



Continuity and change in US legal tradition: Evidence from judicial citation communities

Elliott Ash^{a,1,2} , Caterina Chiopris^{b,1} , Robert Mahari^{c,1}, and Suresh Naidu^{d,1}

Affiliations are included on p. 11.

Edited by Thomas D. Albright, The Salk Institute, La Jolla, CA; received July 9, 2025; accepted April 9, 2026

Legal doctrine is often described as evolving case by case while preserving coherent legal traditions. We introduce a method for identifying and tracing such traditions by applying the Louvain community-detection algorithm to the complete U.S. federal judicial citation graph (1790–2024; 760,000 opinions linked by 8.3 million citations). The resulting citation communities are unsupervised, blind to court labels and topical taxonomies, yet they form meaningful empirical groupings: In validation exercises, they recover known institutional structure and capture coherent substantive and procedural approaches to law. Using these communities as empirical units, we study how legal traditions evolve and respond to legislative change. We show that the introduction of pivotal federal statutes shifts citation patterns toward newer precedent, indicating that legislation can disrupt established traditions and redirect the development of case law.

common-law evolution | citation networks | Louvain communities | legal analytics | US federal law

For more than two centuries scholars have treated the common-law tradition as a living organism. The common law grows case by case, yet somehow maintains overall coherence e.g., refs. 1–3. A central puzzle follows from this metaphor: Federal courts, though arranged in a clear hierarchy, are in practice an atomistic collection of judges. No docket can be centrally planned, nor any individual judge micromanaged. How can so many independent actors generate a body of doctrine that reads as a single, intelligible system of legal traditions? How does that system of traditions absorb statutory changes emanating from democratic politics?

We approach the question with network analysis (4, 5). In a common-law system, the principle of *stare decisis*, or precedent, means that citations to previous cases are important inputs into a successful legal argument. Using the complete graph of citations among all published U.S. federal opinions—millions of edges linking two centuries of decisions—we ask what the pattern of citations alone reveals about the structure and evolution of American jurisprudence. An unsupervised Louvain algorithm (6) partitions the graph into citation communities that we interpret as empirical legal traditions, without any knowledge of jurisdictional labels, levels of authority, or doctrinal headings. The resulting map recovers formal institutional boundaries, for example between circuit courts. The communities also expose forms of legal reasoning, including both statute-anchored domains and case-heavy reasoning modes, that are more granular than traditional legal taxonomies.

Tracing these citation communities over time reveals several regularities about the evolution of legal traditions. Innovation is constant: Each decade generates many small clusters, most of which disappear within twenty years as short-lived “orphans.” At the same time, major fields, such as administrative law, civil procedure, and intellectual property, exhibit enduring lines of precedent that persist for more than a century, with temporary branches that coexist for a period before merging or fading. Doctrinal vintage also varies substantially: Nineteenth-century precedents on constitutional structure and federalism remain deeply integrated with modern case law, whereas many other traditions emerge only after the regulatory expansion of the 1960s–1970s.

We then analyze legislation as a mechanism of persistence or disruption. Major statutory enactments tend to redirect citation behavior within communities, shifting attention toward statutory authority and toward newer precedent.

These findings portray U.S. legal doctrine as a self-organizing network that preserves coherence through persistent traditions while continually experimenting at the margins and adapting to statutory innovations. Given the importance of institutional origins, they also push against a “formalist” null hypothesis that novel and persuasive legal arguments emerge at random and gain traction through sheer intellectual force (7). While our

Significance

Using 760,000 U.S. federal court opinions linked by 8.3 million citations, we identify legal traditions as citation communities recovered from the judicial citation network using the Louvain algorithm. These communities form meaningful empirical units: They reproduce known judicial structure and legal domains. This network perspective reveals how bodies of precedent evolve over time. We show that the introduction of key federal statutes increases reliance on statutory authority and shifts citation patterns toward newer precedent within affected communities. These results demonstrate that legislative change can redirect the trajectory of legal traditions and illustrate how network methods can be used to study the evolution and disruption of complex knowledge systems.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2026 the Author(s). Published by PNAS. This open access article is distributed under [Creative Commons Attribution License 4.0 \(CC BY\)](#).

PNAS policy is to publish maps as provided by the authors.

¹ E.A., C.C., R.M., and S.N. contributed equally to this work.

² To whom correspondence may be addressed. Email: ashe@ethz.ch.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2509763123/-DCSupplemental>.

Published July 20, 2026.

analysis focuses on federal court decisions rather than on the traditional domain of judge-made common law, the underlying mechanism is the same: Judicial reasoning evolves through precedent and interpretation. In this broader sense, the “common law network” we describe captures the dynamic structure of precedent-based reasoning that spans both common-law and statutory domains.

Our analysis suggests a balance in the common law between coherence and flexibility (8). Persistent doctrines act as long-term anchors of tradition, allowing the system to incorporate experimental decisions without jeopardizing its core architecture (4). Far from being a rigid pyramid or an anarchic swarm, federal case law emerges as a network that continually refines itself through the ordinary act of citation. But the internal logic of citation must still adapt to the exigencies of statute, as the introduction of important legislation shifts citation patterns into a new community-specific body of precedent.

Related Literature

A broad theoretical tradition views law as a complex, adaptive system that must remain stable enough to guide private ordering yet agile enough to accommodate social and technological change. As such, forms of legal reasoning are driven not by internal arguments or purely deductive chains from precedents and facts, but instead respond to broad-based changes in statutes and society. The opening of *The Common Law* (1) poses the hypothesis clearly:

The life of the law has not been logic: It has been experience. The felt necessities of the time, the prevalent moral and political theories, intuitions of public policy, avowed or unconscious, even the prejudices which judges share with their fellow-men, have had a good deal more to do than the syllogism in determining the rules by which men should be governed.

This perspective situates legal evolution within its social and institutional context, emphasizing that doctrine adapts as courts reinterpret precedent to meet the practical demands of their time.

Law-and-society scholars have stressed the reciprocal influence of doctrine and culture (3). Classic liberal theorists cast common law as spontaneous order in which rules emerge from decentralized problem solving rather than central design (2). Law-and-economics work formalizes this intuition: Inefficient rules invite costly relitigation and are gradually displaced, nudging doctrine toward efficiency (9–12).

In the common law, rules evolve by establishment of and citation to precedent. Precedent has been modeled as depreciating legal capital whose citation value declines over time (13). Judges have been shown to gradually and coherently extend precedent by making finer distinctions based on the case facts (14). We study U.S. federal courts; since *Erie Railroad Co. v. Tompkins*, 304 U.S. 64 (1938), federal judges do not create substantive law for traditional common-law areas like contracts, property, and tort. Yet federal judges develop and refine doctrine by citing, distinguishing, and extending prior decisions, so the evolutionary dynamics of case law remain fundamentally those of the common-law tradition. The process of development and refinement of precedent is a mechanism by which laws can efficiently adapt to changing societal circumstances.

These ideas have motivated empirical work on how legal structures reflect social structures. Statutory and case-law corpora scale with social complexity much like other knowledge systems

(8, 15, 16). More law is associated with more economic growth (17) and in some domains better governance (18, 19).

At the microlevel of the individual judicial decision, empirical work has explored how citations transmit authority, reasoning, and ideology. Judges signal precedent quality through costly references (20), engage in strategic citation to bolster preferred outcomes (21), and pass along distinctive doctrinal “memes” (22). Early-career citation habits have been shown to forecast later voting behavior (23).

Mapping these links at larger scale uncovers the geometry of judicial knowledge. The citation network is highly skewed, with a power-law degree distribution anchored by a small core of landmark decisions (4, 24, 25). Network-based measures that capture second-degree influences show this hierarchy better than plain citation counts (5, 26). Community detection algorithms can reveal latent structure in the citation network, allowing analysis to partition decisions into cohesive groups (27–29). Hybrid methods adding text features have improved the thematic clarity of clusters (22, 30, 31).

In turn, it has been shown that the macrostructure of the citation network influences local citation links. The number of backward links predicts a case’s eventual prominence (32), while network centrality slows the otherwise rapid depreciation documented for Supreme Court holdings (33). Opinions combining very old with very recent sources gain disproportionate future influence (34). Bursts of expansion occur when landmark rulings redirect influence across layers of law (8).

A closely related study also analyzes the citation network of U.S. federal court opinions (35), focusing on a disruption index—measuring how far individual decisions break from established citation paths. That work documents a long-term decline in such pathway “breaks,” attributing the trend to precedent overload and institutional polarization. Our work complements this approach by shifting the focus from individual-case disruptiveness to the community-level persistence of doctrinal lineages. Our empirical grounding lets us use communities to test mechanisms of persistence and statutory disruption, not just to describe clusters.

Legislation itself can be a deliberate mechanism for disrupting judicial precedent. Congress has been shown to override Supreme Court statutory interpretations and use those overrides to redirect doctrine (36). The “statutory override game” formalizes how legislatures explicitly undo judicial interpretations, leaving shadow precedents in their wake (37). More recently, proposed statutory limits on judicial power have been found to alter judicial behavior (38). We speak to this literature below by providing network-scale evidence that major statutory introductions measurably reshape citation patterns within doctrinal communities.

This work is part of the broader recent literature on the determinants and consequences of idea diffusion using historical data (39). Horizontal diffusion among state courts follows lines of professional prestige and institutional similarity (25, 40, 41). This line of research is poised for significant progress given the emergence of large-scale public resources on legal citation networks and associated metadata (42).

Methods

Data. This work is based on a large-scale analysis of the U.S. federal court citation network, constructed by integrating three datasets. The primary dataset is the Case Law Access Project (CAP), which provides the network’s underlying structure. This is supplemented with two metadata sources on judicial characteristics and the textual context for citations.

The case law access project (CAP) citation graph. The main analysis is built upon the citation graph derived from the Case Law Access Project (CAP), a comprehensive digital collection of over 6.7 million U.S. case law decisions. We utilize the subset of these data corresponding to the U.S. federal courts, which includes decisions from the Supreme Court, the Courts of Appeals, and District Courts. In this dataset, each judicial opinion is a node and a citation from one opinion to another is a directed edge. CAP constructed this network by programmatically parsing the text of each opinion to extract all outgoing citations. For our study, we use the complete dataset of federal opinions published between 1790 and 2024 and included in the database. This results in a network of 760,000 cases and 8.3 million citations, which forms the foundation of our analysis.

Judicial idiosyncrasies metadata. To incorporate information about judicial actors into our analysis, we leverage data from a recent study on the early-career citation patterns of federal judges (23). The authors of that study created a dataset of 283,511 civil lawsuits from U.S. District Courts. They systematically disambiguated judge identities by matching the raw judge names listed in CAP opinions against the biographical data provided by the Federal Judicial Center (FJC). This matching process provides a unique judge ID for each civil district court opinion in their dataset. This allows us to link opinions to specific judges and their associated biographical details, such as their gender, political party of their appointing president, and appointment and termination dates.

The legal passage retrieval dataset (LePaRD). To understand the substantive legal reasoning behind the citation links in our network, we incorporate the Legal Passage Retrieval Dataset (LePaRD) (42). LePaRD is a large-scale corpus containing over 4 million examples of passages from U.S. federal court opinions that are directly quoted in subsequent opinions. For a given citation, the dataset provides the specific text from the cited opinion (the source) that is referenced in the citing opinion (the destination). This allows us to move beyond a purely structural analysis of the citation graph and investigate the flow of legal ideas and arguments. By associating these textual data with the corresponding edges in our network, we add a crucial qualitative dimension to the citation links.

Parsing statutory, regulatory, and constitutional citations. Because statutes and regulations may redirect precedent, we separately parse noncase citations. Beyond case-to-case citations, federal opinions cite statutes (U.S. Code), regulations (Code of Federal Regulations), and constitutional provisions. We extract these noncase citations using a regex parser, which identifies citation type and returns structured tokens (title, section, amendment, article, etc.). Citations to Public Laws and the Statutes at Large are mapped to their corresponding U.S. Code sections using official concordance tables. Citations to the Federal Register and Executive Orders are pooled into the regulatory (CFR) category. The resulting corpus contains 5,445,978 statutory (USC) citations, 686,438 regulatory (CFR) citations, and 88,747 constitutional citations. Finally, for USC statutes, we parsed the enactment dates of each section using the XML files from the Office of the Law Revision Counsel, assigning the earliest listed year. This decomposition lets us track community-level citation mixes and enables the statutory-disruption tests below.

Community Detection. This subsection describes our main approach to unsupervised community detection on the legal citation network. We outline the Louvain algorithm and how it is implemented, and then describe how we interpret the resulting communities.

With the complete directed graph of judicial citations in hand, our goal is to identify coherent groups of cases that cite one another densely but cite outside cases much more sparsely. Such groups, if they exist, correspond to recognizable legal doctrines or traditions. Formally, the task is to find a partition of the citation network that maximizes within-group connectivity relative to between-group connectivity.

We adopt the Louvain algorithm (6), a greedy multilevel procedure for modularity optimization that is both conceptually transparent and computationally efficient at the scale of millions of edges. Modularity, Q , measures how many links fall inside communities compared with what would be expected in a random network with the same degree sequence:

$$Q = \frac{1}{2m} \sum_{i,j} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta(c_i, c_j),$$

where A_{ij} is the (possibly weighted) adjacency matrix, k_i and k_j are the degrees of nodes i and j , $m = \frac{1}{2} \sum_i k_i$ is the total number of edges, c_i is the community label of node i , and δ is the Kronecker delta. A partition with high Q contains many more intracommunity edges than would occur by chance. Note that, because modularity is defined for undirected graphs in the standard Louvain implementation, we treat the citation network as undirected for the purpose of community detection, collapsing each directed citation into an undirected edge between the corresponding opinions.

The Louvain procedure repeats two phases until convergence:

1. **Local moving phase.** Each case starts as its own singleton community. For every node, the algorithm considers each citation neighbor's community as a potential destination. A move is executed if and only if it yields a positive gain in global modularity; otherwise the node remains in its current community.
2. **Aggregation phase.** Once no further local move can increase Q , each community is collapsed into a "supernode," producing a smaller coarse-grained graph whose edge weights equal the total citations between the underlying communities. The local-moving phase is then reapplied to this reduced graph.

The process terminates when an entire sweep yields no modularity gain, delivering an unsupervised partition of the citation graph into legal communities.

Three features of this strategy are critical for the present study. First, its computational efficiency allows us to analyze the full federal corpus without subsampling. Second, Louvain requires no advance choice of the number or minimum size of communities, letting the data themselves reveal doctrinal clusters. Third, by abstaining from external legal taxonomies we can see what doctrinal structure can be detected purely from judges' citation behavior, without imposing any structure of our own.

It is worth noting that although the citation network is built from judicial opinions, it inevitably captures both common-law reasoning and the interpretation of statutes. Statutory enactments generate new clusters of citations as judges construe and apply legislative provisions, and these developments become embedded in the same network structure as judge-made law. Our use of the term "common law network" therefore refers not exclusively to the common law in its narrow sense, but more broadly to the self-organizing structure of legal reasoning expressed through citations among judicial opinions. In this sense, statutory and common-law evolution are treated within a unified empirical

framework: Both produce enduring lines of precedent that reveal how the federal judiciary organizes and updates legal doctrine over time. The analysis below studies this interaction directly.

Validation.

Validation using court-to-court citations. Before getting into case-level doctrinal communities, we conduct a validation exercise that asks whether the Louvain procedure recovers a known judicial structure when one should be present. U.S. judges ordinarily cite precedent from their own circuit more frequently than precedent from other circuits. This is due to own-circuit precedent being binding on judges (you must follow it), while other-circuit precedent being persuasive (you can follow it). Further, this same-circuit preference might strengthen once electronic databases make relevant decisions easier to locate. Consequently, if the algorithm is sensitive to real citation regularities, it should be able to reconstruct the federal circuit map when we aggregate citations at the court level.

We restrict attention to the modern era of computer-assisted legal research by selecting the opinions filed between 1990 and 2018. For each federal district court and calendar year we aggregate every opinion issued that year, yielding a set of court-year nodes. The directed weight on the edge from court-year u to court-year v equals the number of times opinions written by court u in year t cite opinions of court v in year s . The resulting multilayer network thus records the intensity of intercourt citation flows over time.

We run the Louvain procedure on the court-year graph. The algorithm is blind to all metadata: Node labels carry no information about court names, geographic location, or formal circuit membership. Communities are assigned purely on the observed citation frequencies.

Fig. 1 plots every court-year as a point located at the associated federal district court's geographic coordinates. Points belonging to the same court are clustered into circles, with the color of points indicating the community label returned by Louvain. As can be clearly seen in the figure, the algorithm recovers the federal circuits almost perfectly: Nearly all court-years belonging to the same circuit share a common color, while citations that cross circuit lines are too sparse to merge the clusters. This partition emerges without any institutional labels, supporting that our method captures meaningful structure rather than artifacts of graph partitioning.

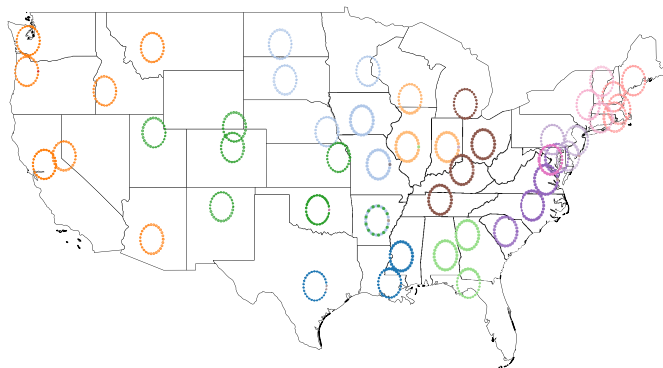


Fig. 1. Louvain applied to citations recovers circuits. Each point represents a federal district court-year (1990–2018), plotted at the court's geographic coordinates. Colors indicate Louvain community assignments, computed solely from citation frequencies with no access to court names, locations, or circuit membership. Nearly all courts within the same federal circuit receive the same community label.

What do citation communities capture? The Louvain method is unsupervised and only returns case IDs, so we interpret communities separately and treat them as empirical units to be tested. Here we use AI as a labeling aid to help interpret the legal content of the communities (without changing any of the communities data used in the analysis).

We prompt the LLM (GPT-4o) to interpret clusters in two ways. As a first exploration, we provide the names of the top ten most-cited cases associated with the cluster to the LLM, and obtain a summary of the overriding topics and legal principles. That provided an initial sense of the themes covered by the cluster. Second, we sampled 100 key passages from opinions in the cluster, which had been quoted by future judges (42). In addition, we identified the federal statute sections most associated with each community (see below), to further ensure consistency.

For both AI approaches, we found the summaries to be intuitive and high-quality. *SI Appendix* provides a list of summaries based on the key quoted passages for the 25 largest communities built from the whole citation graph (used in Fig. 4 below). The resulting communities span various domains relevant to federal law, both in terms of legal approaches and policy priorities. For example, an immigration cluster (#12) is organized around the Illegal Immigration Reform and Immigrant Responsibility Act (8 U.S.C. §1252) and asylum procedures, while an ERISA cluster (#2) centers on plan-administration duties anchored in 29 U.S.C. §1132. Meanwhile, a patent cluster (#9) is structured around claim-construction and obviousness standards, and a securities/antitrust cluster (#7) focuses on preventing anticompetitive behavior and fraud. A Social-Security cluster (#15) revolves around the “substantial-evidence” formula, uniting cases by deference posture across diverse statutes.

Across these examples, communities are defined by distinctive mixes of areas of law, citation types, shared opportunities for legal relief, and aligned reasoning patterns. The interpretable content of these legal groupings provides context for the character of law evolving across time and in response to the treatments analyzed below.

AI disclosure. This study used AI tools in three ways. First, OpenAI GPT-4o was used to interpret and label citation communities: It was provided with i) the names of the ten most-cited cases in each community to generate initial topic summaries, and ii) a sample of 100 key passages quoted from opinions in each community to extract shared legal reasoning patterns. Second, OpenAI GPT-5.2 and Anthropic Claude Opus 4.5 were used to support coding and data analysis. Third, OpenAI GPT-5.2 was used to refine writing and correct grammatical errors.

Evolution of Citation Communities

This section uses our legal Louvain communities to analyze how bodies of precedent evolve over time. We begin with 30-y sliding windows (with a 10-y stride) of cases, and we apply the Louvain procedure to each of the 12 windows. For each window, a large number of communities (up to 1,116) are constructed, but most of them are small and insignificant (many with just 2 cases). We limit to communities with greater than 100 opinions in a window, excluding over 95% of communities but accounting for more than 99% of all citations. In the resulting panel, there are up to 36 active communities at any given period.

We then link communities across adjacent windows based on shared cases. Specifically, we link clusters across periods if they have greater than 0.2 Jaccard similarity, producing 195 valid transitions. This approach allows us to follow the life cycle of

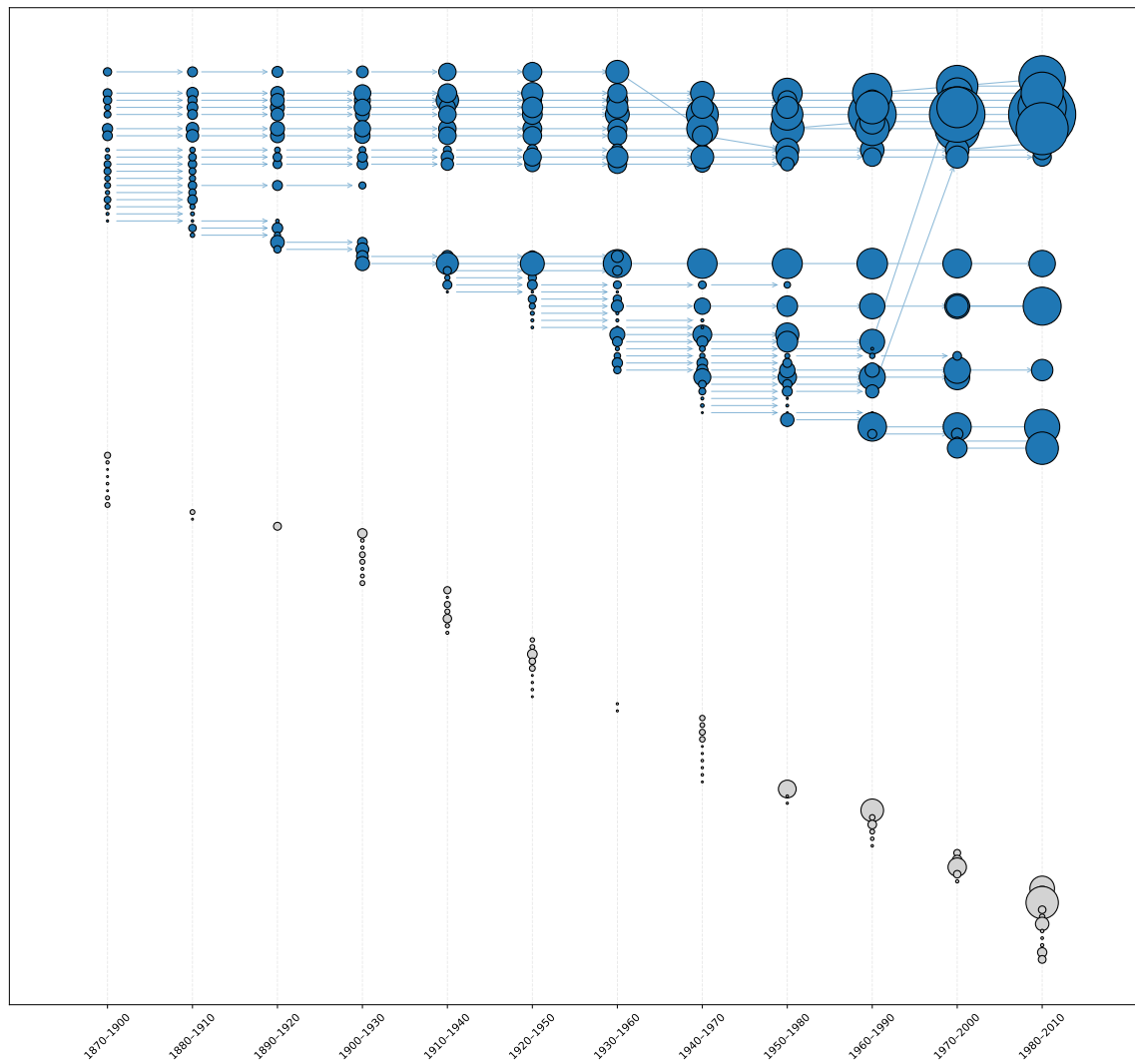


Fig. 2. Evolution of communities: traditions, mutations, and orphans. Each column corresponds to a 30-y sliding window. Blue markers sized proportional to the number of opinions in that community during the window. Horizontal lines connect the same Louvain community across adjacent overlapping windows. Three regimes are visible: persistent traditions (*Top* band), long-lived mutations (*Middle*), and single-period orphans (*Bottom*, gray).

doctrinal clusters and observe whether lines of precedent persist, branch into parallel traditions, or disappear.

Fig. 2 summarizes the resulting trajectories. We observe, first, that some communities (blue, horizontal lines) form persistent traditions, providing long doctrinal backbones across time. Second, long-lived “mutation” branches (blue, diagonal lines) coexist with these traditions for decades before eventually merging or fading. Third, sets of short-lived “orphans” (gray) appear briefly and then disappear. These orphans only count for about 6% of opinions.

Fig. 3 reorganizes the communities from Fig. 2 by canonical doctrinal areas (legend at right), excluding single-period orphans. We see, again, that most areas feature a dominant lineage that persists across many windows. At the same time, many topics produce parallel branches that emerge for a period and then merge or fade, reflecting episodes of experimentation. A few domains, such as military justice, enter the system later and grow rapidly, consistent with statutory or institutional shocks. The presence of both stable lineages and temporary branches suggests the operation of both persistence and disruption.

Next we examine when different legal areas historically developed. We rebuild the citation network using all cases since

1870 and apply Louvain to the full graph, rather than to sliding windows. The algorithm identifies 3,138 communities, with the vast majority (3,100) having only 2 cases each and a handful (13) having between 10 and 1,015 cases. A clear set of 25 key communities emerges, ranging in size from 2,491 cases to 210,683 cases. Quote-based summaries of the content of these communities are listed in *SI Appendix*.

These 25 communities are plotted in Fig. 4, with doctrinal labels at the left. The plot shows kernel-density ridgelines for each community, summarizing the years for which associated decisions were issued. This representation shows the temporal distribution of decisions within each community, allowing us to compare when different doctrinal areas emerged, expanded, and faded away.

Several notable patterns are observed. Constitutional structure and federalism shows a long, continuous thread from the late nineteenth century to the present, indicating a long-running line of precedent. Similarly, patent law, Federal Tort Claims Act, tax, and copyright evolve steadily across more than a century of decisions. By contrast, most other doctrinal areas form coherent communities only after the 1960s or 1970s. Administrative review of agency action, securities and antitrust, ERISA, employ-

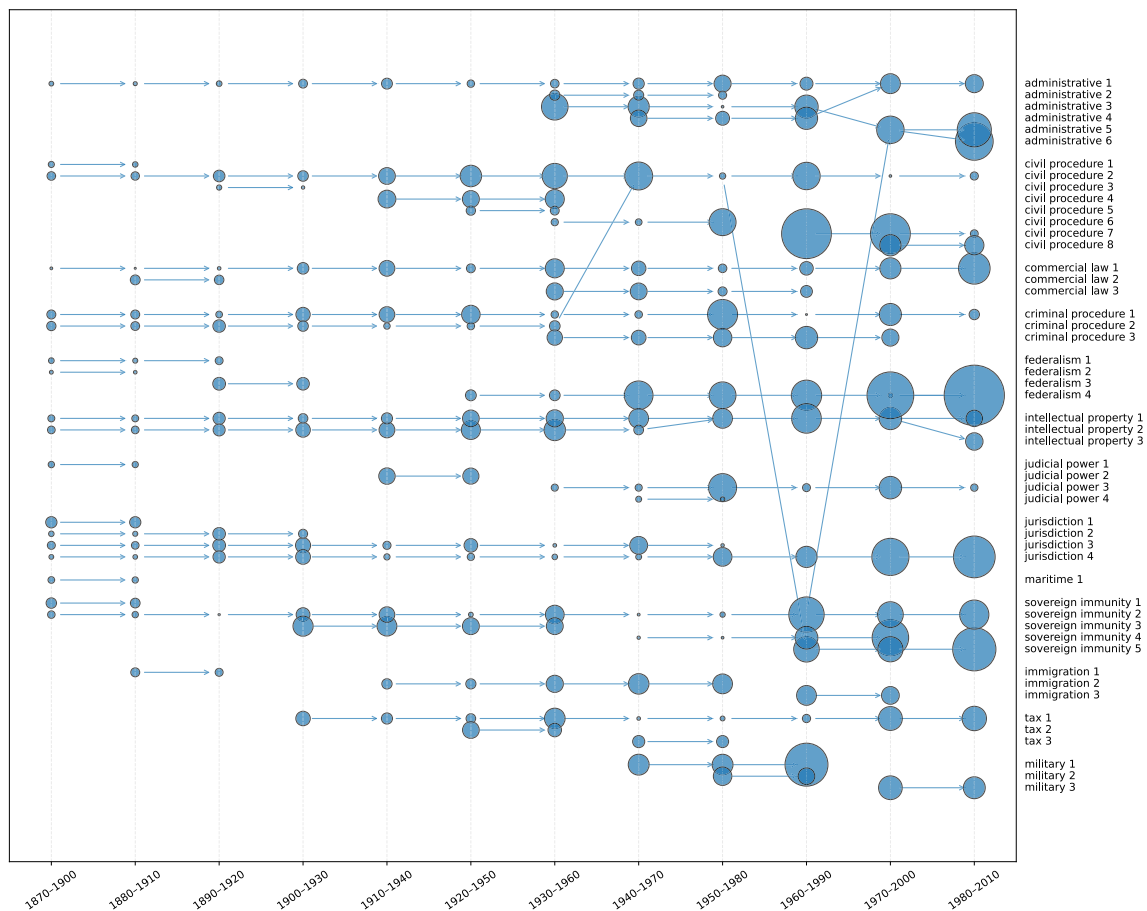


Fig. 3. Evolution of doctrinal communities within topics over time. Communities from Fig. 2 grouped by canonical doctrinal area (legend at Right). Single-period orphans are excluded; only communities surviving at least two windows are shown. Node size is proportional to opinion count within each window. Horizontal lines trace the same community across adjacent windows.

ment discrimination, and special-education litigation all expand sharply in the mid-twentieth century, matching the growth of the modern regulatory and rights-based state. A few topics, such as “military draft obligations” and “FOIA and discovery”, appear as narrow midcentury peaks tied to specific historical episodes.

There could be a number of mechanisms for variation in the persistence of the communities. A formalist null hypothesis would posit that legal innovations can arise from any judge or court and spread through the system purely by virtue of their persuasiveness or analytic strength. In *SI Appendix*, we analyze institutional reasons: Is it driven by initiation by highly influential courts, or is it driven by broad adoption in early years? We find that the latter is more important: Communities that find broad adoption across courts tend to persist longer than others.

Overall, legal Louvain communities help us paint a rich picture. Some legal traditions persist for more than a century, others enter and immediately fade out, while others arise in response to institutional and statutory developments. That supports the view that federal adjudication combines long-term continuity with episodic doctrinal rebuilding.

Legislative Influence on Louvain Communities

The previous sections show that citation communities behave as coherent doctrinal traditions that persist, branch, and sometimes

disappear over time. A natural next question is what forces redirect these trajectories. We explore legislation as a candidate mechanism. Statutes alter the legal environment in which courts adjudicate, potentially shifting the authorities judges rely upon and the precedents they treat as central. We use our citation network data to ask whether the introduction of central federal statutes reshapes citation patterns within communities.

As described above, we have parsed citations to statutes, regulations, and the U.S. Constitution, as well as case law (as used so far). *SI Appendix, Fig. S1* shows the evolution of the share of citations across these four sources. In the common-law system of federal courts, cases are unsurprisingly most important as sources of authority. Statutes are also important, and they have increased in frequency. Citations to regulations are quite infrequent, and citations to the U.S. Constitution (perhaps surprisingly) are actually quite rare. Going forward, we analyze the centrality and impacts of statutes on the caselaw communities.

We match the parsed statutory citations to the caselaw-based communities, described above. For each of the 25 whole-period communities, we compute the share of citations going to the various sources of authority. Our first question is which communities are more oriented around key statutes. First, we compute the share of citations going to USC provisions—measuring how statute-oriented each community’s reasoning is. Second, we identify the most-cited USC provisions and compute the share of all statutory citations going to the five most-cited provisions (the ranking across communities is not sensitive to

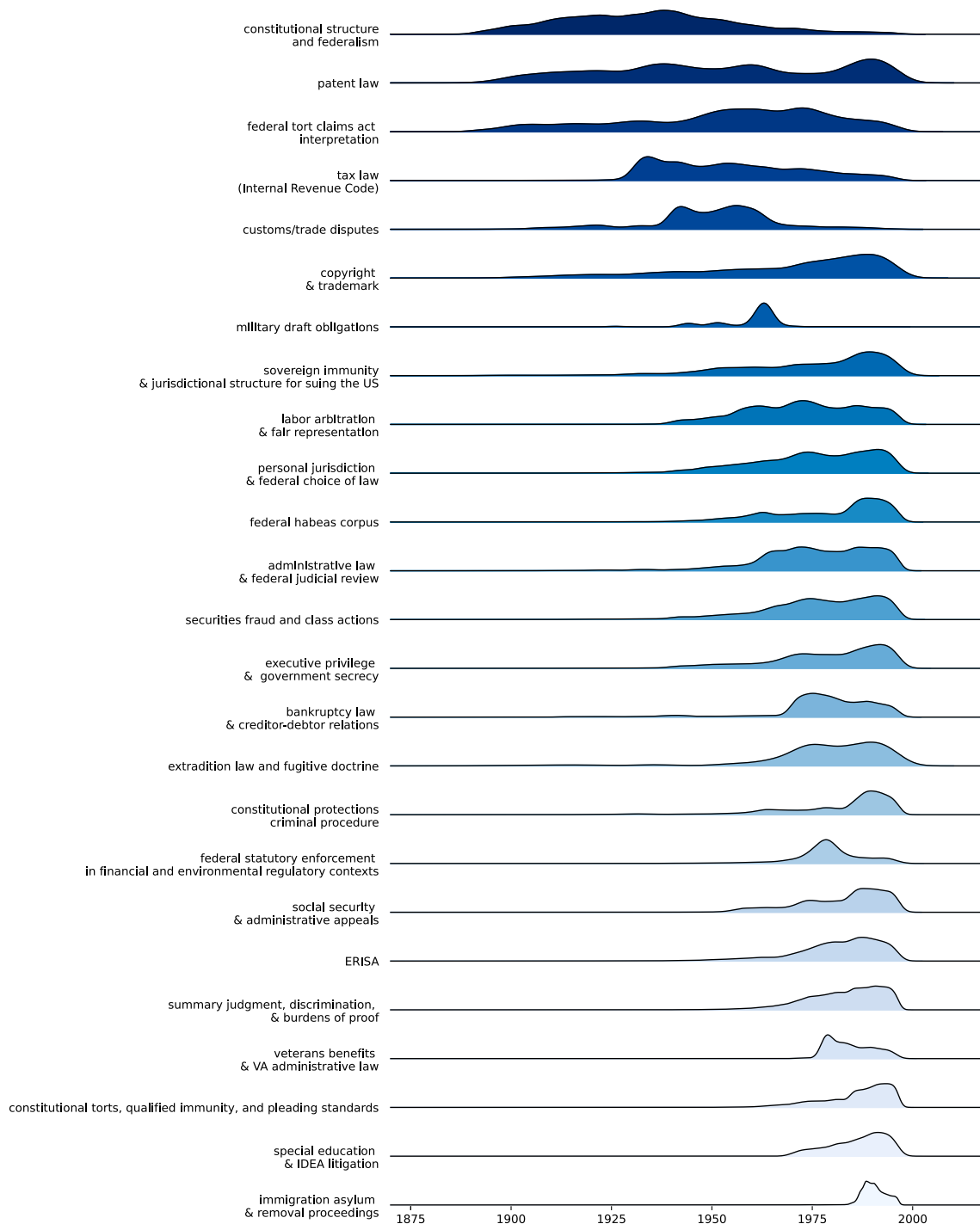


Fig. 4. Temporal evolution of citation communities. Kernel-density ridgelines of decision-year distributions for communities identified by applying Louvain to the full 1870–2024 citation graph. Each row represents one community, labeled by its doctrinal area and ordered from earliest to latest modal development.

including more or fewer key provisions). This gives a measure of statutory concentration by community.

Table 1 characterizes the statutory foundations of each community. The variation is striking. At one extreme, “veterans benefits”, “special education”, and “bankruptcy” devote 30 to 35% of their citations to USC provisions, reflecting communities whose reasoning is deeply statute-anchored. At the other extreme, “military draft obligations” (7%) and “patent law” (13%) rely overwhelmingly on case-to-case citation. Statutory concentration also varies independently of the overall USC share: special education and immigration each direct over 60% of their

USC citations to just five provisions, while administrative law and constitutional structure & federalism spread their statutory references across many provisions (Top-5 Rates of 7% and 4%, respectively). The full citation-source decomposition across all four authority types is reported in *SI Appendix, Table S1*.

Using the key statutes from Table 1, we next investigate the temporal dynamic between federal adjudication and those statutes. Within each community, we pair the 50 most-cited cases with the 5 most-cited statutes and compute the gap between each pair (case decision year — statute enactment year). The idea is to

Table 1. Statutory foundations of legal citation communities

Community	Cases	USC (%)	Top-5 rate	Key statutes
1 Veterans benefits & VA administrative law	3,315	35%	30%	38 U.S.C. §7292, <i>Veterans Claims Court</i> (8%); 38 U.S.C. §7104, <i>Board of Veterans' Appeals</i> (6%); 38 U.S.C. §7261, <i>Veterans Claims Court</i> (6%)
2 Special education & IDEA litigation	2,491	30%	63%	20 U.S.C. §1415, <i>IDEA</i> (30%); 20 U.S.C. §1412, <i>IDEA</i> (9%); 20 U.S.C. §1400, <i>IDEA</i> (8%)
3 Bankruptcy law & creditor-debtor relations	80,040	30%	24%	11 U.S.C. §523, <i>Bankruptcy Code</i> (7%); 28 U.S.C. §157, <i>Bankruptcy Judges</i> (5%); 11 U.S.C. §362, <i>Bankruptcy Code</i> (5%)
4 Pretrial detention & extradition	2,937	27%	36%	18 U.S.C. §3142, <i>Release and Detention</i> (19%); 15 U.S.C. §2802, <i>Petroleum Marketing Practices</i> (7%); 42 U.S.C. §11603, <i>International Child Abduction Remedies</i> (4%)
5 Social security & administrative appeals	22,621	27%	34%	42 U.S.C. §405, <i>Social Security</i> (12%); 42 U.S.C. §423, <i>Social Security</i> (8%); 19 U.S.C. §1677b, <i>Tariff Act</i> (5%)
6 Sovereign immunity & jurisdictional structure for suing the US	24,095	26%	17%	28 U.S.C. §1491, <i>Court of Federal Claims</i> (6%); 31 U.S.C. §3729, <i>False Claims</i> (4%); 31 U.S.C. §3730, <i>False Claims</i> (4%)
7 Financial and environmental regulatory enforcement	4,538	26%	42%	42 U.S.C. §9607, <i>CERCLA</i> (13%); 12 U.S.C. §1821, <i>FDIC</i> (11%); 42 U.S.C. §9601, <i>CERCLA</i> (8%)
8 Administrative law & federal judicial review	114,585	25%	7%	42 U.S.C. §1983, <i>Civil Rights</i> (2%); 5 U.S.C. §706, <i>APA</i> (2%); 28 U.S.C. §1331, <i>Jurisdiction</i> (1%)
9 ERISA	30,832	24%	28%	29 U.S.C. §1132, <i>ERISA</i> (9%); 28 U.S.C. §1332, <i>Jurisdiction</i> (6%); 28 U.S.C. §1441, <i>Removal</i> (5%)
10 Immigration asylum & removal proceedings	48,579	23%	60%	8 U.S.C. §1252, <i>INA</i> (31%); 8 U.S.C. §1158, <i>INA</i> (11%); 8 U.S.C. §1101, <i>INA</i> (8%)
11 Labor arbitration & fair representation	33,347	22%	26%	29 U.S.C. §158, <i>NLRA</i> (8%); 29 U.S.C. §8, <i>Labor Statistics</i> (6%); 29 U.S.C. §160, <i>NLRA</i> (5%)
12 Constitutional protections criminal procedure	210,683	20%	23%	18 U.S.C. §3553, <i>Sentencing</i> (7%); 21 U.S.C. §841, <i>Drug Abuse Prevention</i> (6%); 18 U.S.C. §924, <i>Firearms</i> (4%)
13 Federal habeas corpus	93,163	18%	51%	28 U.S.C. §2254, <i>Habeas Corpus</i> (22%); 28 U.S.C. §2253, <i>Habeas Corpus</i> (11%); 28 U.S.C. §2255, <i>Habeas Corpus</i> (10%)
14 Securities & antitrust	55,891	17%	13%	15 U.S.C. §1, <i>Sherman Act</i> (3%); 18 U.S.C. §1962, <i>RICO</i> (3%); 15 U.S.C. §78j, <i>Securities Exchange Act</i> (2%)
15 Copyright & trademark	23,216	17%	22%	15 U.S.C. §1125, <i>Lanham Act</i> (10%); 15 U.S.C. §1114, <i>Lanham Act</i> (4%); 17 U.S.C. §101, <i>Copyright Act</i> (3%)
16 FOIA & discovery	25,986	17%	33%	5 U.S.C. §552, <i>FOIA</i> (21%); 28 U.S.C. §455, <i>Judicial Disqualification</i> (4%); 28 U.S.C. §1291, <i>Courts of Appeals</i> (3%)
17 Federal tort claims act interpretation	46,090	16%	16%	28 U.S.C. §1346, <i>Jurisdiction</i> (6%); 28 U.S.C. §2680, <i>Tort Claims Procedure</i> (4%); 46 U.S.C. §688, <i>Jones Act</i> (2%)
18 Tax law (Internal Revenue Code)	60,235	16%	9%	26 U.S.C. §7422, <i>Tax Judicial Proceedings</i> (2%); 28 U.S.C. §1346, <i>Jurisdiction</i> (2%); 26 U.S.C. §6323, <i>Tax Liens</i> (2%)
19 Employment discrimination	137,803	15%	17%	42 U.S.C. §2000e, <i>Title VII</i> (5%); 42 U.S.C. §1981, <i>Civil Rights Act of 1866</i> (4%); 42 U.S.C. §2000e-5, <i>Title VII</i> (3%)
20 Personal jurisdiction & federal choice of law	32,178	14%	27%	28 U.S.C. §1404, <i>Venue</i> (10%); 28 U.S.C. §1391, <i>Venue</i> (7%); 28 U.S.C. §1332, <i>Jurisdiction</i> (4%)
21 Constitutional torts	100,823	14%	35%	42 U.S.C. §1983, <i>Civil Rights</i> (21%); 28 U.S.C. §1291, <i>Courts of Appeals</i> (5%); 28 U.S.C. §1915, <i>In Forma Pauperis</i> (5%)
22 Customs/trade disputes	25,863	14%	26%	28 U.S.C. §1581, <i>Court of International Trade</i> (7%); 19 U.S.C. §1514, <i>Tariff Act</i> (6%); 19 U.S.C. §1401a, <i>Tariff Act</i> (5%)
23 Constitutional structure and federalism	121,447	13%	4%	28 U.S.C. §41, <i>Courts of Appeals</i> (1%); 49 U.S.C. §1, <i>Interstate Commerce Act</i> (1%); 28 U.S.C. §1331, <i>Jurisdiction</i> (1%)
24 Patent law	35,763	13%	37%	35 U.S.C. §102, <i>Patent Act</i> (10%); 35 U.S.C. §103, <i>Patent Act</i> (9%); 35 U.S.C. §112, <i>Patent Act</i> (9%)
25 Military draft obligations	3,024	7%	24%	28 U.S.C. §2241, <i>Habeas Corpus</i> (10%); 10 U.S.C. §673a, <i>Active Duty</i> (4%); 28 U.S.C. §1331, <i>Jurisdiction</i> (4%)

Notes: All 25 whole-period Louvain communities, ranked by USC citation share. USC (%): share of citations that go to USC provisions. Top-5 Rate: share of USC citations going to the five most-cited provisions. Parenthetical percentages: each provision's share of total USC citations within that community. Key statutes include the most cited statutes from the community cases, with the associated USC chapter headings.

understand whether key cases precede key statutes, or vice versa, and whether that varies across areas of law.

Fig. 5 displays the distribution of these distances as a ridge plot, ranked by mean delay. We see substantial heterogeneity across communities. Constitutional structure & federalism shows a mean gap of -42 y: Landmark cases in this community predate the key statutes, consistent with a body of law built on constitutional principles that were later partially codified. At the other extreme, immigration (+42 y) and securities &

antitrust (+40 y) feature cases that were mostly decided decades after their foundational statutes, reflecting a pattern of ongoing judicial interpretation of established legislative frameworks. Most communities cluster in the +10 to +25 y range, where statutes provide the initial impetus and case law develops in the subsequent decades.

Now we focus more closely—for each community, we identify the single most-cited statutory provision and its year of enactment. These central statutes are listed in *SI Appendix, Table S2*.

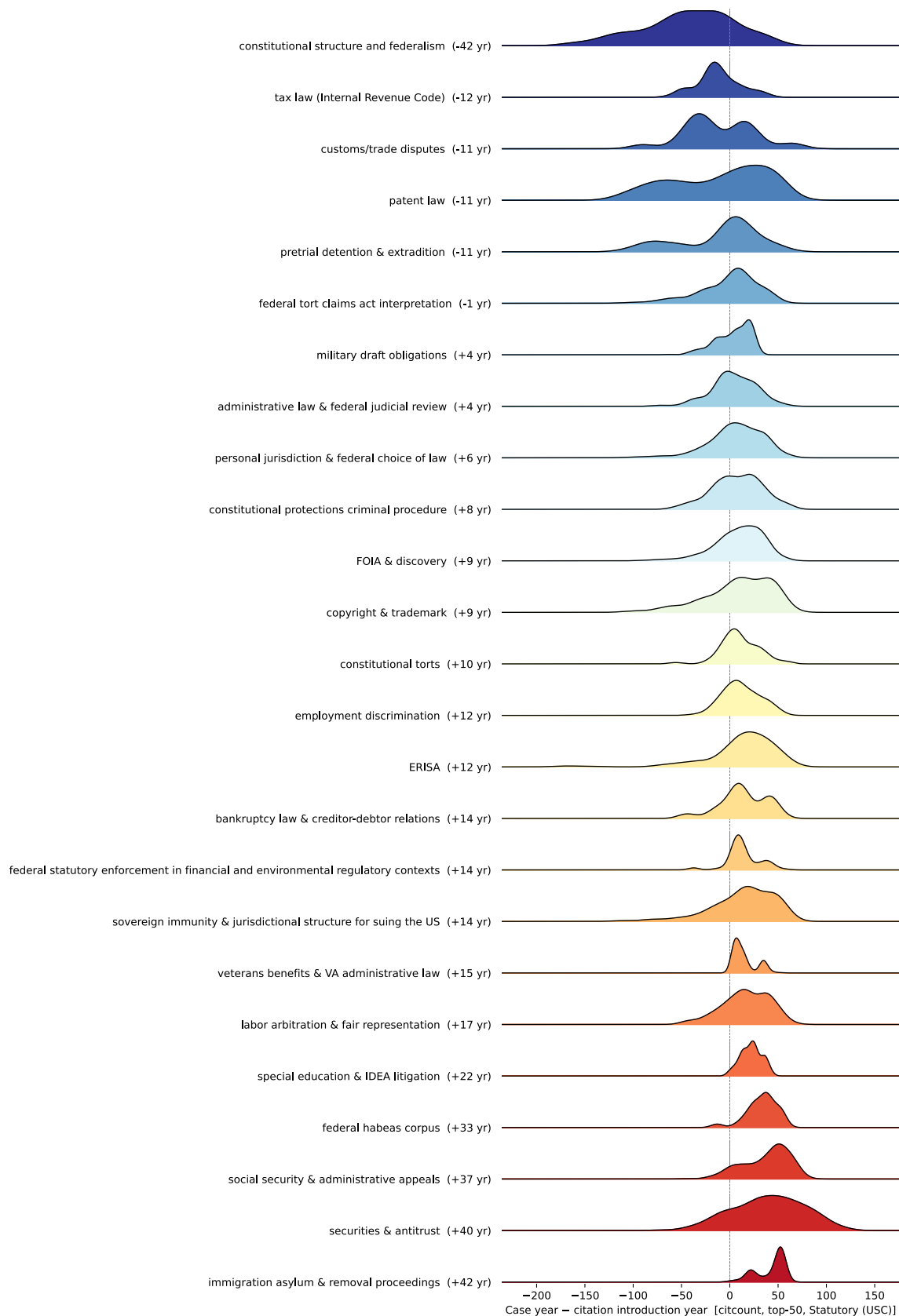


Fig. 5. Temporal distance between landmark cases and their central statutes. Each row represents one community. Distributions show the temporal gap (case decision year – statute enactment year) computed by pairing the 50 most-cited cases in each community with the 5 most-cited USC provisions. Mean distance is shown in parentheses. Negative values indicate landmark cases predate the key statutes; positive values indicate cases follow statutory enactment. Communities are ranked by mean gap.

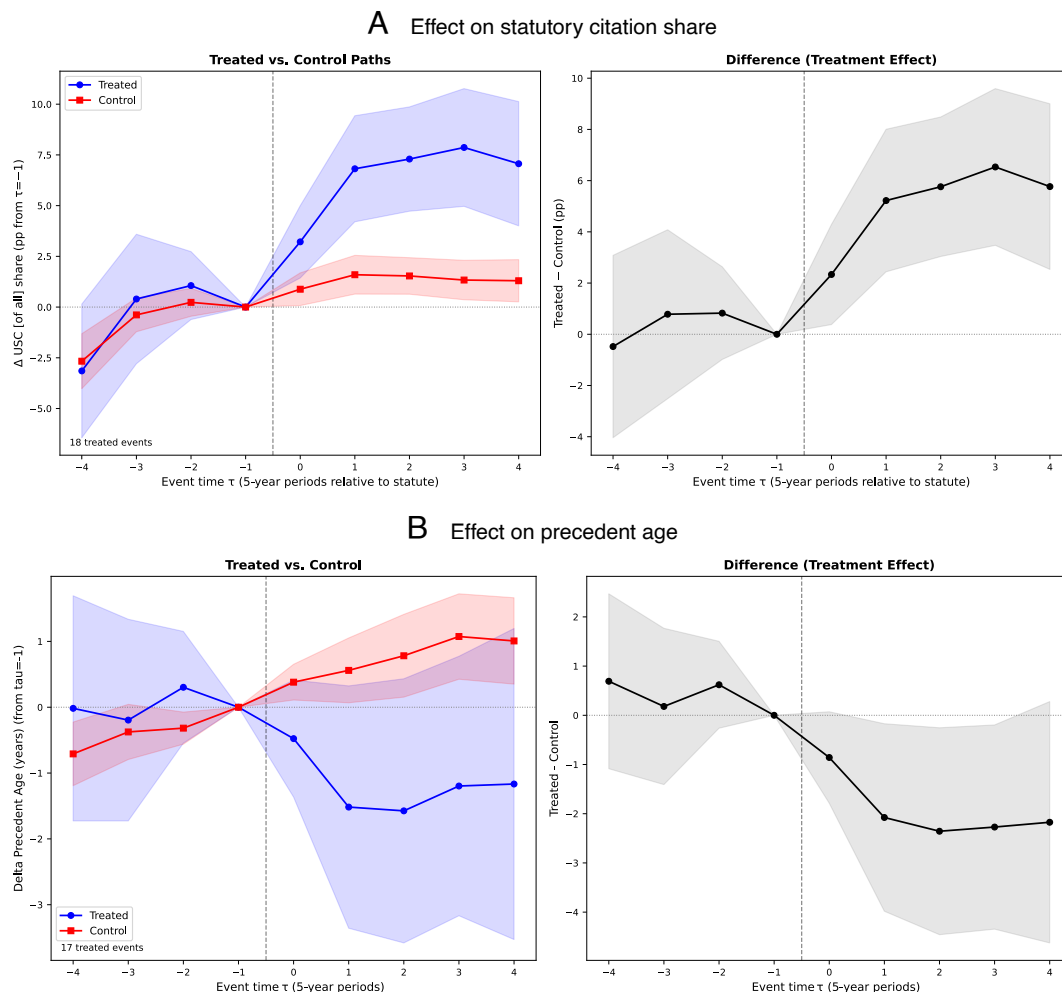


Fig. 6. Event studies: effect of central statute introduction on citation patterns. Pooled difference-in-differences event study across 18 communities with identifiable central statutes. In each panel, *Left* plots show treated (blue) and control (red) paths; *Right* plots show the difference. Event time τ is measured in 5-y bins relative to statute enactment, normalized to $\tau = -1$. Controls are communities whose own central statute year is > 20 y from the treated event. Shaded bands show 95% CIs. Panel (A): outcome is USC citation share (percentage points). Panel (B): outcome is average within-community precedent age (mean age of cited cases, in years); a posttreatment decline indicates judges cite newer cases after the central statute is introduced.

The introduction years span from 1890 (Sherman Antitrust Act) to 1987 (veterans benefits), providing natural variation for a quasi-experimental design. Our goal is to estimate a causal effect of statutory introduction on the citation patterns of the courts, using the communities as potentially treated units.

Specifically: For each of the 18 communities with an identifiable central statute, we define the treated unit as that community's citation trajectory centered on the statute's enactment year. The control group consists of all other communities whose own central statute year is more than 20 y away from the event, ensuring controls are not contaminated by their own statutory shocks. Event time τ is measured in 5-y bins relative to the statute introduction, normalized to $\tau = -1$. We then estimate a stacked event study for how citation outcomes vary for the treated communities relative to their respective controls.

First we estimate the effect on citations to statutes. This can be understood as a "first stage" effect of the key statute, which should mechanically shift citations toward that legislative framework. Fig. 6 presents the event-study estimates. The result is as expected: Following introduction of the central statute, the treated community's share of citations to USC authority rises sharply relative to control communities. The divergence begins at $\tau = 0$ and grows to approximately 6 percentage

points by $\tau = +3$ (15 y after enactment). The pretreatment paths are roughly parallel, supporting the identifying assumption. Central statute introduction measurably reshapes the citation mix within communities, shifting judicial attention toward statutory authority.

A more substantive question is whether these central statutes also disrupt established lines of caselaw precedent. We investigate by applying the same event-study design to a different outcome: the average age of within-community case-to-case citations (precedent age). If statutes merely add a new layer of citations without displacing existing case law, precedent age should be unaffected. If, instead, statutes trigger a break from prior reasoning, precedent age should decline as judges build new lines of precedent.

Fig. 6 presents the results. The treated path shows a clear posttreatment decline in precedent age relative to control. The difference reaches approximately -2 y by $\tau = +2$, meaning that judges in treated communities begin citing materially newer cases after the central statute is introduced. The pretrend is flat, with no systematic divergence before the event. This evidence indicates that major statutory introductions are disruptive: They do not merely supplement existing case law but redirect judicial attention toward newer precedent, consistent with statutes opening new

lines of legal reasoning that partially displace older citation patterns.

These results add evidence that legal Louvain communities capture meaningful legal structure. Communities have distinct statutory foundations that vary systematically in concentration and timing. The introduction of central statutes causally increases citation to statutory authority (the first stage) and simultaneously disrupts established precedent patterns (the reduced form). These results show that identifiable legislative events reshape citation patterns.

Conclusion

This paper offers an empirical framework for studying how legal traditions evolve. By applying the Louvain algorithm to the complete federal citation network, we identify communities of cases that behave as coherent doctrinal units. These communities recover known institutional structures and display distinctive patterns of legal authority, suggesting that citation behavior contains enough information to reveal the underlying organization of doctrine. In this sense, the structure of precedent resembles other complex knowledge systems whose evolution can be traced through network analysis (4, 8).

The results speak to long-standing debates about the evolution of law. Formalist accounts emphasize the internal logic of doctrine, while realist perspectives stress the institutional context in which legal reasoning develops (1). Some legal traditions are

long-lived and others are short-lived. We find that the staggered introduction of central statutes increases reliance on statutory authority and shifts judicial attention toward newer precedent. Legislation therefore does not simply layer new rules onto existing doctrine but can redirect the trajectory of precedent itself.

Taken together, these findings depict the common-law system as a large-scale adaptive network. Persistent doctrinal traditions provide structural continuity, while statutory interventions periodically redirect the flow of legal influence through the citation graph. This perspective connects the evolution of legal doctrine to broader patterns observed in other knowledge systems, where innovation occurs through a balance of stability and disruption (2). Future work could build on this framework by integrating textual analysis of legal reasoning, modeling interactions between case law and statutes as multiplex networks, and developing predictive models of doctrinal change.

Data, Materials, and Software Availability. Replication code and data are available at Zenodo: <https://doi.org/10.5281/zenodo.19982763> (43). The Case Law Access Project data are available at <https://case.law> (44). The LePaRD dataset is available at <https://huggingface.co/datasets/rmahari/LePaRD> (45).

Author affiliations: ^aDepartment of Social Sciences, ETH Zurich, Zurich 8092, Switzerland; ^bDepartment of Political Science, Columbia University, New York, NY 10027; ^cStanford Law School, Stanford University, Stanford, CA 94305; and ^dDepartment of Economics, Columbia University, New York, NY 10027

Author contributions: E.A., C.C., R.M., and S.N. designed research; performed research; analyzed data; and wrote the paper.

- O. W. Holmes Jr., *The Common Law* (Routledge, 1881).
- F. A. Hayek, *Law, Legislation and Liberty, Volume 1: Rules and Order* (University of Chicago Press, Chicago, 1973).
- L. M. Friedman, *A History of American Law* (Simon & Schuster, New York, ed. 1, 1975).
- T. A. Smith, The web of law. *San Diego Law Rev.* **44**, 315–358 (2007).
- J. H. Fowler, T. R. Johnson, J. F. I. Spriggs, S. Jeon, P. J. Wahlbeck, Network analysis and the law: Measuring the legal importance of supreme court precedents. *Polit. Anal.* **15**, 324–346 (2007).
- V. D. Blondel, J. L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks. *J. Stat. Mech. Theory Exp.* **2008**, P10008 (2008).
- R. A. Posner, *The Problems of Jurisprudence* (Harvard University Press, 1993).
- D. M. Katz, M. J. I. Bommarito, J. Blackman, Complex societies and the growth of the law. *Sci. Rep.* **10**, 18737 (2020).
- G. Calabresi, *The Costs of Accidents: A Legal and Economic Analysis* (Yale University Press, New Haven, 1970).
- R. A. Posner, *Economic Analysis of Law* (Little, Brown & Co., Boston, ed. 1, 1973).
- P. H. Rubin, Why is the common law efficient? *J. Leg. Stud.* **6**, 51–63 (1977).
- G. L. Priest, The common law process and the selection of efficient rules. *J. Leg. Stud.* **6**, 65–82 (1977).
- W. M. Landes, R. A. Posner, Legal precedent: A theoretical and empirical analysis. *J. Law Econ.* **19**, 249–307 (1976).
- N. Gennaioli, A. Shleifer, The evolution of common law. *J. Polit. Econ.* **115**, 43–68 (2007).
- E. Ash, C. Goessmann, S. Naidu, Scaling laws: Legal and social complexity in US localities. *Philos. Trans. R. Soc. A* **382**, 20230151 (2024).
- S. Kojaku et al., Uncovering the universal dynamics of citation systems: From science of science to law of law and patterns of patents. *arXiv [Preprint]* (2025). <https://arxiv.org/html/2501.15552v1>.
- E. Ash, M. Morelli, M. Vannoni, More laws, more growth? Evidence from U.S. states. *J. Polit. Econ.* **133**, e734874 (2025).
- J. B. Ruhl, Thinking of environmental law as a complex adaptive system: How to clean up the environment by making a mess of environmental law. *Houst. Law Rev.* **34**, 933–1004 (1997).
- D. M. Katz, M. J. I. Bommarito, Measuring the complexity of the law: The United States code. *Artif. Intell. Law* **22**, 337–374 (2014).
- R. A. Posner, An economic analysis of the use of citations in the law. *Am. Law Econ. Rev.* **2**, 381–406 (2000).
- Y. Lupu, J. H. Fowler, Strategic citations to precedent on the U.S. supreme court. *J. Leg. Stud.* **42**, 151–186 (2013).
- S. Verma, A. Parthasarathy, D. L. Chen, *The Genealogy of Ideology: Predicting Agreement and Persuasive Memes in the U.S.* (United Kingdom, ACM, London, 2017), pp. 253–256.
- R. Mahari, S. C. Lera, Early career citations capture judicial idiosyncrasies and predict judgments. *arXiv [Preprint]* (2024). <https://arxiv.org/abs/2410.00725> (Accessed 22 May 2026).
- F. B. Cross, J. F. I. Spriggs, The most important (and best) supreme court opinions and justices. *Emory LJ* **60**, 407 (2010).
- H. Derlén, J. Lindholm, Characteristics of precedent: European Court of Justice case law in three dimensions. *Ger. Law J.* **19**, 1073–1098 (2018).
- G. Leibon, S. Pauls, D. Rockmore, S. Sullivan, Bending the law: geometric tools for quantifying influence in the multinetwork of legal opinions. *Artif. Intell. Law* **26**, 145–167 (2018).
- M. J. I. Bommarito, D. M. Katz, J. Zelner, "Law as a seamless web? Comparison of various network representations of the united states supreme court corpus (1791–2005)" in *Proceedings of the 12th International Conference on Artificial Intelligence and Law* (Association for Computing Machinery, New York, NY, 2009).
- A. Mirshahvalad, P. Holme, M. Rosvall, Significant communities in large sparse networks. *PLoS one* **7**, e33721 (2012).
- J. R. Clough, J. Gollings, T. V. Loach, T. S. Evans, Transitive reduction of citation networks. *J. Complex Netw.* **3**, 189–203 (2015).
- J. Soh, Simulating subject communities in case law citation networks. *Front. Phys.* **9**, 665563 (2021).
- E. Mohammadi, M. Hildebrandt, T. van der Meer, Combining topic modelling and citation network analysis to study case law from the European Court of Human Rights on the right to respect for private and family life. *Artif. Intell. Law*, 1–24 (2025).
- J. T. Carmichael, M. J. I. Bommarito, D. M. Katz, Predicting supreme court decisions using citation networks of lower court opinions. *J. Empir. Leg. Stud.* **14**, 307–359 (2017).
- R. C. Black, J. F. I. Spriggs, The citation and depreciation of U.S. supreme court precedent. *J. Empir. Leg. Stud.* **10**, 325–358 (2013).
- R. Whalen, Judicial citations & common law evolution: A chronometric analysis of high court citations. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2627725. Accessed 22 May 2026.
- S. Lee, T. Kim, J. Yoon, H. Youn, When common law ages: Two centuries of growing inertia in us judicial opinions. *arXiv [Preprint]* (2025). <http://arxiv.org/abs/2410.04493> (Accessed 22 May 2026).
- W. N. J. Eskridge, Overriding supreme court statutory interpretation decisions. *Yale Law J.* **101**, 331–423 (1991).
- D. A. Widiss, Shadow precedents and the statutory override game. *Notre Dame Law Rev.* **84**, 597–680 (2009).
- L. Hager, A. Uribe-McGuire, Court-curbing bills, the supreme court, and judicial review. *J. Law Court.* **12**, 303–324 (2024).
- C. Chiopris, "The diffusion of ideas." Working Paper (2024). https://www.caterinachiopris.com/_files/ugd/b45409_ba6a9e005f5c428ba5811d3dc219580.pdf.
- G. A. Caldeira, The transmission of legal precedent among state supreme courts. *Am. Polit. Sci. Rev.* **79**, 178–194 (1985).
- Y. Lupu, E. Voeten, Precedent in international courts: A network analysis of case citations by the European court of human rights. *Br. J. Polit. Sci.* **42**, 413–439 (2012).
- R. Mahari, D. Stambach, E. Ash, A. Pentland, "Lepard: A large-scale dataset of judicial citations to precedent" in *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (ACL 2024), Volume 1: Long Papers* (Association for Computational Linguistics, Bangkok, Thailand, 2024), vol. 1, pp. 9863–9877.
- E. Ash, C. Chiopris, R. Mahari, S. Naidu, Replication code and data for "Continuity and change in US legal tradition: Evidence from judicial citation communities." Zenodo. <https://doi.org/10.5281/zenodo.19982763>. Accessed 22 May 2026.
- CaseLaw Access Project, Case Law Access Project data. Harvard Law School Library. <https://case.law>. Accessed 22 May 2026.
- R. Mahari, D. Stambach, E. Ash, A. Pentland, LePaRD dataset. Hugging Face Datasets. <https://huggingface.co/datasets/rmahari/LePaRD>. Accessed 22 May 2026.

SI Appendix

Continuity and Change in U.S. Legal Tradition: Evidence from Judicial Citation Communities

Elliott Ash, Caterina Chiopris, Robert Mahari, Suresh Naidu

A.1 Community Topic Validation Using Quotes Data

In the main text, we relied on LLM constructions of topics from case names. We validate the main legal topics covered by the different citation network communities here by extracting the texts of quotes from cases in those communities. We draw on the LeParD dataset, developed previously by co-authors Mahari and Ash [1]. We linked our community assignments to the “destination” case (that is, the quoted case), and took the associated text of the quote.

For each of the 25 topics covered in Figure 3 of the main text, we sampled 100 quotes. We then prompted an aligned LLM (gpt-4o) to extract the shared legal reasoning, using the following prompt:

You are analyzing legal quotes from court cases. Below are 100 legal quotes that have been categorized under the same row/category.

Please analyze these quotes and extract the distinct line of legal reasoning that distinguishes and unifies these texts. Provide a short answer up to 3 sentences that identifies the common legal principle, doctrine, or reasoning pattern that makes these quotes belong to the same category.

A sample of summaries are contained below. We first manually compared the summaries to the quotes for each of the topics and found that they were satisfactory. We then manually compared the quote-based summary to the topic extracted via the case names. Overall, the fit was very good. We made some minor adjustments to the topic labels based on the quotes analysis. Overall, we were impressed with the ability of the aligned LLM to extract coherent legal topics from either case names or quote texts.

1. **Personal jurisdiction & federal choice of law.** The common legal principle unifying these quotes is the doctrine of personal jurisdiction and its interplay with international law and procedural fairness. The quotes frequently reference the necessity of establishing “minimum contacts” with a forum state to satisfy due process requirements, ensuring that legal proceedings do not offend “traditional notions of fair play and substantial justice.”
2. **ERISA.** The legal quotes from this community primarily revolve around the principles of federal jurisdiction, removal, and the preemption of state law by federal law, particularly in the context of ERISA and civil rights. The unifying legal reasoning pattern is the delineation of federal court jurisdiction over cases that arise under federal law, the conditions under which cases can be removed from state to federal court, and the preemption of state laws by federal statutes.

3. **Federal tort claims act interpretation.** The common legal principle unifying these quotes is the doctrine of governmental and maritime liability, particularly focusing on the discretionary function exception and the liability of the United States under the Federal Tort Claims Act (FTCA). These quotes explore the boundaries of government liability, emphasizing the discretionary nature of certain governmental actions and the conditions under which the government can be held liable.
4. **Administrative law & federal judicial review.** The legal quotes from this community primarily revolve around the principles of judicial review, administrative law, and the limits of judicial power. They emphasize the necessity for courts to adhere to statutory mandates, respect agency discretion, and avoid overstepping into areas reserved for legislative or executive branches unless there is a clear legal basis or constitutional violation.
5. **Labor arbitration & fair representation.** The common legal principle unifying these quotes is the protection and enforcement of employee rights and benefits, particularly in the context of labor relations, arbitration, and veterans' reemployment rights. The quotes emphasize the importance of maintaining seniority, status, and pay for employees, the enforceability of arbitration agreements, and the equitable treatment of veterans returning to civilian employment.
6. **Bankruptcy law & creditor-debtor relations.** The common legal principle unifying these quotes is the treatment of fiduciary duties and the consequences of breaches within the context of bankruptcy law. The quotes frequently address issues such as fraud, defalcation, and the responsibilities of individuals acting in fiduciary capacities, particularly in relation to the discharge of debts and the protection of creditors' rights.
7. **Securities & antitrust.** The common legal principle unifying these quotes is the regulation and enforcement of fair competition and trade practices, primarily under antitrust and securities laws. The quotes reflect a focus on preventing anti-competitive behavior, fraudulent practices, and ensuring fair market conditions, often through class actions and regulatory oversight.
8. **FOIA & discovery.** The common legal principle unifying these quotes is the doctrine of confidentiality and privilege in legal proceedings, particularly concerning the protection of sensitive information from disclosure. This includes the protection of inter-agency communications, personal privacy, and privileged materials prepared in anticipation of litigation.
9. **Patent law.** The common legal principle unifying these quotes is the doctrine of "obviousness" in patent law, particularly as it relates to the assessment of whether an invention is patentable. This doctrine requires evaluating the scope and content of prior art, the differences between the prior art and the claimed invention, and the level of ordinary skill in the pertinent art to determine if the invention would have been obvious at the time it was made.

10. **Sovereign immunity & jurisdictional structure for suing the US.** The common legal principle unifying these quotes is the doctrine of sovereign immunity and its exceptions, particularly in the context of government contracts and claims against the United States. The quotes frequently address issues related to the conditions under which the government consents to be sued, the necessity of privity of contract for claims, and the limitations on claims for damages against the government.
11. **Tax law (Internal Revenue Code).** The common legal principle unifying these quotes is the interpretation and application of tax law, particularly focusing on procedural aspects, taxpayer obligations, and the government's authority in tax collection and enforcement. The quotes reflect a consistent theme of ensuring compliance with statutory requirements and the balance between federal authority and state law in determining tax liabilities and rights.
12. **Immigration asylum & removal proceedings.** The legal quotes in this community primarily revolve around the principles and standards governing asylum and deportation proceedings, particularly focusing on the evidentiary requirements and legal standards for establishing a well-founded fear of persecution. The common legal reasoning pattern involves the necessity for claims to be supported by "reasonable, substantial, and probative evidence."
13. **Veterans benefits & VA administrative law.** The common legal principle unifying these quotes is the specialized adjudicative framework governing veterans' benefits claims, particularly focusing on the standards and procedures for reviewing and challenging decisions made by the Department of Veterans Affairs and the Board of Veterans' Appeals. This includes the principles of "clear and unmistakable error," the duty to assist claimants, and the interpretation of statutes and regulations in favor of veterans.
14. **Copyright & trademark.** The common legal principle unifying these quotes is the protection and enforcement of intellectual property rights, particularly focusing on trademark and copyright law. The quotes collectively address issues such as infringement, likelihood of confusion, the functionality doctrine, and the balance of equities in granting relief.
15. **Social security & administrative appeals.** The common legal principle unifying these quotes is the standard of "substantial evidence" in the context of disability determinations under social security law. This standard requires that decisions be supported by more than a mere scintilla of evidence, meaning such relevant evidence as a reasonable mind might accept as adequate to support a conclusion.
16. **Special education & IDEA litigation.** The common legal principle unifying these quotes is the interpretation and application of the Individuals with Disabilities Education Act (IDEA). The quotes collectively emphasize the procedural safeguards and

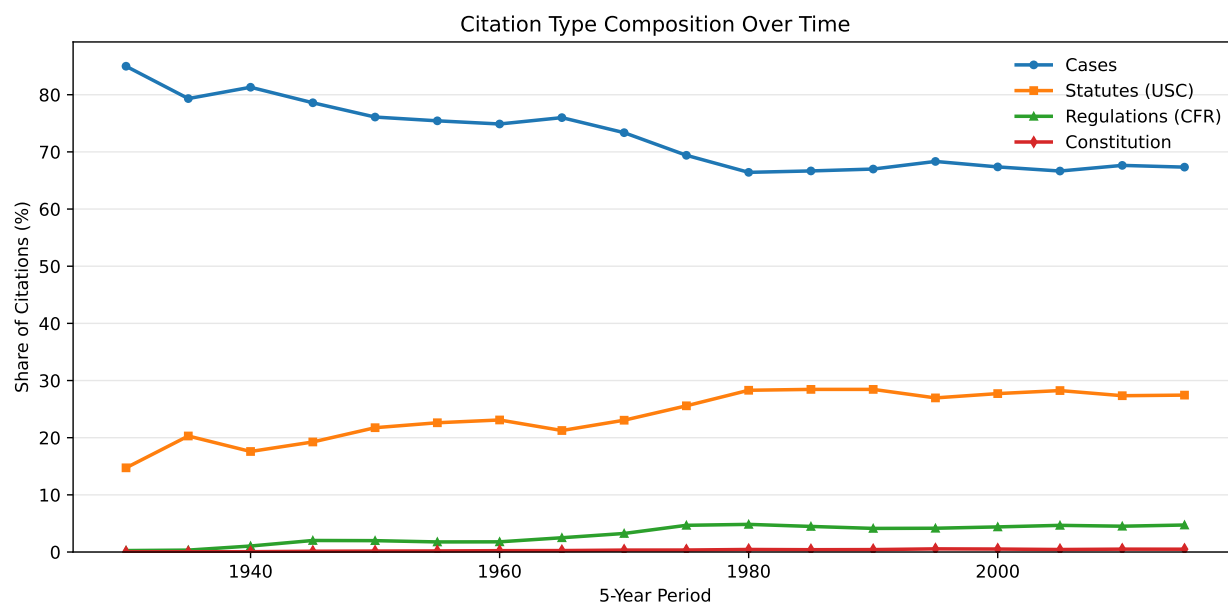
substantive standards required to ensure that children with disabilities receive a “free appropriate public education” tailored to their unique needs.

17. **Customs/trade disputes.** The common legal principle unifying these quotes is the interpretation and application of customs and tariff laws, specifically focusing on the classification, valuation, and duty assessment of imported merchandise. The quotes reflect a consistent emphasis on the precise definition and categorization of goods, the use of common or dictionary meanings for tariff terms, and the importance of legislative intent.
18. **Pretrial detention & extradition.** The legal quotes from this community primarily revolve around the principles and procedures of international law concerning extradition and the Hague Convention on the Civil Aspects of International Child Abduction. The unifying legal reasoning pattern is the emphasis on the procedural and jurisdictional aspects of international treaties and conventions.
19. **Military draft obligations.** The common legal principle unifying these quotes is the procedural and substantive framework governing conscientious objector status and exemptions from military service under the Selective Service Act. The quotes emphasize the importance of procedural regularity, the burden of proof on the registrant to demonstrate eligibility for exemptions, and the limited scope of judicial review in these matters.
20. **Constitutional protections & criminal procedure.** The common legal principle unifying these quotes is the doctrine of proportionality and reasonableness in the context of legal sanctions, forfeitures, and procedural fairness. The quotes collectively emphasize the need for legal actions, such as forfeitures and penalties, to be proportionate to the offense’s gravity and for procedural decisions to be grounded in reasonableness and fairness. This includes considerations of probable cause, the impact of legal actions on individuals’ rights, and the necessity for judicial discretion to ensure just outcomes.
21. **Federal habeas corpus.** The common legal principle unifying these quotes is the distinct legal framework and procedural considerations governing military justice and its interaction with civilian legal principles. These quotes reflect the unique aspects of military law, such as the separation of military and civilian legal systems, the specific rights and procedures applicable to servicemembers, and the balance between military discipline and individual rights.
22. **Employment discrimination.** The common legal principle unifying these quotes is the doctrine of summary judgment and the standards for its application. Many quotes emphasize the requirement that there be “no genuine issue as to any material fact” and that the “moving party is entitled to judgment as a matter of law,” reflecting the procedural mechanism by which courts can resolve cases without a full trial when there is no dispute over the key facts.

23. **Federal statutory enforcement in financial and environmental regulatory contexts.** The common legal principle unifying these quotes is the emphasis on federal supremacy and the development of federal common law in areas where federal interests are predominant. This includes the application of federal law over state law in matters involving federal agencies like the FDIC, the need for uniformity in federal commercial transactions, and the federal courts' role in crafting rules in the absence of specific congressional guidance.
24. **Constitutional torts.** The common legal principle unifying these quotes is the doctrine of qualified immunity and the related concept of absolute immunity for government officials. This doctrine protects government officials from liability for civil damages as long as their conduct does not violate clearly established statutory or constitutional rights that a reasonable person would have known.
25. **Constitutional structure and federalism.** The common legal principle that unifies these quotes is the application and interpretation of equitable doctrines and principles in various legal contexts. This includes the necessity for written agreements in certain transactions, the role of equity in complex accounting and financial disputes, the equitable treatment of creditors in bankruptcy, the application of laches, and the equitable discretion of courts in matters such as contract reformation.

A.2 Citation Centrality Exhibits

Figure S1: Citation type composition over time



Notes: Lines plot the share of all citations over time, going to case law (blue), U.S. Code (orange), Code of Federal Regulations (green), and the Constitution (red), computed in 5-year bins from 1930 to 2020.

Table S1: Sources of Authority Across Legal Citation Communities

	Community	Cases	Case	USC	CFR	Const.
1	veterans benefits & VA administrative law	3,315	54.6%	29.3%	16.0%	0.2%
2	social security & administrative appeals	22,621	61.0%	22.4%	16.5%	0.0%
3	special education & IDEA litigation	2,491	63.1%	27.6%	9.1%	0.3%
4	sovereign immunity & jurisdictional structure for suing the US	24,095	69.6%	23.9%	6.2%	0.3%
5	administrative law & federal judicial review	114,585	69.6%	22.9%	6.4%	1.1%
6	bankruptcy law & creditor-debtor relations	80,040	69.9%	29.5%	0.4%	0.2%
7	immigration asylum & removal proceedings	48,579	70.2%	21.4%	8.4%	0.0%
8	federal statutory enforcement in financial and environmental regulatory contexts	4,538	71.7%	24.5%	3.6%	0.2%
9	pretrial detention & extradition	2,937	72.2%	26.8%	0.4%	0.6%
10	ERISA	30,832	74.4%	23.7%	1.7%	0.2%
11	labor arbitration & fair representation	33,347	77.6%	21.4%	0.9%	0.1%
12	constitutional protections criminal procedure	210,683	79.2%	20.0%	0.4%	0.4%
13	securities & antitrust	55,891	80.6%	16.8%	2.5%	0.1%
14	military draft obligations	3,024	81.0%	6.2%	12.6%	0.3%
15	federal habeas corpus	93,163	81.8%	17.6%	0.2%	0.4%
16	FOIA & discovery	25,986	81.8%	16.2%	1.8%	0.2%
17	federal tort claims act interpretation	46,090	82.4%	15.9%	1.6%	0.1%
18	employment discrimination	137,803	82.5%	15.1%	2.3%	0.1%
19	copyright & trademark	23,216	82.7%	16.5%	0.6%	0.2%
20	tax law (Internal Revenue Code)	60,235	82.8%	15.5%	1.5%	0.1%
21	customs/trade disputes	25,863	84.2%	13.5%	2.2%	0.0%
22	constitutional torts	100,823	84.8%	13.8%	0.8%	0.6%
23	personal jurisdiction & federal choice of law	32,178	85.1%	14.1%	0.5%	0.3%
24	constitutional structure and federalism	121,447	85.7%	13.3%	0.7%	0.3%
25	patent law	35,763	86.1%	12.6%	1.2%	0.1%

Notes: Share of unique cited authorities by type. Each authority counted once per citing case. Sorted by case citation share (ascending).

Table S2: Central statutes and introduction years by community

	Community	Central Statute	Year	Citing (%)
1	constitutional protections criminal procedure	18 U.S.C. §3553	1983	14.0%
2	federal habeas corpus	28 U.S.C. §2254	1948	33.5%
3	employment discrimination	42 U.S.C. §2000e	1963	12.4%
4	personal jurisdiction & federal choice of law	28 U.S.C. §1404	1948	16.4%
5	ERISA	29 U.S.C. §1132	1973	26.2%
6	constitutional torts	42 U.S.C. §1983	1979	51.4%
7	administrative law & federal judicial review	42 U.S.C. §1983	1979	12.3%
8	labor arbitration & fair representation	29 U.S.C. §158	1935	21.7%
9	bankruptcy law & creditor-debtor relations	11 U.S.C. §523	1977	15.3%
10	securities & antitrust	15 U.S.C. §1	1890	10.7%
11	FOIA & discovery	5 U.S.C. §552	1965	14.8%
12	constitutional structure and federalism	28 U.S.C. §41	1948	1.0%
13	patent law	35 U.S.C. §102	1952	11.0%
14	tax law (Internal Revenue Code)	26 U.S.C. §7422	1954	1.6%
15	immigration asylum & removal proceedings	8 U.S.C. §1252	1952	61.5%
16	copyright & trademark	15 U.S.C. §1125	1946	18.3%
17	social security & administrative appeals	42 U.S.C. §405	1935	45.9%
18	special education & IDEA litigation	20 U.S.C. §1415	1969	78.9%

Notes: Central statute per community, selected as the most-cited USC provision with a known introduction year. Year = year the statutory provision was first enacted. Citing (%) = share of cases in the community citing the central statute at least once. Communities without an identifiable central statute year are omitted.

Table S3: Regulatory Foundations of Legal Citation Communities

	Community	Top-5		Key Regulations
		Cases	Rate	
1	immigration asylum & removal proceedings	48,579	52%	8 C.F.R. §1003.2 (23%); 8 C.F.R. §208.13 (8%); 8 C.F.R. §208.16 (8%)
2	social security & administrative appeals	22,621	28%	20 C.F.R. §404.1520 (14%); 20 C.F.R. §404.1527 (7%); 20 C.F.R. §416.920 (3%)
3	ERISA	30,832	26%	29 C.F.R. §2560.503-1 (16%); 29 C.F.R. §2560.503 (3%); 29 C.F.R. §2510.3-1 (3%)
4	pretrial detention & extradition	2,937	25%	22 C.F.R. §95.2 (6%); 22 C.F.R. §95.3 (6%); 22 C.F.R. §95.4 (5%)
5	securities & antitrust	55,891	24%	17 C.F.R. §240.10b-5 (13%); 17 C.F.R. §10 (5%); 12 C.F.R. §226.8 (2%)
6	military draft obligations	3,024	23%	32 C.F.R. §1625.2 (10%); 32 C.F.R. §1631.7 (4%); 32 C.F.R. §456 (3%)
7	federal statutory enforcement in financial and environmental regulatory contexts	4,538	19%	40 C.F.R. §300.700 (6%); 40 C.F.R. §300.430 (4%); 40 C.F.R. §300.68 (4%)
8	patent law	35,763	17%	37 C.F.R. §1.56 (12%); 37 C.F.R. §1.175 (2%); 37 C.F.R. §1.131 (2%)
9	labor arbitration & fair representation	33,347	16%	29 C.F.R. §102.69 (6%); 29 C.F.R. §102.67 (3%); 29 C.F.R. §8 (3%)
10	employment discrimination	137,803	15%	29 C.F.R. §1630.2 (11%); 29 C.F.R. §1614.105 (2%); 29 C.F.R. §541.118 (1%)
11	federal habeas corpus	93,163	15%	28 C.F.R. §2.20 (5%); 28 C.F.R. §2.52 (4%); 28 C.F.R. §2.47 (2%)
12	copyright & trademark	23,216	15%	37 C.F.R. §202.1 (4%); 37 C.F.R. §202.3 (4%); 19 C.F.R. §133.21 (2%)
13	veterans benefits & VA administrative law	3,315	14%	38 C.F.R. §3.303 (3%); 38 C.F.R. §3.159 (3%); 38 C.F.R. §3.105 (3%)
14	bankruptcy law & creditor-debtor relations	80,040	13%	7 C.F.R. §46.46 (6%); 34 C.F.R. §685.209 (3%); 7 C.F.R. §46.2 (2%)
15	customs/trade disputes	25,863	12%	19 C.F.R. §177.10 (4%); 19 C.F.R. §177.9 (3%); 19 C.F.R. §177.12 (2%)
16	special education & IDEA litigation	2,491	10%	34 C.F.R. §1415 (2%); 34 C.F.R. §300.502 (2%); 34 C.F.R. §300.8 (2%)
17	federal tort claims act interpretation	46,090	7%	28 C.F.R. §14.2 (3%); 49 C.F.R. §1005.2 (1%); 28 C.F.R. §14.3 (1%)
18	FOIA & discovery	25,986	7%	28 C.F.R. §16.11 (3%); 28 C.F.R. §16.3 (1%); 45 C.F.R. §164.512 (1%)
19	personal jurisdiction & federal choice of law	32,178	6%	17 C.F.R. §240.10b-5 (3%); 31 C.F.R. §596.201 (1%); 31 C.F.R. §535.201 (1%)
20	constitutional torts	100,823	6%	12 C.F.R. §560.2 (2%); 12 C.F.R. §226.23 (1%); 26 C.F.R. §301.7433-1 (1%)
21	constitutional protections	210,683	6%	31 C.F.R. §103.22 (2%); 17 C.F.R. §240.10b-5 (1%); 36 C.F.R. §4.23 (1%)
22	criminal procedure	60,235	4%	26 C.F.R. §301.6402-2 (1%); 26 C.F.R. §301.6323 (1%); 26 C.F.R. §31.3121 (1%)
23	tax law (Internal Revenue Code)	24,095	4%	5 C.F.R. §1201.56 (1%); 48 C.F.R. §15.306 (1%); 5 C.F.R. §1201.115 (1%)
24	sovereign immunity & jurisdictional structure for suing the US	121,447	3%	25 C.F.R. §83.7 (1%); 19 C.F.R. §24.24 (1%); 25 C.F.R. §531.1 (1%)
25	constitutional structure and federalism	114,585	2%	40 C.F.R. §1508.27 (1%); 50 C.F.R. §402.14 (0%); 40 C.F.R. §1508.9 (0%)
	administrative law & federal judicial review			

Notes: All 25 whole-period Louvain communities, ranked by CFR centrality. **Top-5 Rate:** share of CFR citations going to the five most-cited provisions. Parenthetical percentages: each provision's share of total CFR citations within that community.

Table S4: Constitutional Provisions Cited by Legal Citation Communities

	Community	Cases	Top-5	Key Constitutional Provisions
			Rate	
1	veterans benefits & VA	3,315	85%	Art. 3 (31%); Amend. 5 (28%); Art. 1 (16%)
2	administrative law			
3	special education & IDEA litigation	2,491	79%	Amend. 14 (33%); Art. 3 (17%); Amend. 1 (10%)
4	patent law	35,763	78%	Art. 1, §8, cl. 8 (32%); Art. 3 (24%); Art. 1, §8 (11%)
5	social security & administrative appeals	22,621	73%	Amend. 5 (30%); Art. 3 (26%); Art. 4, §1 (8%)
6	constitutional protections	210,683	73%	Amend. 4 (31%); Amend. 5 (18%); Amend. 6 (18%)
7	criminal procedure			
8	customs/trade disputes	25,863	71%	Art. 1, §8 (31%); Art. 1, §9, cl. 5 (14%); Art. 1, §8, cl. 1 (10%)
9	federal habeas corpus	93,163	69%	Amend. 6 (36%); Amend. 5 (12%); Amend. 14 (12%)
10	immigration asylum & removal proceedings	48,579	68%	Amend. 5 (42%); Art. 3 (10%); Art. 1, §9, cl. 2 (7%)
11	copyright & trademark	23,216	67%	Art. 1, §8, cl. 8 (34%); Art. 1, §8 (11%); Art. 3 (10%)
12	federal statutory enforcement in financial and environmental regulatory contexts	4,538	65%	Art. 3 (28%); Amend. 5 (17%); Art. 6 (8%)
13	ERISA	30,832	62%	Art. 3 (39%); Art. 6 (8%); Art. 3, §2 (6%)
14	bankruptcy law & creditor-debtor relations	80,040	62%	Art. 3 (19%); Amend. 5 (12%); Amend. 7 (12%)
15	constitutional torts	100,823	60%	Amend. 14 (21%); Amend. 4 (14%); Amend. 1 (9%)
16	FOIA & discovery	25,986	60%	Amend. 5 (29%); Amend. 1 (10%); Art. 3 (9%)
17	sovereign immunity & jurisdictional structure for suing the US	24,095	60%	Amend. 5 (38%); Art. 3 (12%); Art. 1 (6%)
18	securities & antitrust	55,891	60%	Art. 3 (33%); Amend. 5 (8%); Amend. 14 (8%)
19	military draft obligations	3,024	59%	Amend. 13 (14%); Amend. 5 (14%); Art. 1, §8 (13%)
20	employment discrimination	137,803	58%	Amend. 14 (22%); Art. 3 (14%); Amend. 1 (9%)
21	personal jurisdiction & federal choice of law	32,178	57%	Amend. 14 (26%); Art. 3 (13%); Amend. 5 (8%)
22	labor arbitration & fair representation	33,347	55%	Art. 6 (16%); Art. 3 (16%); Art. 6, cl. 2 (11%)
23	federal tort claims act interpretation	46,090	50%	Art. 3, §2 (15%); Art. 3 (12%); Art. 6 (10%)
24	administrative law & federal judicial review	114,585	48%	Art. 3 (13%); Amend. 14 (11%); Amend. 1 (10%)
25	pretrial detention & extradition	2,937	46%	Art. 3 (19%); Art. 6 (8%); Art. 2 (7%)
26	tax law (Internal Revenue Code)	60,235	44%	Amend. 5 (14%); Amend. 16 (12%); Art. 1 (6%)
27	constitutional structure and federalism	121,447	40%	Art. 1, §8, cl. 3 (15%); Art. 1, §8 (7%); Art. 3 (6%)

Notes: All 25 whole-period Louvain communities, ranked by constitutional centrality. **Top-5 Rate:** share of constitutional citations going to the five most-cited provisions. Parenthetical percentages: each provision's share of total constitutional citations within that community.

A.3 Drivers of Persistent Legal Tradition

What predicts whether a newly formed doctrinal lineage endures? A formalist null hypothesis would posit that legal innovations can arise from any judge or court and spread through the system purely by virtue of their persuasiveness or analytic strength. By contrast, a more realist perspective—echoing Holmes Jr.’s view that “the life of the law has not been logic, it has been experience”—suggests that new legal arguments emerge from institutional structures or in response to broader social transformations [2].

From this perspective, persistence should depend not only on the content of legal arguments but also on how they enter the judicial system. One possibility is that influential, high-volume courts play a decisive role, so doctrines introduced in those venues endure. Another is that durable traditions emerge when many courts confront similar statutory or social pressures at once, producing parallel lines of reasoning that diffuse across the system. To distinguish between these mechanisms, we compare persistence driven by early concentration of influence among major courts with persistence driven by broad early diffusion across courts, a question central to debates about how different courts contribute to the evolution of federal doctrine [e.g. 3, 4, 5].

We work with the panel of 54 sliding-window Louvain communities. For each linked community cluster, we zoom in on the “starter” community as the first decade where it emerged – that is, communities with no incoming continuation link. For each starter community g we record its start and end year and define Persistence_g as its end year minus start year.

Next we operationalize the competing views of how legal persistence is established. First, let $\text{cases}_{c,g}$ be the number of cases from court c that fall in starter community g , representing the court’s contribution to the initial formation of that doctrinal community. Second, let $\text{cases}_c^{\text{total}}$ be the total number of cases written by court c in the federal corpus, measuring the court’s overall institutional importance in the federal judiciary. Define

$$\text{share}_{c,g}^{\text{comm}} = \frac{\text{cases}_{c,g}}{\sum_{c'} \text{cases}_{c',g}}, \quad \text{share}_c^{\text{total}} = \frac{\text{cases}_c^{\text{total}}}{\sum_{c'} \text{cases}_{c'}^{\text{total}}},$$

giving the initial prevalence of court c in community g , as well as the overall prevalence of court c in the broader federal caselaw corpus. Next, define a court’s *community weight* as

$$\text{weight}_{c,g} = \text{share}_{c,g}^{\text{comm}} \times \text{share}_c^{\text{total}},$$

which is large only if a court is both prominent in the community g and important in the overall system.

There are two explanatory variables of interest. The first, LargeCourts_g , measures the degree to which a new doctrinal community is dominated by high-volume courts at its inception. It is computed as the sum of $\text{weight}_{c,g}$ across all federal courts (excluding SCOTUS):

$$\text{LargeCourts}_g = \sum_{c \in \mathcal{C}_{\text{nonSC}}} \text{weight}_{c,g}.$$

Higher values indicate that a small number of influential courts contribute disproportionately to the formation of the community.

The second explanatory variable, Breadth_g , captures how broadly the community is distributed across courts. It is measured as the number of distinct courts whose initial case count in community g meets a minimum threshold $\bar{n}$:

$$\text{Breadth}_g = |\{c : \text{cases}_{c,g} \geq \bar{n}\}|.$$

By default we set $\bar{n} = 1$, so we count every court contributing at least one opinion to the starter community (robustness $\bar{n}$ in Figure S2b). Higher values of Breadth_g indicate that a new doctrinal community draws early participation from a wider set of courts.

The baseline OLS model is

$$\text{Persistence}_g = \beta_1 \text{LargeCourts}_g + \beta_2 \text{Breadth}_g + \beta_3 \mathbf{1}\{\text{start}_g < 1900\} + \varepsilon_g, \quad (1)$$

where $\mathbf{1}\{\text{start}_g < 1900\}$ is an indicator for communities originating before 1900. We report heteroskedasticity-robust (HC3) standard errors when constructing confidence intervals.

Together, LargeCourts_g and Breadth_g operationalize two competing mechanisms of persistence—one driven by institutional dominance and the other by broad diffusion. The coefficients of interest are β_1 and β_2 . A positive β_1 indicates that early dominance by large, high-volume courts predicts longer-lived doctrinal lineages, while a negative value would indicate that such dominance is associated with shorter-lived communities. A positive β_2 suggests that broader early participation across courts promotes persistence, whereas a negative value would indicate that wider initial participation is associated with shorter-lived communities.

This regression is unlikely to capture causal effects of institutional features on the persistence of law. While a number of unobserved factors may influence both the initial composition of a community and its eventual persistence, we can address two particularly salient sources of confounding. First, larger communities may naturally persist longer simply because they encompass more cases and legal material. To account for this, we compute each starter community’s initial size—the total number of opinions it contains at inception—and include this *starter size* as an additional control. Second, the influence of the Supreme Court may independently affect persistence, as doctrines originating in or heavily cited by the Court could be reinforced through vertical authority rather than diffusion. We therefore compute the *Supreme Court share* as the sum of $\text{weight}_{c,g}$ over Supreme Court cases and include it as an additional control.

Figure S2 reports coefficient estimates and 95% confidence intervals for the three models: baseline, court size controls, and court size controls + Supreme Court influence controls. We see, first, that initial breadth matters: the coefficient on Breadth_g is positive and statistically significant. Lineages whose initial window draws participation from more courts tend to persist longer. Second, the coefficient on LargeCourts_g is null when included alongside the breadth measure. Early dominance by large producers of case law does not, on its own, forecast longevity. Both of these results hold across specifications. Controlling for the size of the initial community, or the influence of the Supreme Court, does not change the result

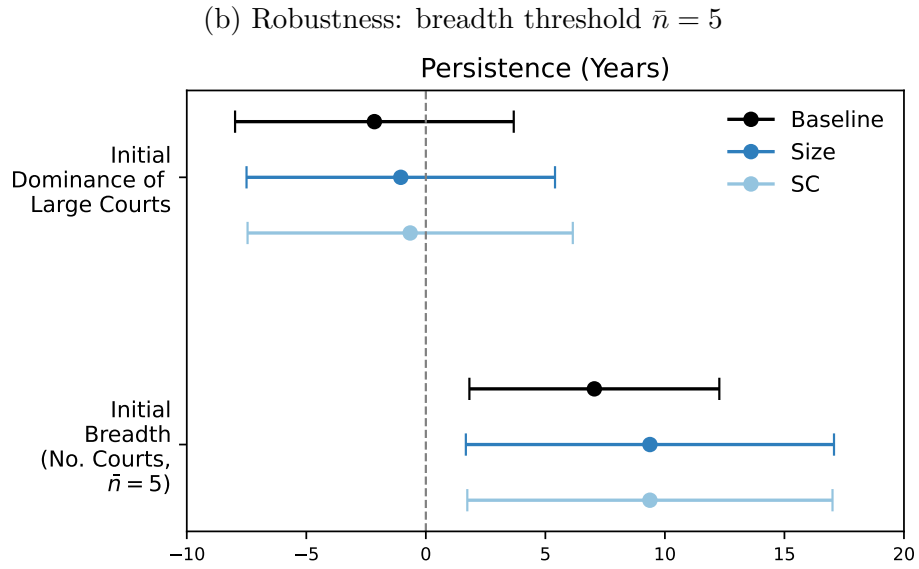
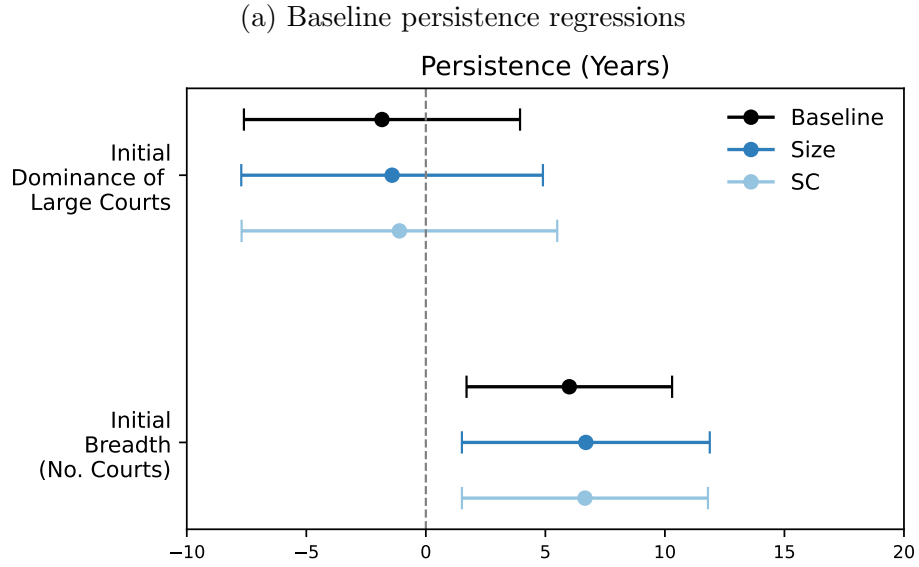


Figure S2: Determinants of Persistence

Notes: OLS coefficient estimates with 95% confidence intervals (HC3 robust standard errors). The dependent variable is community persistence (end year – start year). Panel (a) shows three specifications: baseline (with pre-1900 indicator), adding starter community size control, and adding Supreme Court share control. Panel (b) repeats the analysis using an alternative breadth measure in which $Breadth_g$ counts only courts contributing ≥ 5 opinions to the starter community. The unit of observation is a starter community (the first decade in which a linked community cluster appears).

that broader initial court participation leads to higher persistence.

These results add evidence on the underlying mechanics of persistence across legal Louvain communities. We find that precedent tends to preserve what the system backs broadly, not what a few influential venues push early. Durable innovations start with many lower-court authors adopting a shared legal approach, then accrete authority over time. This complements our evolution results: judges keep core lines where settlement is valuable, yet rebuild doctrine when social and statutory change demands it. The rebuilding that endures tends to be *national* in scope from the outset, reaching across many courts rather than emanating from a single venue.

The view that lasting doctrinal change must originate in the Supreme Court finds little support here. In the evolution of doctrine, the Supreme Court plays an amplifying role in addition to a generative one. While landmark opinions can consolidate or reframe existing lines of reasoning, new, durable traditions have to be established broadly in the lower courts as well.

Taken together, these results challenge a “formalist” view of legal evolution in which persuasive arguments arise at random and spread purely through their intrinsic intellectual force. Common law is an adaptive, experience-driven system responsive to social and institutional conditions, rather than the projection of any single tribunal’s internal logic [2]. Doctrinal innovations that secure early, broad footing across many courts, rather than originating in a single influential venue, are far more likely to persist.

References

- [1] Robert Mahari, Dominik Stammach, Elliott Ash, and Alex Pentland. Leopard: A large-scale dataset of judicial citations to precedent. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pages 9863–9877. Association for Computational Linguistics, 2024.
- [2] Oliver Wendell Holmes Jr. *The common law*. Routledge, 1881.
- [3] Hersel W Perry. *Deciding to decide: agenda setting in the United States Supreme Court*. Harvard University Press, 1991.
- [4] Tom S Clark and Jonathan P Kastellec. The supreme court and percolation in the lower courts: an optimal stopping model. *The Journal of Politics*, 75(1):150–168, 2013.
- [5] Seth Davis and Michael Coenen. Percolation’s value. *Stanford Law Review*, 73, 2021.