sped up and new cases were more rapidly cited by treatise authors, and that in other periods the relevant law grew older over time.

Perhaps unsurprisingly, the age of authorities in general increased over time as the amount of case reporting—and thus the number of older precedents—proliferated. But the rate of change was not uniform. For two decades in the middle of the nineteenth century, the average age of authorities flattened at around 25 years old, indicating that recent case law was cited about the

same amount as early case law establishing first principles. The relative age of authorities also varied significantly by body of law, as the following figure shows.

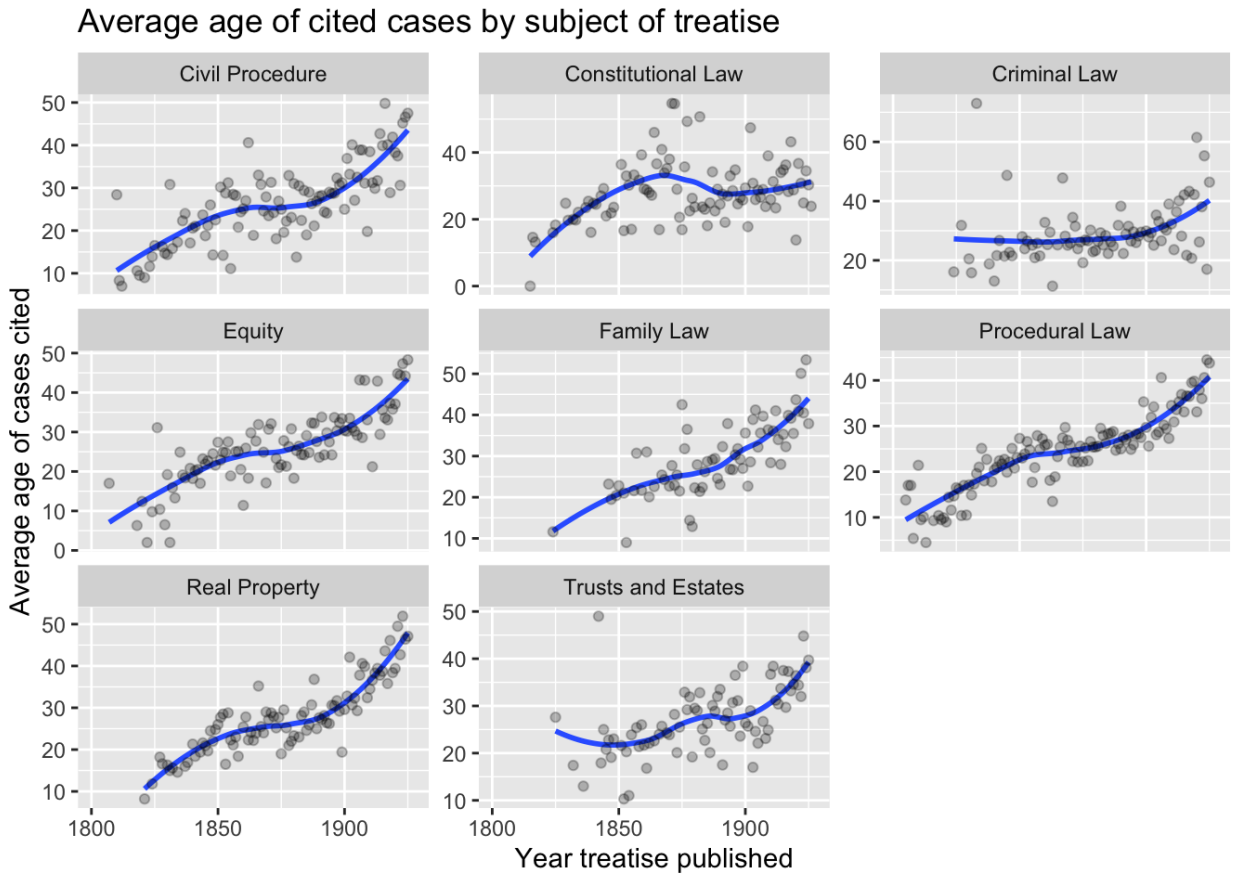


Figure 6. The aging of cited authorities was not uniform across time, but neither was it uniform across types of law. A steeper positive slope indicates a relative lack of development in the law, while a more horizontal line indicates a relatively rapid development of the law. In constitutional law, for example, the average age of cited authorities actually decreased in the 1880s and 1890s, as treatise writers grappled with the changes brought about by the Civil War and the Reconstruction Amendments.

The fields of civil procedure and equity show the most similarity to the general curve. Both bodies of law were significantly impacted by a novel code of remedies issued in New York in 1848 and swiftly copied around the nation thereafter.[40] Like the general curve, the more procedurally oriented fields seemingly stopped aging in the mid-nineteenth century as treatises explaining the new codes relied on recent interpretations of their novel provisions. Constitutional law is the only field for which authorities grew *younger* for a time, starting in the 1860s with the ratification of the Reconstruction Amendments, a radical disjuncture in the law not unlike the procedure codes of a decade before.

[40] See Kellen Funk & Lincoln A. Mullen. "The Spine of American Law: Digital Text Analysis and U.S. Legal Practice." 123 *American Historical Review* 132 (2018).

To an extent, all fields show the effects of increasingly novel law at mid-century, except for criminal law—a body of law unaffected by the procedure codes and largely unaffected by the Reconstruction Amendments and developments in constitutional law.

Significantly, the mechanism of citation to cases in legal treatises appears demonstrably different from the mechanism of citation from one case to another in this regard. The average age of a case cited in a treatise—by the end of the nineteenth century, at least—could be fairly old. But when one case cited another, the typical age of the case cited was quite young, a matter of only a few years. Legal treatises built up their reasoning from first principles by citing the oldest and most influential cases, while cases cited relatively recent cases to identify—and finesse—rules at the legal frontier.

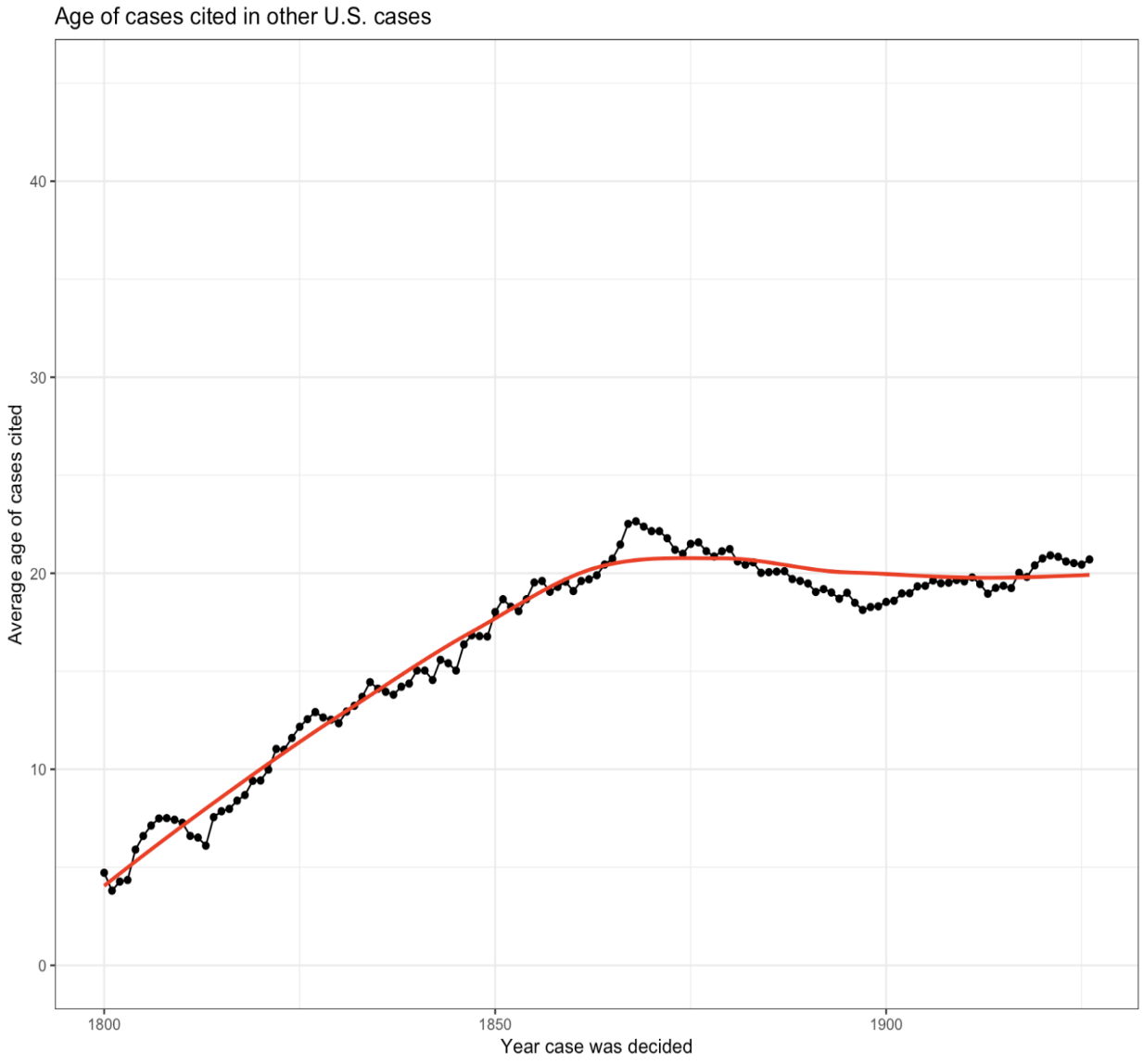


Figure 7. The typical age for a case that was cited by another case was comparatively more recent than the age of a case cited in a treatise.

B. The bibliographic coupling of treatises

The next avenue we are pursuing is understanding the relationship between the treatises themselves. Over time, treatises tended to become less general and more specific. Blackstone could claim to cover the entirety of the common law in four volumes, but by the end of the century, it would take five volumes at least to cover *Fire Insurance Cases*.[41]

How can we demonstrate how treatises were connected to one another? Certainly the treatises are labeled by Gale according to subject matter, such as “Procedural Law” or “Constitutional Law.” Even leaving aside obvious errors in categorization,[42] such categorizations are not to be accepted a priori or ahistorically. Rather, what we are seeking to demonstrate is how the categories of modern legal thought came to develop, including any transitional periods.

There are a number of ways that we can show how treatises were related to one another. The most straightforward is through the concept of bibliographic coupling.[43] The dataset we have created is, in essence, a directed graph showing that Treatise A cited Case X. Two treatises can be considered to be linked in an undirected graph if they both cite the same case. For example, if Treatise A cites Case X, and Treatise B cites Case X, then Treatises A and B can be considered linked with a weight of 1. If they were to cite more of the same cases, then the weight of their connection would increase.

The concept of bibliographic coupling makes intuitive sense for the historical problem we are trying to understand. Writers of legal treatises understood themselves to be crafting an understanding of the law using the findings of cases as their evidence. A case might be regarded as determining some point about contracts, or torts, or the law of railroads. Two authors of competing treatises on contracts would cite a similar, if not necessarily identical, body of case law. An author of a treatise on railroads, however, might cite some of the same cases that the author of a treatise on torts cited. (The railroads, after all, caused many of the injuries or wrongs that led to the development of the law of torts.[44]) The two treatises on contracts would thus be strongly linked together in a bibliographic coupling graph, since they share most of the same cases, while the treatise on railroads and the treatise on torts would be only moderately linked, since the overlap in cases cited was smaller but nevertheless significant.

---

[41] Edmund H. Bennett, *Fire Insurance Cases*, 5 vols (Hurd and Houghton, 1872–1877).
[42] For example, in the MOML corpus, *The Works of Alexander Hamilton* are categorized as being about “Real Property.” Even more bizarrely, a Spanish translation of the works of John C. Calhoun is categorized as “Constitutional Law.”
[43] Mark E. J. Newman, *Networks: An Introduction* (Oxford University Press, 2010), section 6.4.1. Bibliographic coupling is a concept related to, but distinct from, the more commonly used concept of co-citation. In co-citation, two documents A and B are considered linked if they are *both cited by* a document X. In bibliographic coupling, two documents A and B are considered linked if they *both cite* a document X.
[44] See John Fabian Witt, *The Accidental Republic: Cripple Workingmen, Destitute Widows, and the Remaking of American Law* (2006).

The network that results can be seen in the figure below. This figure is necessarily simplified for the purposes of visualization, being limited to treatises published before 1860, and showing only treatises which are so strongly connected to other treatises that they share citations to at least 300 of the same cases.

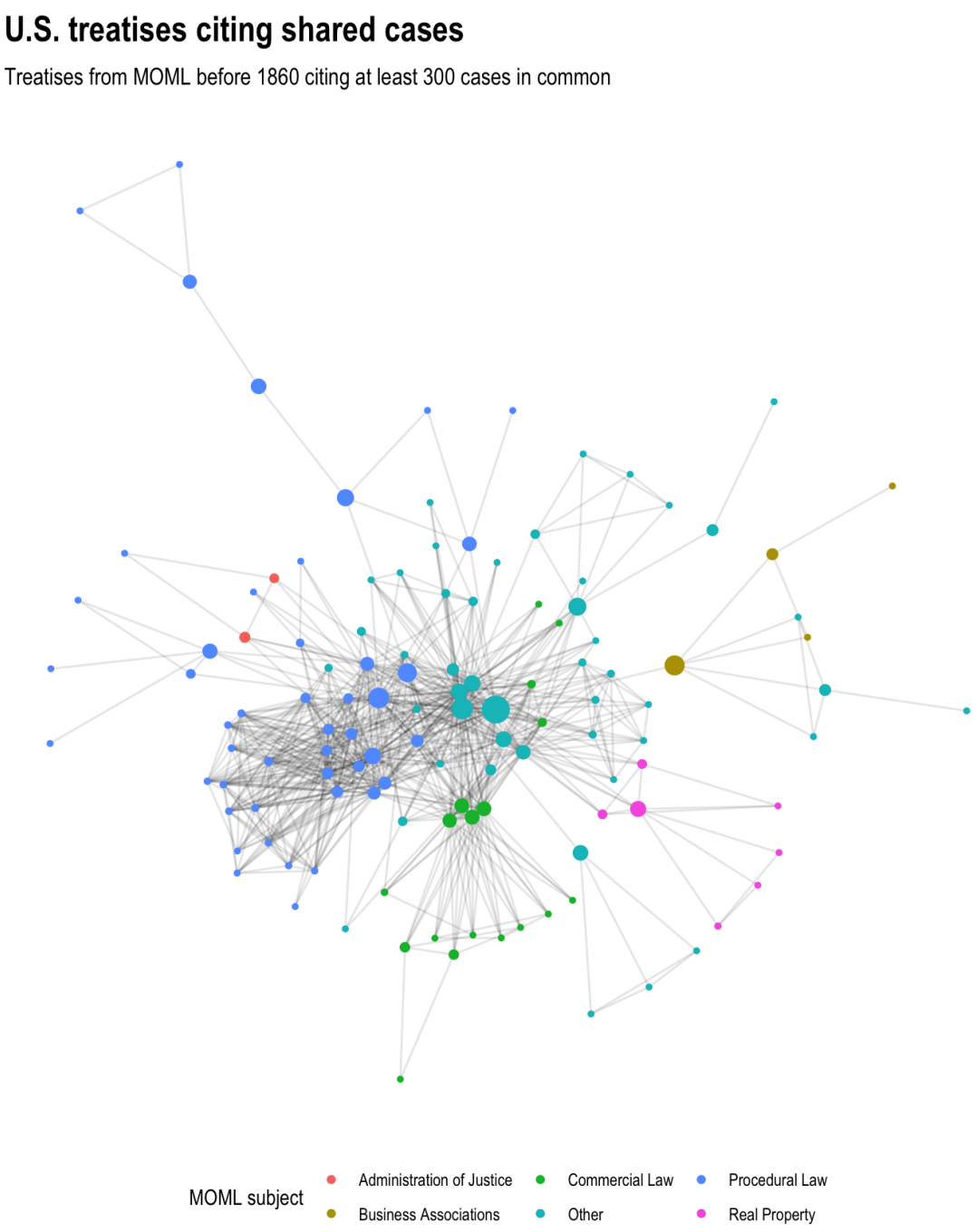


Figure 8. This network shows treatises published before 1860. The treatises are connected to one another if they cited at least 300 of the same cases. This network shows only the single

largest connected subgraph; other subgraphs could also be displayed. The treatises are colored according to their classification in the Making of Modern Law. The apparent structure of the network, which is more detailed than what is captured by the MOML categories, hints at the ways in which the categories of modern legal genres formed.

Though this network is considerably simplified for visualization purposes, it demonstrates an approach to understanding how the categories of legal modernity developed.[45] The network shows a more detailed degree of structure than is captured by the atemporal library card catalog categories assigned by the Making of Modern Law database.

For example, "commercial law" shows distinct subgroupings. The more centrally connected treatises are those titled *A Compendium of Mercantile Law* and *The Elements of Mercantile Law*. The more peripheral treatises labeled as "commercial law" are several treatises (in several different editions) about the law of promissory notes, bills of exchange, and bank notes. In other words, the treatises on mercantile law generally are connected to even more generic areas of law such as contracts and pleadings. But those general treatises on mercantile law are also connected to more specific treatises on an area of concern to merchants. Real property shows a similar variegation between general treatises on the subject and more specific treatises on the law of mortgages.

The network does not only show ways in which the treatises became increasingly variegated over time. It also shows how treatises combined what are now regarded as distinct categories of modern law. For example, the section of the network which contains treatises labeled as "business associations" also shows them to be connected in a subgroup with other treatises that do not share that label. Some of the treatises in that section of the graph are about *The Law of Private Corporations Aggregate* and other such treatises about corporations; but their subtitles indicate them to be not just about *The Law of Agency* but specifically about agency *As a Branch of Commercial and Maritime Jurisprudence*. In other words, modern categories of corporate, commercial, and maritime law were not held distinct by antebellum treatise writers, but rather were treated together. In the act of citing together cases on corporate, commercial, and maritime law, the treatise authors were creating a new category of law. But that category of law

[45] There is a robust literature on citational networks within (usually much smaller corpora of) judicial case law. Our treatment of bibliographic coupling here is suggestive of an approach unique to the treatise literature, but we expect to draw heavily upon the judicial network models as the project unfolds. See especially Stephen M. Marx, "Citation Networks in the Law," 10 *Jurimetrics Journal* 121 (1970); Gregory Caldeira, "Legal Precedent: Structures of Communication between State Supreme Courts," 10 *Social Networks* 29 (1988); James H. Fowler, Timothy R. Johnson, James F. Spriggs II, Sangick Jeon & Paul J. Wahlbeck, "Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court," 15 *Political Analysis* 324 (2007); Martin Gelter and Mathias Siems, "Networks, Dialogue, or One-Way Traffic? An Empirical Analysis of Cross-Citations Between Ten of Europe's Highest Courts," 8 *Utrecht Law Review* 88 (2012); Ryan Whalen, "Legal Networks: The Promises and Challenges of Legal Network Analysis," 2016 *Michigan State Law Review* 539 (2016); Mattias Derlen & Johan Lindholm, "Peek-a-Boo, It's a Case Law System: Comparing the European Court of Justice in Three Dimensions," 16 *German Law Journal* 16 (2015); Jens Frankenreiter, "Network Analysis and the Use of Precedent in the Case Law of the CJEU—A Reply to Derlen and Lindholm," 18 *German Law Journal* 687 (2017); Wolfgang Alschner & Damien Charlotin, "The Growing Complexity of the International Court of Justice's Self-Citation Network," 29 *European Journal of International Law* 83 (2018).

was, crucially, neither the same categories that the new republic inherited from English common law, nor the fully variegated categories of modern law. They were, instead, a transitional period of development as law was made modern by the treatise authors.

C. Canon formation

The network above provides a simplified visualization of a massively large and complex set of interactions across a fairly broad corpus. Further down, we will focus on particular cases—tiny nodes within the vast sea of data. In prior work, we have extolled the virtues of "medium data," and it is especially at the middle scale where our data may be most useful to domain specialists, as we illustrate in this section. Rather than ranging over the whole corpus of treatises or cases, we can hone in on particular domains of knowledge as they were undergoing significant formation and construction to study patterns of development.

Our network graph above risks anachronism by using a twenty-first-century library card catalog to classify nineteenth-century texts. This time around, we can draw on a classification from our period to define a set of relevant texts: namely, required coursebooks for law school classes. While some of America's leading law schools trace their roots to the early nineteenth century, most law departments remained trivially small or disorganized until the 1870s. But starting in that decade, law schools hired up to a dozen expert lecturers and created defined course schedules to guide students through two or three years of instruction. Along with the schedules came an increasingly standardized list of required texts. Some fields quickly coalesced around a standard textbook—"Bishop on Contracts," for instance, was the nearly universal requirement in that field for almost fifty years. Other domains show more variation and lack a clear dominant text. Salient in the latter category is the field of Civil Procedure, taught as a foundational course at every law school, but with a diverse array of textbooks.

The following table of Civil Procedure curricula is drawn from law school bulletins found in the law library archives of their respective schools. While all schools required instruction in common law pleading, most over time included additional courses on code procedure and special procedures in equity.

School	Textbook	Ed.	From	Until	Topic
Yale	Gould on Pleading	3rd	1870	1890	com. law
Yale	Stephen on Pleading	3rd	1874	1880	com. law
Yale	Heard on Civil Pleading	1st	1881	1913	com. law
Yale	Bryant on Code Pleading	1st	1896	1914	code

Yale	Ames's Cases	Rev.	1912	1919	com. law
Yale	Whittier's Cases on Common Law Pleading	1st	1914	1919	com. law
Yale	Hinton's Cases on Code Pleading	1st	1915	1919	code
Yale	Cook & Hinton's Cases on Common Law Pleading	pre	1917	1925	com. law
Yale	Thompson's Cases on Equity Pleading	1st	1916	1917	equity
Yale	Redfield's Cases on Code Pleading	1st	1917	1919	code
Yale	Sunderland's Cases on Code Pleading	1st	1920	1925	code
Yale	Morgan & Whittier's Cases on Com. Law Pleading	1st	1920	1925	com. law
Columbia	Ames on Pleading	1st	1895	1925	com. law
Columbia	Stephen on Pleading	3rd	1895	1925	com. law
Columbia	Perry on Pleading	1st	1901	1925	com. law
Columbia	Redfield's Cases on Code Pleading	1st	1903	1925	code
Columbia	Thompson's Cases on Equity Pleading	1st	1903	1925	equity
Columbia	Webb's Cases under the Code	1st	1916	1925	code
Columbia	Cook & Hinton's Cases on Common Law Pleading	1st	1919	1925	com. law
Harvard	Stephen on Pleading	7th	1874	1875	com. law
Harvard	Langdell on Equity	1st	1875	1901	equity
Harvard	Ames on Pleading	1st	1876	1904	com. law

Harvard	Ames on Equity	1st	1902	1925	equity
Harvard	Ames on Pleading	2nd	1905	1925	com. law
Michigan	Stephen on Pleading	3rd	1890	1900	com. law
Michigan	Bliss on Code Pleading		1880	1900	code
Michigan	Lube on Equity Pleading		1882	1900	equity
Michigan	Perry on Pleading	1st	1900	1905	com. law
Michigan	Thompson's Cases on Equity Pleading	1st	1903	1915	equity
Michigan	Pomeroy on Code Pleading	4th	1905	1910	code
Michigan	Ames on Pleading	2nd	1910	1915	com. law
Michigan	Sunderland's Cases on Common Law Pleading	1st	1915	1925	com. law
Michigan	Sunderland's Cases on Code Pleading	1st	1915	1925	code
Georgetown	Stephen on Pleading	3rd	1883	1895	com. law
Georgetown	Mitford's and Tyler's Equity Practice and Pleading	1st	1883	1886	equity
Georgetown	Heard's Equity Pleading	2nd	1887	1888	equity
Georgetown	Barton's Suit in Equity	Rev.	1889	1895	equity
Georgetown	Ames on Pleading	1st	1894	1895	com. law

Figure 9. Assigned coursebooks in courses titled “Pleading” or “Procedure” at the law schools of Harvard, Yale, Columbia, Michigan, and Georgetown from ca. 1878 to 1925.

There is some overlap among the five law departments—Stephen’s third edition on common law pleading was fairly popular in the 1870s, and the works of James Barr Ames found use outside

of Harvard in the twentieth century. But more striking is the variety. Especially when it came to instruction in equity, or on the relatively new codes of procedure, each school had its own favored texts.

But does a variety of casebooks indicate variety in the curriculum? Or to put the question another way, did the casebooks focus on the same cases and differ only in presentation and supporting materials? To explore the question, we can turn to a network visualization that preserves more of the complexity of the data than our simplified network above. Each casebook in our lists excerpts and cites cases (and the excerpts themselves contain further citations to case authorities). Altogether, 15,000 cases are cited across the curriculum, a surprisingly high number for such a narrowly circumscribed corpus. The following graph visualizes each one of those cases as a point, connected to the schools in whose curriculum it appears.

The Emerging Canon of Procedural Law in the U.S., 1875–1925

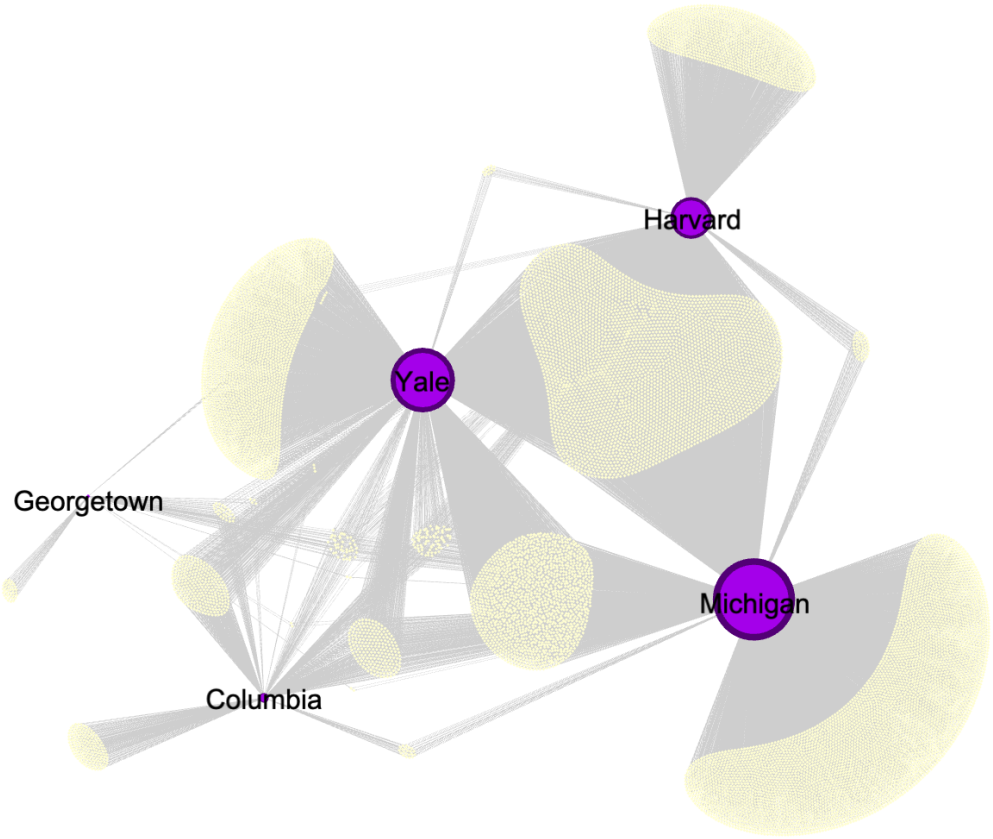


Figure 10. This network graph depicts each of 15,000 cases in procedural law as a (small yellow) node connected to (purple, larger) nodes of schools that included those cases in their curriculum. Clusters at the edges, in the lower right corner for instance, indicate a body of cases taught only at the one school to which they are connected. Clusters in the center indicate cases common to multiple curricula. For instance, the largest cluster of cases in between Yale,

Harvard, and Michigan indicates the set of courses taught in all three schools, while the cluster to its lower left are courses taught at Yale and Michigan, but not at Harvard.

The graph indicates that each school taught procedure, in part, by using a body of cases that were unique to it, i.e., that were not shared by the other schools. Indeed, Michigan's instruction covered a much larger array of cases unique to the midwestern school's curriculum than common to those of other schools. Nevertheless, significant clusters of cases were shared among curricula, albeit with disuniform coverage. Harvard, Yale, and Michigan appear to have coalesced around a common core of procedural cases that Columbia and Georgetown did not share, though each had notable common curricula with Yale and with each other.

Drawing on the metadata linked to each case in CAP, we can chart other patterns, such as the regional contributions to each field. Where did the 15,000 cases cited in the curriculum come from? The following figure provides some of the details.

	Common Law	Code	Equity	Overall
New York	15.4%	26.9%	13.3%	19.9%
Massachusetts	7.6%	1.5%	8.1%	5.2%
Indiana	8.6%	9.9%	3.5%	8.8%
California	5.4%	9.2%	2.3%	6.7%
U.S. Federal	4.4%	2.5%	16.3%	4.3%
Missouri	3.6%	8.0%	3.4%	5.4%
Alabama	3.6%	0.7%	7.7%	2.6%
Illinois	3.4%	0.8%	4.2%	2.4%
New Jersey	0.1%	0.3%	3.7%	0.8%
Wisconsin	3.3%	8.9%	2.8%	5.5%
Iowa	2.7%	4.4%	2.9%	3.4%

Figure 11. U.S. jurisdictions by share of citations in the Civil Procedure curriculum of Harvard, Yale, Columbia, Michigan, and Georgetown law schools.

As we might expect, New York dominates all categories, and there are a few other unsurprising patterns: States that retained a separate court of chancery (New Jersey, Iowa, Alabama) figure prominently in equity procedure; states that rejected codification (Massachusetts, Illinois) contribute basically nothing to education in code pleading. But some unexpected patterns also stand out. Despite nearly a third of the corpus being made up of textbooks taught at Michigan

Law School, Michigan cases were cited too few times to even register. Instead, Indiana appears to be far and away the preeminent midwestern jurisdiction. The U.S. federal courts enjoyed a respectable but not dramatic influence, except when it came to equity, where their preeminence outstripped even New York's. Finally, California shows that even a state that had begun its existence at midcentury could achieve significant juridical influence by the century's end.

D. Significant cases

The approaches we have shown above about the age of cited authorities and a bibliographic coupling necessarily obscure individual cases and even individual treatises by showing aggregate patterns. But the dataset we have assembled also allows us to descend back to the level of a specific case, or set of cases. While we will continue to refine our approach to citation detection, including by whitelisting more genuine cites distorted by textual variants or messy OCR, we can preliminarily compile a list of the top-cited cases in the treatises. Such a list looks much like one would expect: dominated by Supreme Court rulings which were binding in all places and bodies of law, roughly in the same order of importance in which they are regarded in the constitutional canon today: *Dartmouth College v. Woodward*, *Osborn v. Bank of the United States*, *McCulloch v. Maryland*, *Munn v. Illinois*, *Fletcher v. Peck*, *Gibbons v. Ogden*, *Marbury v. Madison*, *Swift v. Tyson*, and so forth. Setting aside federal decisions one would necessarily expect to have widespread coverage, we derive the following list of top-ten cited cases by the number of treatise volumes that cite them.

Case	Jurisdiction	Date	Cites
Kellogg v. Ingersoll	Mass.	1804	1650
Curtis, Graham & Blatchford v. Leavitt	N.Y.	1857	444
Gold v. Eddy	Mass.	1804	440
Fairlee v. People	Ill.	1849	396
Salem Bank v. Gloucester Bank	Mass.	1820	392
New York & New Haven Railroad v. Schuyler	N.Y.	1865	386
People v. Mather	N.Y.	1830	384
Farwell v. Boston & Worcester Rail Road Corp.	Mass.	1842	373
Stackpole v. Arnold	Mass.	1814	369
Despatch Line of Packets v. Bellamy Man. Co.	N.H.	1841	332

Figure 12. The ten most-cited state-level case decisions in MOML treatises, by number of citing volumes.

Even without controlling for age, the list is quite varied in terms of early and late decisional law, private commercial law or criminal law, but not by jurisdiction. Massachuesetts and New York enjoyed outsized influence, as many legal histories of those jurisdictions have attested.

Of course the mere fact that a case was frequently cited does not adequately reveal its place in the development of law. A next step is to see the changing rate at which legal treatises cited the cases over time. As the figure below shows, the trends in the rates at which treatises cited these cases were not uniform. Most were gradually recognized as significant by treatise writers shortly after their date of decision, and gradually made their way into the treatise literature. But other cases, such as *Fairlee v. People*, were not recognized as significant by most treatise writers for several decades past their date of decision, after which they rapidly entered the treatises.[46] Most cases declined in significance to the treatises over time.

[46] *Fairlee v. People*, 11 Ill. 1 (1849).

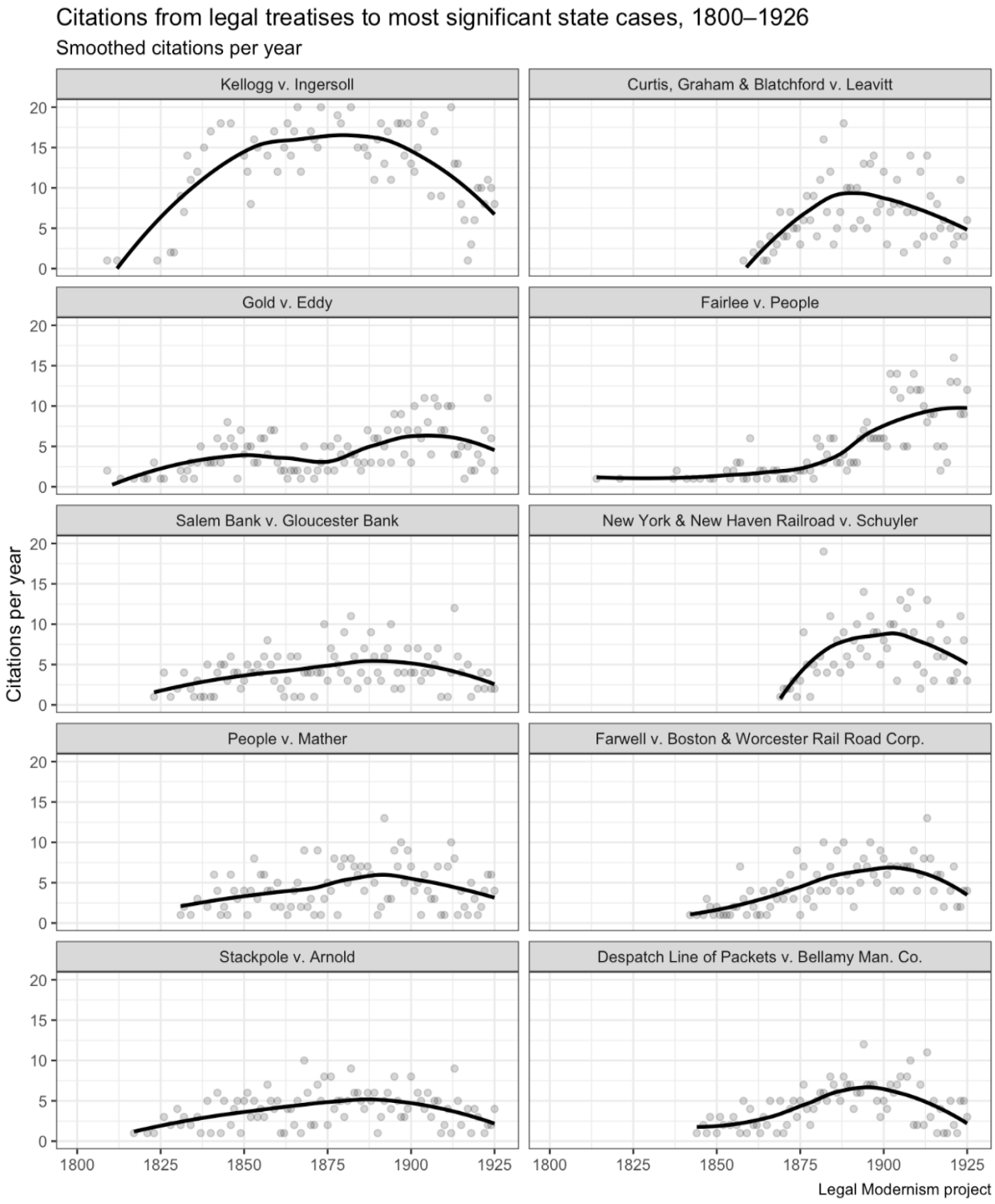


Figure 13. The state cases that were most cited by legal treatises show different trends in the rate of citation. The significance of most was recognized gradually by treatises but soon after their decision date. But some—such as Fairlee v. People—were not cited with any frequency until decades after their decision date.

IV. Conclusion

The approaches illustrated here can help us begin to understand the work of law in an age of legal and technological modernization. Relative age of authorities can detect patterns of rapid innovation and stagnation. Bibliographic coupling can show genre formations evolving over time before the calcifying process of library cataloging set in. Curricular networks can investigate the deliberate formation of canons by professors in emerging fields. And calibrating a network around a single case citation can show how one particular authority was transmitted across time and space but also how that authority itself was transformed to bear on different bodies of law.

This is only the beginning. Legal historians have long relied on treatise tables of contents to support their arguments about the structural evolution of the law.[47] Others have invested significant resources in closely reading through treatises to detect, for instance, how cases involving women and children developed rules about the duties of care owed by "the reasonable man."[48] We expect our project to extend such inquiries at much higher scales and with greater degrees of sophistication. In time we expect topic-modeling, stylometry, and natural language processing techniques may allow for other inquiries into legal modernization not yet addressed here. We intend to build out our database of citations by using our detection method on the CAP corpus, and likely on a corpus of the English Reports. Metadata from CAP may allow for targeted inquiries concerning the identities of particular judges, levels of courts, or the arguments of legal counsel.

At this early stage, we certainly welcome ideas and suggestions and hope readers and commenters enjoy following along as our project develops.[49]

[47] For particularly salient examples, see Hartog, *Public Property and Private Power*; Janet Halley, "What Is Family Law."
[48] Welke, *Recasting American Liberty*, 103.
[49] To follow our work on this historical problem, please see the project website. *Legal Modernism*, Roy Rosenzweig Center for History and New Media and Columbia Law School (2022–): https://doi.org/10.31835/legalmodernism.