THE PHYSICS OF THE LAW: LEGAL SYSTEMS THROUGH THE PRISM OF COMPLEXITY SCIENCE

EDITED BY: Pierpaolo Vivo, Daniel Martin Katz and J. B. Ruhl
PUBLISHED IN: Frontiers in Physics

frontiers Research Topics

ISSN 1664-8714
ISBN 978-2-88976-129-6
DOI 10.3389/978-2-88976-129-6

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

THE PHYSICS OF THE LAW: LEGAL SYSTEMS THROUGH THE PRISM OF COMPLEXITY SCIENCE

Topic Editors:
Pierpaolo Vivo, King's College London, United Kingdom
Daniel Martin Katz, Illinois Institute of Technology, United States
J. B. Ruhl, Vanderbilt University, United States

Citation: Vivo, P., Katz, D. M., Ruhl, J. B., eds. (2022). The Physics of the Law: Legal Systems Through the Prism of Complexity Science. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88976-129-6

Table of Contents

frontiers in Physics

ORIGINAL RESEARCH
published: 31 March 2021
doi: 10.3389/fphy.2020.625241

Check for updates

A Multinetwork and Machine Learning Examination of Structure and Content in the United States Code

Keith Carlson[1], Faraz Dadgostari[2], Michael A. Livermore[3]* and Daniel N. Rockmore[1,4]

[1]Department of Computer Science, Dartmouth College, Hanover, NH, United States, [2]Department of Systems Engineering, University of Virginia, Charlottesville, VA, United States, [3]School of Law, University of Virginia, Charlottesville, VA, United States, [4]The Santa Fe Institute, Santa Fe, NM, United States

This paper introduces a novel linked structure-content representation of federal statutory law in the United States and analyzes and quantifies its structure using tools and concepts drawn from network analysis and complexity studies. The organizational component of our representation is based on the explicit hierarchical organization within the United States Code (USC) as well an embedded cross-reference citation network. We couple this structure with a layer of content-based similarity derived from the application of a “topic model” to the USC. The resulting representation is the first that explicitly models the USC as a “multinetwork” or “multilayered network” incorporating hierarchical structure, cross-references, and content. We report several novel descriptive statistics of this multinetwork. These include the results of this first application of the machine learning technique of topic modeling to the USC as well as multiple measures articulating the relationships between the organizational and content network layers. We find a high degree of assortativity of “titles” (the highest level hierarchy within the USC) with related topics. We also present a link prediction task and show that machine learning techniques are able to recover information about structure from content. Success in this prediction task has a natural interpretation as indicating a form of mutual information. We connect the relational findings between organization and content to a measure of “ease of search” in this large hyperlinked document that has implications for the ways in which the structure of the USC supports (or doesn’t support) broad useful access to the law. The measures developed in this paper have the potential to enable comparative work in the study of statutory networks that ranges across time and geography.

Keywords: multinetwork, statutory network, United States Code, topic modeling, assortativity, law search

OPEN ACCESS

Edited by:
Daniel Martin Katz,
Illinois Institute of Technology,
United States

Reviewed by:
Satyam Mukherjee,
Indian Institute of Management
Udaipur, India
Jie Cao,
Nanjing University of Finance and
Economics, China

****Correspondence:***
Michael A. Livermore
mlivermore@virginia.edu

Specialty section:
This article was submitted to
Social Physics,
a section of the journal
Frontiers in Physics

Received: *02 November 2020*
Accepted: *30 December 2020*
Published: *31 March 2021*

Citation:
Carlson K, Dadgostari F, Livermore MA and Rockmore DN (2021) A Multinetwork and Machine Learning Examination of Structure and Content in the United States Code. Front. Phys. 8:625241. doi: 10.3389/fphy.2020.625241

1 INTRODUCTION

In this paper we present a network-based framing and analysis of the United States Code (USC), the legal corpus comprising the federal statutes of the United States. Statutes therein possess hierarchical structure via their organization into titles, sections, sub-sections and the like, and a cross-reference structure in which provisions of a code cite to other provisions for purposes of sharing definitions or establishing legal relations. These overlapping structures of hierarchy and cross-references intertwine content which can also be studied as a network through a similarity structure derived from its defining documents. Taken together, these interleaved network structures define the USC – and more generally, any statutory network corpus – as a *mutilayered network* or *multinetwork*, a complex network structure increasingly of interest in the areas of network science and complex systems [1]. The contribution of this work is the first framing of the USC as a multinetwork and with that, a first network-based analysis of the USC that includes the calculation of various network-related metrics (degree, betweeness centrality, hubs and authorities measures) and other quantifiable characteristics of the USC as well as detailing the relationships between the layers. In doing so, we present a defining framework for the notion of a "statutory network" and a collection of attendant measures that will enable future work in comparative and intrinsic statutory analysis.

As discrete, information dense, and legally important corpora, bodies of statutes have proven particularly attractive to researchers who are interested in applying computational tools to legal texts [2]. The hierarchical and networked structure we see in the USC is a hallmark of complex systems, which include regulatory structures as well as evolved organizational structures ranging from corporations to societies and ecosystems (*see* e.g., [3, 4]).

As mentioned, network characteristics, as well as the large body of work in quantitative textual analysis of document corpora enable an empirical approach to the study of statutory networks. This puts the study of statutes squarely within the vibrant body of research that is bringing to bear the tools of complex systems and machine learning on legal documents and the law (*see* [5] for a recent survey). For example, it may be that there are temporal or geographic determinants of a statutory network that may influence the diffusion of legal culture over space and time [6]. There may also be consequences of certain statutory network features, such as regulatory complexity, that increase legal transaction costs [7] or lead to hidden "cumulative" costs of regulation [8]. In this vein, new measures of statutory complexity are needed to move the field beyond fairly primitive proxies, such as word counts or simple *n*-gram style metrics (see e.g., [9, 10], as well as [11] for analogous efforts in the banking industry).

Network features of the law can also be exploited to study certain types of legal behavior. Prominent in this regard is the body of work that has been done related to the citation network of Supreme Court opinions, notably the groundbreaking work of Fowler and his collaborators in their study of precedent [12]. The use of text analysis tools in the law is more recent. Early uses of the machine learning *topic modeling* approach used herein include novel analyses of the impact of Supreme Court opinions on litigation [13], the influence of clerks on opinion-writing [14] and a more general study of opinions as genre [5]. Extensions to the much larger corpus of Circuit Court opinions has resulted in a new revelation of publication bias in that setting [15].

The multinetwork framework of SCOTUS opinions combining citations with topic-similarity has further served to produce a geometric framework for their study [16] and is used to study the problem of "law search" [17]. This work provides useful insights into an important category of legal behavior that has been difficult to study using traditional tools. Indeed, as far back as Jeremy Bentham, legal philosophers have recognized that the diffuse nature of the law (especially in common law systems) poses important normative problems [18]. Given the growth of publicly available legal datasets[1], advances in understanding the internal organization of the law and using that understanding to facilitate search of legal materials for non-experts and experts alike may also facilitate broader access to the law in a comprehensible format. These questions are similarly germane to the USC.

This paper is the first to merge information on statutory structure with the semantic content of statutory text. We thus build and expand on prior work, such as [2, 19]; that focuses on organizational features (specifically hierarchical structure and cross-references) of statutory codes. The inclusion of semantic content is particularly important: while the hierarchical and cross-reference structure of a statutory corpus provides some valuable information about the nature of a legal regime, the structure itself has no legal effect. Statutory structure establishes relationships and order, but it is the semantic content that ultimately is the legal *materiel* stitched together through structure. A better understanding of the relationship between statutory structure and content may also prove particularly useful in developing better search tools for statutory law. Because statutory language is relatively parsimonious—especially compared to the lengthy narrative documents produced by courts—common search approaches (including Boolean searches and various natural language augmentations) can be ineffective at identifying relevant statutory authority for a given legal matter. Systematic patterns in the overlap of structure and substance could be leveraged in search tools that were specifically designed to lower the costs of identifying relevant statutory text. While we focus on the example of the United States Code, the techniques described here could be applied to any statutory system. Empirical comparative extensions are likely to yield particularly worthwhile insights into different system-level characteristics of legal orders and the relationship of those characteristics to outcomes of interest, which could vary from sociological legitimacy to regulatory compliance costs.

For our analysis of the semantic content, we rely on the machine learning tool of *topic modeling* [20] and specifically the *structural topic model* (STM) introduced in [21]. A great deal of work has gone into the development and refinement of topic models and they have become widespread within social sciences (e.g., [22],) the humanities (e.g., [23–26],) and other text-centric disciplines. Several recent papers have applied topic models to legal documents [5, 27–29]. The STM approach builds on the conventional and widely used Latent Dirichlet Allocation (LDA) topic model.

In **Section 2** we discuss prior work and summarize the contributions of our analysis. In **Section 3** we describe our data derived from the current (online) version of the USC. This includes some basic descriptive statistics of the structural network of the USC including centrality and hub/authority measures of the underlying title network. **Section 4** contains the meat of the analysis. After providing some additional detail on

[1]See e.g., Court Listener https://www.courtlistener.com/.

topic models and their utility in producing basic characterizations of unstructured collections of documents, we report the results of our topic model of the USC as well as some intuitive descriptions of the relationships between topics and statutory structure. The topic modeling of the USC text is new and a main contribution of this work. Our next main contributions build on the results of the topic modeling to generate a set of measures to examine the relationship between semantic features of the USC and its structure, the latter represented in its cross-references and hierarchical organization into higher-level titles. We do this first by using an *assortativity* measure that connects cross-reference and semantic structure (using the topic model output). This shows significant relationship between title and content. We then construct new relational measures inspired by mutual information that investigate the degree to which semantic content can predict connectivity. Using an SVM machine learning approach, we achieve predictive accuracy of 60% in using topic proportions to predict titles. We create a second measure using the law search model developed in [16, 17] to predict cross-reference citations from topics and achieve similar performance as has been achieved for an analogous experiment using the corpus of Supreme Court opinions. Taken together, these results suggest good alignment between the structural and content layers of the multinetwork, with attendant positive implications for searchability. We close in **Section 5** with a description of future work enabled by an anticipated merging of the USC data with other legal corpora.

2 GENERAL CHARACTERISTICS OF LEGAL SYSTEMS

Statutes are the laws enacted by a legislative body. In common law jurisdictions such as the United Kingdom, statutory law (e.g., the enactments of Parliament) can be contrasted with judge-made law that accretes through the decisions of courts in individual cases. In the United States, most areas of federal and state law have statutes at their foundation, with courts charged with the task of statutory interpretation through the application of statutory language to particular cases. Administrative agencies, themselves established and empowered via statutes, frequently have a role in elucidating broad statutory commands through more detailed regulations.

Statutes are distinct from constitutions, which are adopted and altered through special procedures rather than the typical legislative process. Taking the United States as an example, Article I, **Section 7** of the U.S. Constitution sets out the procedure for Congress to adopt statutes, via majority vote in the House of Representatives and Senate and presentment to the President, and in the case of a presidential veto, a two-thirds vote in both chambers. Article V of the Constitution describes the (very difficult) procedure for amending the Constitution itself, which requires that any proposed amendment be ratified by three fourths of the states.

Statutes are a longstanding object of analysis for empirical legal study: for example, there are a substantial number of papers that examine the effects of the death penalty—a statutory provision—on crime (e.g., [30].) In addition to investigating the consequences of particular policy choices as embodied in statutes, empirical legal scholars also examine factors that affect the decision of whether or not to adopt a new law, such as prior success in another jurisdiction or geographic proximity to other adopters [31].

In addition to studying *specific* legislative enactments, scholars have also focused on certain *general* characteristics of legal systems. For example, scholars in the “legal origins” tradition have argued that certain legal characteristics that are correlated with whether a country has a common law or civil law system are associated with macro-social outcomes such as economic development [32, 33]. This literature has been broadly influential and has shaped recent political discourse on law and development [34]. For criticism of this work, see [35, 36] among others.

A related literature focuses specifically on the notion of *legal complexity* and the question of relationships between the complexity of a legal system and a variety of societal outcomes [37, 38]. At a high level of abstraction [39] argues that societal complexity along a variety of fronts eventually contributes to the disintegration of politically organized groups. More concretely, scholars have argued that legal complexity hampers economic development through several channels, including by the lowering of returns on capital and thus impeding innovation [40–43]. In recent years, several scholars have attempted to use data on aggregate regulatory levels to draw conclusions about the costs and benefits of various regulatory regimes [9, 44, 45]. In the realm of political discourse, trade associations representing regulated industry frequently bemoan legal complexity and the cumulative cost of regulations [46].

Rigorous work on legal complexity has been hampered by inadequate definition and measurement of the underlying concept (*see* e.g., [37].) Simple measures, such as the number of pages in the U.S. Federal Register, or counting *n*-grams that target “command” type language (i.e., “shall” or “must”) have been used as rough proxies [45], but their shortfalls are fairly obvious. As domestic legal regimes cope with continued economic growth and global integration, legal complexity (broadly understood) is likely to increase, and social scientific study of this phenomenon will take on even greater importance. But that work will continue to struggle without reliable and accurate measures of the phenomenon.

A separate and related vein of scholarship examines the practice of *law search* – the process whereby agents seek out relevant legal authority to apply to a given legal problem [16, 17]. Other things being equal, a legal system in which it is more difficult to locate relevant authority can be understood as less comprehensible in way that is often attributed to its greater “complexity.” In short, a legal system of greater complexity is simply one in which it is more difficult for legal actors of any level of background to learn, know, and understand their rights and responsibilities. In [17]; the authors introduce the notion of “convergence” in a legal system, which is the tendency for legal participants to *converge* via law search to a similar set of legal authorities that are relevant to a given legal question. Convergence may be inversely associated with complexity.

When possible, quantitative measures – possibly tied to the field of complexity science – have the promise of making rigorous discussions of legal complexity. In particular, when a network framework makes sense, the tools of network science can be brought to bear effectively on the subject [37] and similarly for computational text analysis [17]. Both of these approaches are germane in the case of the study of the USC and statutory networks generally. The work described in this paper contributes to the literature of legal complexity and law search, and more generally to the field of computational analysis of legal texts (and statutes specifically). The work presented herein is the first project that we are aware of that combines information on statutory structure with semantic data on legal content to study the relationship between structure and substance in a statute-based legal order. Our findings regarding the United States Code can help set the stage for comparative work that examines similar relationships within other legal systems. In addition, the techniques we use to capture semantic content and overlay content and structure (and in particular the use of topic models) can inform future efforts to define measures of legal complexity. Finally, prior efforts to study law search have focused on judicial opinions, which are particularly information-dense, and therefore relatively easy to navigate. Our work can be leveraged to expand those analyses to statutory (and regulatory) texts, where the costs of law search may be even more pronounced.

3 DATA

As per the U.S. Constitution, the fundamental role of Congress is to exercise the "legislative power," much of which is embodied in the statutes it adopts. Statutes are simply the laws adopted by Congress. These laws cover a wide range of public and private conduct – everything from the tax rate and the penalties for kidnapping to provisions establishing the authority of the Environmental Protection Agency to issue air quality standards.

When a statute is successfully adopted (e.g., via majority votes in both houses and a presidential signature), it is issued as a Public Law and published as a session law, and is compiled chronologically in the Statutes at Large.[2] These session laws are the exact text enacted by Congress, and so represent "the law" as a direct exercise of Congress's power.

These public laws can be difficult to navigate. They are not organized by subject matter. Furthermore, rather than evolving through edits, subsequent revisions are recorded as separate session laws which then operate on earlier versions. This in turn creates a complex relation of interlocking dependencies and cross-references. For several decades after the founding of the republic, the work of compiling and publishing a comprehensive representation of the current law fell to private publishers. These documents were useful for lawyers, but had no official legal status. In the 1870s, Congress undertook an official codification, the Revised Statues of the United States, which was meant to capture the existing state of the law. Subsequent efforts at official codification faltered until the USC was approved by Congress in 1926.[3] Eventually the maintenance of the USC was brought into the U. S. government and now the USC (or "the Code") is maintained and published by the U.S. Office of the Law Revision Counsel (OLRC).[4] The OLRC is also responsible for the organization of the Code and compiling relevant changes as they are enacted.[5]

The representation used in the following analyses is based on a one-time "snapshot" of the Code based on the information published by the OLRC. Data collection occurred during the period October 2016 through January 2017. For this project, we do not examine dynamic over-time effects as the USC is altered through the legislative process. Instead, we focus on static features of the Code as it existed during the data-collection period.

As will be discussed in more detail in the next section, the topic model approach we use to engage in semantic analysis relies on our treating the USC as a corpus of "documents." The quotation marks call out the distinction between a "document" in the sense of a topic model and a document in a more colloquial sense. The former is a contiguous collection of words (or even more generally, character strings extracted via a standard sort of text processing format) while the latter might suggest something with some recognizable narrative form. The USC is organized in a nested fashion with multiple levels wherein *section* is the core organizational unit, containing thematic blocks of text.[6] We treat these sections as the documents for topic modelling purposes.

The structure of the USC as a hierarchical information repository is somewhat complicated, involving different levels across the Code. These include title, (possible) subtitle, chapter, parts and subparts, subsections, paragraphs, clauses, and items. The highest (i.e., broadest) organizational unit is the "title." Titles can be thought of as a general subject heading: examples include "Armed Forces" (Title 10) and "Public Health and Welfare" (Title 42). We use only sections and titles as the relevant unit of analysis and ignore the other levels, in part because they are treated differently depending on the title. With this definition of the document unit there are 41,138 total documents in the USC organized into 44 titles.

[2]https://www.loc.gov/law/help/statutes-at-large/

[3]Pub.L. 69–441, 44 Stat. 778, enacted June 30, 1926.

[4]https://uscode.house.gov/

[5]There are some additional complications concerning how the Code is compiled from the Statutes at Large and how courts give effect to various texts, but they are not important for this study. For more background on the Code, see [52].

[6]For example, **Section 112** of Title 42 is "Removal of revenue officers from port during epidemic" and states "Whenever, by the prevalence of any contagious or epidemic disease in or near the place by law established as the port of entry for any collection district, it becomes dangerous or inconvenient for the officers of the revenue employed therein to continue the discharge of their respective offices at such port, the Secretary of the Treasury, or, in his absence, the Undersecretary of the Treasury, may direct the removal of the officers of the revenue from such port to any other more convenient place, within, or as near as may be to, such collection district. And at such place such officers may exercise the same powers, and shall be liable to the same duties, according to existing circumstances, as in the port or district established by law. Public notice of any such removal shall be given as soon as may be."

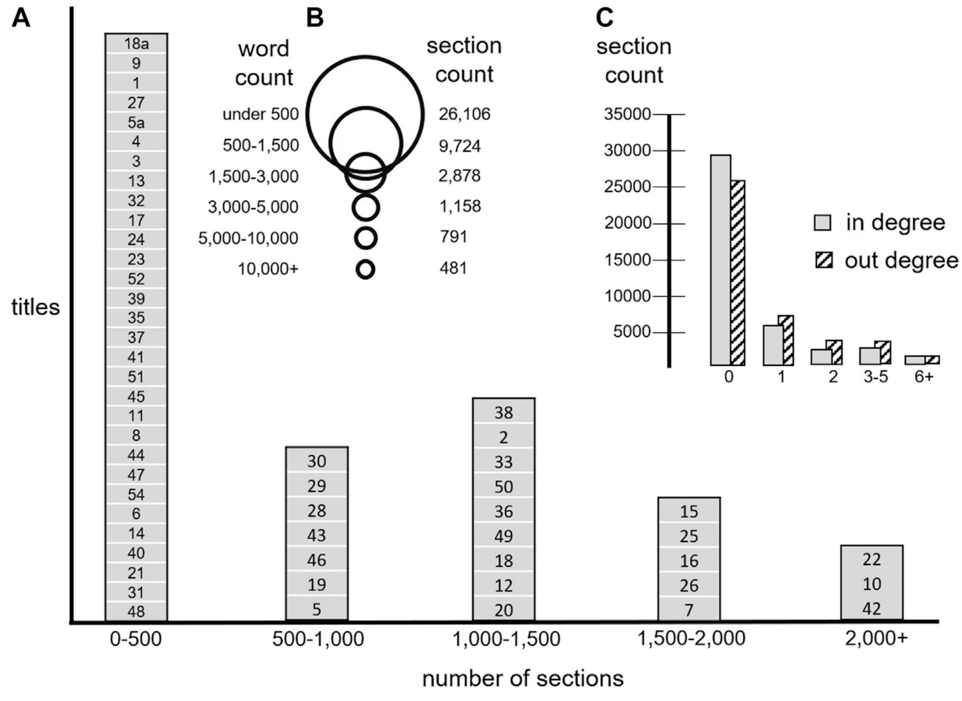


FIGURE 1 | The number of sections in each title varies considerably, and is right skewed. Panel **(A)** arranges each title from least to most sections and places the titles into four buckets. Title 18a has the smallest number of sections (17) and title 42 has the largest (4,581). There is also considerable variation in the word counts in each section, with a large number of fairly small sections (under 500 words) and a small number of very long sections (over 10,000 words). Panel **(B)** illustrates this variation. Panel **(C)** provides histograms of in-degree and out-degree for cross-references at the section level.

Figure 1 reports basic information on the word counts within sections and the counts of sections within titles.[7] There is considerable variation in the section count of the titles and the word count of the sections. Document length is right-skewed, with a modal length less than 500 words and a small number of much longer outliers. The average number of words per document is 1,036. There is a right-skewed distribution for section within titles as well, with many titles containing only a few hundred or even just several dozen sections. A second common size of around one to two thousand sections also is found. There are a few very large outliers, with Title 42 ("Public Health and Welfare") the largest.

Our data also includes information on cross-references within the Code. These section-to-section links (along with the sections) produce the (second) inherent network structure of the USC. A directed edge is created if one document (i.e., section) references another. **Figure 1** also reports histograms of the number of in- and out-citations by section, showing that most sections have neither incoming nor outgoing edges, and few have greater than six (either incoming or outgoing).

There may be multiple citations between documents, but in our representation, a single edge is constructed between two documents if there is one or more citations. We do not construct edges for citations above the section level (e.g., when a citation is to an entire title or chapter). With these caveats in place there are 38,399 cross-references between sections. Note that in the citation network, there are fewer edges than there are nodes. In the USC, there are a large number of sections that do not include cross-references to any other sections, and are not cross-referenced by any other section. This is not necessarily surprising. Many sections are self-contained and do not need to make reference to other sections for shared definitions or other purposes. Likewise, many sections provide no more general terms that must be referenced elsewhere. In addition, there are other relations that exist between statutory sections that may not be called out via cross-reference. For example, the location of a section within a chapter or other supra-section category may be meaningful and denote certain types or relationships. Some sections include intra-section cross-references. These show up as loop edges in the citation network. As will be discussed later in this paper, there is a substantial amount of overlap between the cross-reference network and the hierarchical structure of the USC, with many cross-references occurring within title. However, there is a fair amount of inter-title cross-referencing as well.

We calculate several additional network statistics on the basis of the analysis above, including graph density, average total degree, variance of in-degree and out-degree, and the ratio of the number of components over the number of nodes. They are reported in the Appendix in **Supplementary Table S1**.

[7]A list of title names is provided in the Appendix in **Table A7**.

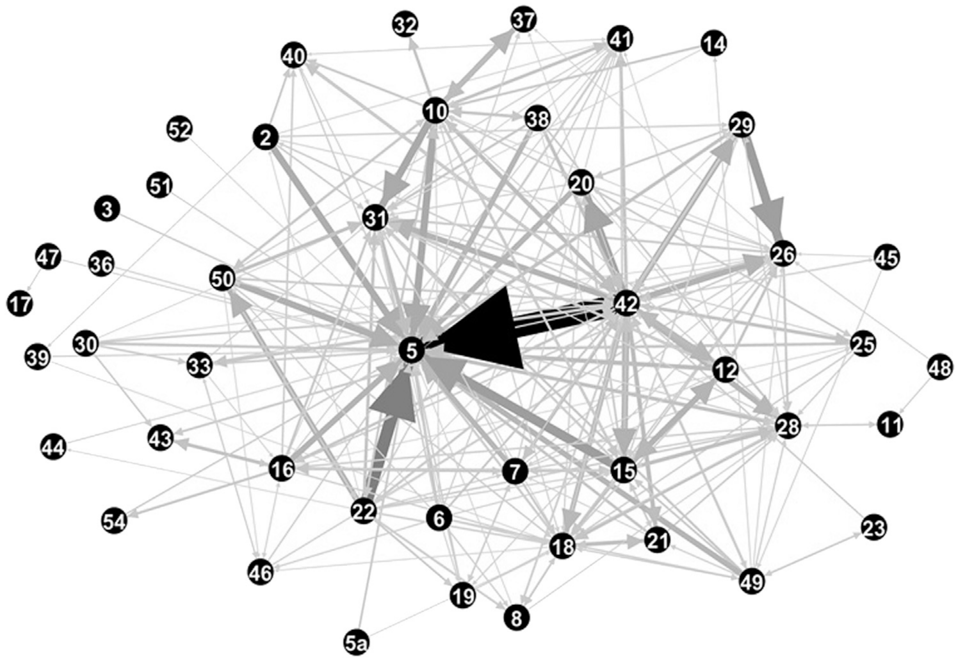


FIGURE 2 | Network representation of cross-references at the title level. Title 42 "Public Health and Welfare" has the largest number of outgoing citations, and Title 5 "Government Organization and Employees" has the most incoming citations.

Figure 2 provides an illustration of the citation network, highlighting only inter-title references. The edge thickness reflects the number of inter-title cross-references. As is visually apparent some titles play a more central role in the inter-title network than others and some titles are largely self-contained.[8] The most obvious pattern found in this figure is that Title 5, "Government Organization and Employees," has a very large number of incoming citations—presumably because it contains language or arrangements that are quite general and apply across legal categories—whereas the massive and hodgepodge Title 42 ("Public Health and Welfare") has a large number of outgoing citations, presumably because of its size and catchall nature.

To quantitatively validate these insights we calculate eigenvector centrality, betweenness centrality, hub scores, and authority scores [47]. These measures are all run on the network described above and pictured in **Figure 2**. This is a directed network with USC titles as nodes and edges with weight (thickness) equal (and thus proportional to the thickness of the edge in the figure) to the number of sections in one title which cite sections in another title. These results are reported in the Appendix in **Supplementary Table S2–S5**.

These more formal measures of centrality largely confirm the visual impression from **Figure 2**. Title 42 has by far the highest Hub Score: twice as large as the next highest, which is Title 22 ("Foreign Relations and Intercourse"), another subject with deep connections to the rest of the Code. Title 5 has the highest Authority Score, more than three times larger than that of the second highest. The Betweenness Centrality estimates highlight the importance of two additional Titles beyond 42 and 5, which are Titles 26 ("Internal Revenue Code") and 18 ("Crimes and Criminal Procedure"). Given the importance of taxes and crime in the life of the law, substantial connections between these Titles and other parts of the law is unsurprising. The Eigenvector Centrality calculation again emphasizes the centrality of Titles 42 and 5; the other two Titles with large centrality values are 31 ("Money and Finance") and 28 ("Judiciary and Judicial Procedure"). Title 31 deals with matters such as budgeting and procurement that are cross-governmental in nature. It is also natural that the Title that deals most directly with courts connects to various other areas of the law.

4 TOPIC MODELING THE USC

Network structure is just the skeleton of the USC. The semantic content provides the meat of the law. To engage in meaningful analysis of the semantic content of statutory texts, we rely on the method of *topic modeling*, which is well suited to constructing information-rich but low-dimensional representations of large textual corpora. The following section provides a short overview of the topic modeling technique and discusses the results of a topic model applied to the USC.

Loosely speaking, a topic model is a machine learning technique that produces a description of any document in a corpus as a probability distribution (weighted sum) of a fixed (and derived) set of "topics," which should be interpreted as akin

[8]Note that this graph is based on raw count numbers of inter-title cross-references (following the procedure of counting a max of only one edge between two documents) and is not normalized by the number of documents or words in a title.

to subject matter categories [22]. Formally, a *topic* is itself a probability distribution over the vocabulary of the corpus.[9] So that in fact, a document – which in any given corpus is an *a priori* determined contiguous set of words that generally respects textual boundaries such as paragraph ends, etc. – ultimately is represented as a distribution of distributions on the vocabulary. Topics are inferred from the corpus on the basis of a set of assumptions concerning a particular kind of parametrized generative model of document construction. For a standard topic model, the parameters of interest are typically restricted to the topic-word distributions (which describe the association between topics and words) and the document-topic distributions (which describe, for each document, the probability of finding words associated with each topic). The Latent Dirichlet Allocation (LDA) model is a common prior placed on the distributions. Standard topic modeling is supervised in the sense that the number of topics is specified *a priori* rather than discovered (e.g., according to some notion of parsimonious representation). *See* [48] for a good general introduction.

We have produced a topic model representing the USC with 100 topics. A list representing each topic by its five most heavily weighted words (or word stems) is presented in the Appendix in **Supplementary Table S6**. This gives a sense of the natural subject matter category that best corresponds to a given topic. Many of the topics appear to be legally meaningful in the sense that the most heavily weighted words suggest a recognizable theme. For example, the most heavily weighted words in Topic 26 are "bank, institut, financi, feder, insur, credit." This cluster of words conforms to a legal category of banking regulation. The most heavily weighted words in Topic 99 are "educ, school, student, institut, agenc," which conform to education. Some of the topics, on the other hand, do not match substantive areas but appear to be fairly generic collections of lawmaking words: these include Topic 1 ("transfer, section, titl, codif, former") and Topic 81 ("subchapt, part, titl, section, purpose"). The prevalence of a topic in a document captures (in a useful sense) the degree to which the document is "about" each topic. More formally, the document level distribution is a latent variable that is a best fit with the observed words in the document, given a set of topic distributions. For purposes of analysis, topic prevalence is a measure of the associated semantic content of each document.

Topic prevalence can capture subtleties that would be otherwise difficult to quantitatively describe. A statutory text that could be hand-coded by a researcher might be categorized according to some set of legal subject matter categories, such as a criminal law issue or environmental law. But such issue categorizations are binary and fail to capture the mix of topics that might be present in a document. Topic prevalence, by contrast, is a set of continuous variables (i.e., representing shares for each topic) that characterize each document. *Structured topic models* (STMs) are a class of topic model that builds on this basic architecture and has been described in the peer reviewed political science literature [21].[10]

4.1 Topics and Titles

Given the description of the documents (sections) according to topic distributions, we can derive a notion of similarity between documents and with that, another kind of edge connecting document to document. In this analysis we are interested in understanding the relationship between textual similarity and cross-referencing. We first offer some intuitive illustrations. We then estimate measures of *assortativity* for the document network and consider a relational measure between content and structure inspired by the notion of *mutual information* This is based on the predictive success of algorithms using one source of data (such as topic distributions) to predict the other (such as title or cross-references).

Supplementary Table S7 (in the Appendix) provides an intuitive sense of the substantive overlap between titles and topics by showing the topic that is most closely associated with each title (as estimated via the topic share of the documents within each title).[11] Note that the numbering of the topics is arbitrary. There is a very substantial intuitive overlap between titles and their associated topics. To give just a few examples, the topic that is most associated with Title 54 ("National Park Service") has top words of "park, secretari, nation, shall, land" and the topic that is most associated with Title 38 ("Veterans' Benefits") has top words of "veteran, renumb, secretari, disabl, administer." The overlap of topics and these substantive categories nicely illustrates the power of topic models to naively discover subject matter trends within textual corpora.

Assortativity Measures

Assortative mixing in networks is "the tendency for vertices. . .to be connected to other vertices that are like (or unlike) them in some way" [49]. Assortative mixing is observed in many natural networks—for example, partisan affiliation predicts connection on social networks [50] and is akin to what is referred to in the sociological literature as *homophily* [51]. We use measures of network assortative mixing or assortativity from [49]. Assortativity is calculated over edges in a graph and is the likelihood that an edge connects two nodes with the same characteristic.

Our analysis of assortativity measures is reported in **Figure 3**. For each title, we calculate two assortativity estimates: one for title itself and the second for the topic that is most closely associated with that title (as reported in **Supplementary Table S7**). More precisely, let e_{ij} denote the fraction of document-to-document

[9]Even the term "vocabulary" is used in a somewhat non-standard fashion: a "word" is sometimes a word fragment, as per the common technique of "stemming" used in topic modeling as well as other kinds of natural language processing algorithm and the extraction of the vocabulary is more or less a standardized process.

[10]The defining feature of STMs is the ability to use metadata when constructing topics. For this paper, we do not take advantage of this feature, and so the STM we use is equivalent to the correlated topic model (CTM) described in [20], which is an extension of the LDA approach. The authors of the STM have made the model publicly available through an *R*-package at http://www.structuraltopicmodel.com/.

[11]For this analysis, we exclude several 'generic' topics that do not appear to be related to substantive legal categories. These excluded topics are: 1, 7, 16, 20, 36, 42, 43, 73, 81, and 94.

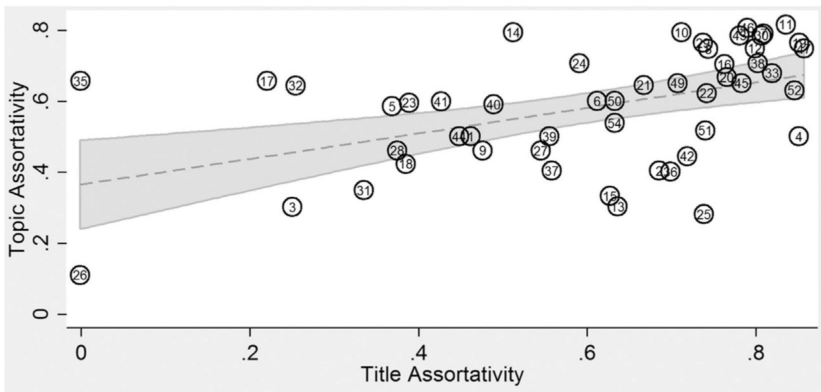


FIGURE 3 | Relationship between cross-reference assortativity at the title and topic level. Titles are matched with the topic that is most closely associated with that title as reported in **Supplementary Table S7**. Cross-references often track subject matter (as proxied by title and topic).

links (out of all links) that connect documents in titles/topics i and j (with the topic proviso above), then the *assortativity coefficient* (relative to the given labeling) is defined as

$$r = \frac{\sum_i e_{ii} - \sum_i a_i b_i}{1 - \sum_i a_i b_i} \tag{1}$$

where

$$a_i = \sum_j e_{ij} \qquad b_j = \sum_i e_{ij}$$

Higher assortativity is associated with a stronger correlation between title/topic and cross-references with an estimate of 1.0 implying perfect correlation [49].

In **Figure 3** observations are titles, and are labeled as such, and the dotted line and shadow is a simple linear fit with a 95% confidence interval. As is visually apparent, there is a relationship between a title's propensity to include intra-title cross-references and for its associated topic to self-cite. This finding tends to confirm that there is an important overlap between structure and content in the USC (An OLS regression with topic assortativity as the dependent variable and title assortativity as the predictor variable shows a relatively tight relationship, with an R^2 of 0.22 and a p-value less than 0.001.)

We conduct a further analysis of the extent to which topics explain citations between titles. Our question is, for any given topic are there titles that are typically cited to from statutory sections that are closely associated with that topic?[12] We call this the *authority* measure. Slightly more formally, for a topic U and Title T we calculate $Auth(U, T)$ as the correlation between the proportion of Topic U in the source of an edge and an indicator variable for Title T as the target of that edge. This is calculated using all edges except those whose source node is in T. For each topic we took the title with the highest correlation which had a p-value of less than 0.001. The results appear to align with intuition. For example Title 21 (Food and Drugs) has the highest authority for topics 49 and 82 with top words of "control, substance, drug, chemic, test" and "product, drug, food, secretari, provid" respectively. This means that when other titles discuss these topics their citations are most likely to be to Title 21. The results are reported in **Supplementary Table S8**.

Mutual Information-like Measures

We also engage in predictive exercises to test the degree to which content and structure carry mutual information. Informally, the mutual information between two random variables attempts to measure the degree to which the observation of one random variable may assist in the prediction of the observation of a second random variable. There are formal mutual information measures, such as *Kullback-Leibler divergence*, but there are none that we are aware of that fit well with the mixed data that we are considering. Instead we use a *prediction* task to operationalize an informal understanding of mutual information: two domains of structure and content have mutual information if it is possible to use information in one domain to make predictions concerning the other. Of course, different predictive approaches (and predictive targets) may be better or worse at leveraging certain kinds of information. Nevertheless, predictive performance using actual machine-learning algorithms provides something of a sense of mutual information, as data in two completely uncorrelated spaces would not be useful for generating predictions across domains.

The first predictive task utilizes a support vector machine (SVM) algorithm trained on 40,000 (out of 41,138) randomly drawn documents with the topic proportions in the documents used to predict the title where that document was found. Testing on the held out $1,138$ documents gives an accuracy of 60%, far more than would be expected from chance.[13]

[12]For this analysis, we dropped titles that were not closely associated with any topic. These were titles 1, 3, 4, 5a, 9, 13, 14, 18a, 24, 29, 32, 35, 44, 45, and 51.

[13]The weighted average F_1 score, which accounts for both precision and recall, is 0.59. For this analysis, we relied on the scikit-learn models in Python. The model used was sklearn.svm.LinearSVC. Interestingly, we saw a very large boost in accuracy when we switched from sklearn.svm.SVC to LinearSVC, with the primary difference between the learners being the use of a linear kernel instead of a radial basis function and handling multiclass labeling as one-to-many rather than one-to-one. The relative performance of these two learners may provide insight into the underlying statutory structure.

TABLE 1 | *Search model performance at predicting cross-references.*

Method	Average performance					
	precision@10 (%)	precision@20 (%)	precision@50 (%)	recall@10 (%)	recall@20 (%)	recall@50 (%)
Proximity	7.5	3.8	1.7	2.99	3.09	3.12
Covering	16.5	11.2	5.1	6.19	8.46	12.48

Our second analysis is based on the search framework described in [16, 17]. The goal of this framework is to generate a computational model of human law search in the navigation of a multinetwork representation of a legal corpus where edges are formed through citation and semantic similarity (as instantiated via a topic model). The authors refer to this multinetwork representation as a "legal landscape." This landscape is traversed by navigating from a source document—which is an exogenously identified member of the corpus—based on two general strategies: a "proximity strategy" and a "covering strategy." The proximity strategy identifies a set of documents that are "closest" to the source document (with distance defined via a specific network-based measure described in [16] that is called PageDist – see the paper for details). The covering strategy, by contrast, attempts to "cover" the range of subjects or issues within a legal document by setting off over a related range of the landscape from the source document.

One method used to test the performance of these strategies, described in detail in [17] relies on information that is embedded in the documents. In brief, the method begins by selecting a source document and then reconstructing the landscape without that document. The source document is then stripped of citation information (leaving a "Citation Free Legal Text" or CFLT). The information in the CFLT is quite coarse-grained because semantic content is represented as topic proportions only. Based on its topic proportions the CFLT is mapped onto the legal landscape (recall that its place in the multinetwork is generated by both citation structure and semantic content, the former now removed, but the latter still intact), and the proximity or covering algorithm is deployed. The success of the model (landscape + algorithm) is tested against the actual citations that were contained in that CFLT. In [17] traditional measures of performance precision and accuracy for a given number of predicted citations are reported. (Note that the number of citations to be generated is set exogenously rather than learned through the model, under an assumption that different searchers will weigh search costs vs. information benefits differently).

[16, 17] use the opinions generated by the Supreme Court of the United States (SCOTUS) as a test corpus. Herein we extend the methodology to the USC. It is worth noting differences between SCOTUS opinions and the USC. The primary and most important difference is that the citation network for SCOTUS opinions is extremely dense, with opinions containing dozens of citations and few opinions with a very small number of citations. By comparison, the USC citation network is very sparse, with zero as the modal and median number of citations. Because citation-less documents cannot be easily incorporated into the landscape, we exclude them from this analysis. Further, because precision and recall are difficult to estimate with a very small number of citations, we limit our CFLTs to the documents that have at least five citations (for this analysis, only outgoing citations were used). It is also worth noting that the semantic content of the two corpora are very different. SCOTUS opinions are meant to persuade and generally conform to the norms of the judicial genre [5] such as stating the facts of a case in a narrative voice and offering reasons for the decision delivered. It is unclear whether the different semantic styles in the two corpora will lead to different navigation behavior on the part of law searchers.

Table 1 presents average precision and recall for different number of recommendations based on roughly 300 CFLTs. These results are roughly commensurate with the model's performance for the SCOTUS corpus (*see* [17].) For the (better performing) covering algorithm, out of the first 10 recommendations, a bit under two would be accurate matches. When the model generates 50 recommendations, a bit over 10% of the actual citations are identified.

As would be expected, there is an inverse relationship between precision and recall as the number of potential cross-references that are identified increases. An additional finding is that the covering algorithm outperforms the proximity algorithm in both the SCOTUS and USC context, indicating that it may be a more robust general approach for simulating human navigation of even very different legal corpora.

It is worth noting a further point of comparison with [17]. In that paper, comparison is made between the model's performance to human research assistants on basic research tasks, finding that although the models do not perfectly simulate natural search behavior, the degree of overlap of the model to the research assistants was not so much less than the overlap of the researchers with each other. Although we do not undertake the same analysis here, it is plausible to speculate that a similar performance would be achieved for the USC, given the model's relatively similar performance on the citation prediction task.

5 CONCLUSION AND FURTHER WORK

This preliminary analysis shows a strong association between structure and content on the USC. The former is embodied in the organizational hierarchy and section-level cross-reference link structure. The latter is quantified through the application of a structural topic modeling of the sections – a contribution in and of itself to the study of statues. There are several potential extensions that we discuss briefly in this section along with a summary of our findings.

The USC is an important corpus with profound legal effect. We find that there is a substantial degree of overlap (in the sense of correlation) between statutory structure and content via two measures: assortativity and prediction. We find a relatively large amount of topic assortativity, as well as intuitive matches between topics and statutory titles and also strong correlations between title assortativity and the assortativity of matched topics. We also find that there is sufficient mutual information between structure and content in the sense that the topic share information of a document can be used to predict structural information (titles and cross-references) using both a trained machine learning classifier (SVM)—for titles—and via the legal landscapes approach borrowed from [16]—for cross-references.

Although these observations are interesting, it is admittedly somewhat difficult to interpret the relationships described in this paper absent a baseline for analysis. There are two possible data sources for such a baseline: domestic statutes from a comparative context, and other U.S. domestic legal orders, such as state statutory regimes. An important extension of the work reported in this paper would apply similar techniques to other statutory corpora. All of the methods described here are general and would be applicable to similarly structured statutory systems. So long as citation information can be extracted, statutory structure can be captured through the same network notions, and topic models are language-agnostic so long as the underlying texts are machine readable. What we describe in this paper then amounts to an off-the-shelf methodology that can, in principle be applied to any statutory regime. Comparative work along these lines is likely to be particularly fruitful because it provides a means of examining the relative strength of the relationship between structure and content in different statutory regimes. We cannot know from this initial analysis whether the United States is an outlier, or if similarly situated legal orders tend to have similar levels of content/structure interrelatedness. In another direction, comparative work of state-level statutory networks would also be interesting.

A second extension would delve deeper into the USC itself, perhaps by linking this corpus to other legal texts, such as the Code of Federal Regulations. Especially with a larger total corpus, it would be possible to make intra-USC comparison between, for example, different titles or different substantive areas (as estimated either naively via a topic model or through expert labeling). It would be worth investigating whether some areas of the law are more "self-contained" than others, and whether any measure of self-containment correlated with other characteristics, such as the size of the industry that was regulated or partisan dynamics (either temporally at the national level or geographically at the state level).

A final set of extensions are more practical in nature. Statutory texts are notoriously difficult to navigate, in part because they lack the kind of identifying information that is contained in judicial opinions. Research into the relationship between statutory structure and content (and the relationships between statutes and other legal documents) could be used as the foundation for new search/navigation tools. Such tools could be used by practitioners to lower the transaction costs associated with identifying relevant statutory texts. Given the substantial private expenditures on law search, any technique that lowered those costs by even a small percentage would create a substantial amount of economic value.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://uscode.house.gov/.

AUTHOR CONTRIBUTIONS

FD and KC did much of the data extraction and computation. All authors analyzed the results. All authors contributed to the writing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2020.625241/full#supplementary-material.

REFERENCES

1. Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA. Multilayer networks. *J Complex Networks* (2014) 2(3):203–71. doi:10.1093/comnet/cnu016
2. Bommarito MJ, Katz DM. A mathematical approach to the study of the United States Code. *Physica A: Stat Mech Appl* (2010) 389(19):4195–200. doi:10.1016/j.physa.2010.05.057
3. Corominas-Murtra B, Goñi J, Solé RV, Rodríguez-Caso C. On the origins of hierarchy in complex networks. *Proc Natl Acad Sci* (2013) 110(33):13316–21. doi:10.1073/pnas.1300832110
4. Mengistu H, Huizinga J, Mouret JB, Clune J. The evolutionary origins of hierarchy. *PLoS Comput Biol* (2016) 12:e1004829. doi:10.1371/journal.pcbi.1004829
5. Livermore MA, Riddell AB, Rockmore DN. The Supreme court and the judicial genre. *Ariz L Rev* (2017) 59(4):837–901.
6. Rockmore DN, Fang C, Foti NJ, Ginsburg T, Krakauer DC. The cultural evolution of national constitutions. *J Assoc Inf Sci Technol* (2017) 69(3):483–94. doi:10.1002/asi.23971
7. Katz DM, Bommarito MJ. Measuring the complexity of the law: the United States Code. *Artif Intell L* (2014) 22(4):337–74. doi:10.1007/s10506-014-9160-8
8. Sunstein CR. *Simpler: the future of government*. New York, NY: Simon & Schuster (2013).
9. Al-Ubaydli O, McLaughlin PA. RegData: a numerical database on industry-specific regulations for all United States industries and federal regulations, 1997–2012. *Regul Governance* (2017) 11(1):109–23. doi:10.1111/rego.12107
10. Carey, MP. Counting regulations: an overview of rulemaking, types of federal regulations, and pages in the federal register (2016). Congressional Research Service. Technical report.
11. Lumsdaine RL, Rockmore DN, Foti NJ, Leibon G, Farmer JD. The intrafirm complexity of systemically important financial institutions. *J Financial Stab* (2020) 52:100804. doi:10.1016/j.jfs.2020.100804

12. Fowler JH, Johnson TR, Spriggs JF, Jeon S, Wahlbeck PJ. Network analysis and the law: measuring the legal importance of precedents at the U.S. Supreme Court. *Polit Anal* (2007) 15(3):324–46. doi:10.1093/pan/mpm011
13. Rice D. The impact of Supreme court activity on the judicial agenda. *L Soc'y Rev* (2014) 48:63–90. doi:10.1111/lasr.12056
14. Carlson K, Livermore MA, Rockmore DN. A quantitative analysis of writing style on the U.S. Supreme court. *Wash U L Rev* (2016) 93:1461.
15. Carlson K, Livermore MA, Rockmore DN. The problem of data bias in the pool of published U.S. Appellate court opinions. *J Empirical Leg Stud* (2020) 17(2): 224–61. doi:10.1111/jels.12253
16. Leibon G, Livermore M, Harder R, Riddell A, Rockmore D. Bending the law: geometric tools for quantifying influence in the multinetwork of legal opinions. *Artif Intell L* (2018) 26(2):145–67. doi:10.1007/s10506-018-9224-2
17. Livermore MA, Rockmore DN, Dadgosari F, Guim M, Beling PA. Modeling law search as prediction. *Artifi Intell Law* (2020) 29:3–34. doi:10.1007/s10506-020-09261-5
18. Postema GJ. *Bentham and the common law tradition*. Oxford, United Kingdom: Oxford University Press (1989).
19. Badawi AB, Dari-Mattiacci G. Reference networks and civil codes In: MA Livermore DN Rockmore, editors *Law as data: computation, text, and the future of legal analysis*. Chap. 12, San Fante, NM: SFI Press (2019). p. 339–65.
20. Blei DM, Lafferty JD. A correlated topic model of Science. *Ann Appl Stat* (2007) 1(1):17–35. doi:10.1214/07-aoas114
21. Roberts ME, Stewart BM, Tingley D, Lucas C, Leder-Luis J, Gadarian SK, et al. Structural topic models for open-ended survey responses. *Am J Polit Sci* (2014) 58(4):1064–82. doi:10.1111/ajps.12103
22. Quinn KM, Monroe BL, Colaresi M, Crespin MH, Radev DR. How to analyze political attention with minimal assumptions and costs. *Am J Polit Sci* (2010) 54(1):209–28. doi:10.1111/j.1540-5907.2009.00427.x
23. Riddell AB. How to read 22,198 journal articles: studying the history of German studies with topic models In: M Erlin L Tatlock, editors *Distant readings: topologies of German culture in the long nineteenth century*. Rochester, NY: Camden House (2014). p. 91–114.
24. Jockers ML, Mimno D. Significant themes in 19th-century literature. *Poetics* (2013) 41(6):750–69. doi:10.1016/j.poetic.2013.08.005
25. Roe G, Gladstone C, Morrissey R. Discourses and disciplines in the enlightenment: topic modeling the French encyclopédie. *Front Digit Humanit* (2016) 2:8. doi:10.3389/fdigh.2015.00008
26. Schöch C. Topic modeling genre: an exploration of French classical and enlightenment drama. *DHQ: Digital Humanities Q* (2017) 11(2):266–85. doi:10.5281/zenodo.166356
27. Macey J, Mitts J. Finding order in the morass: the three real justifications for piercing the corporate veil. *Cornell L Rev* (2014) 100(1):99–155.
28. Law DS. Constitutional archetypes. *Tex L Rev* (2016) 95(2):153–243.
29. Lauderdale BE, Clark TS. Scaling politically meaningful dimensions using texts and votes. *Am J Polit Sci* (2014) 58(3):754–71. doi:10.1111/ajps.12085
30. Chalfin A, Haviland AM, Raphael S. What do panel studies tell us about a deterrent effect of capital punishment? A critique of the literature. *J Quant Criminol* (2013) 29(1):5–43. doi:10.1007/s10940-012-9168-8
31. Graham ER, Shipan CR, Volden C. The diffusion of policy diffusion research in political science. *Br J. Polit. Sci.* (2013) 43(03):673–701. doi:10.1017/s0007123412000415
32. La Porta R, Lopez-de SIlanes F, Shleifer A, Vishny RW. Law and finance. *J Polit Economy* (1998) 106(6):11131–55. doi:10.1086/250042
33. Porta RL, Lopez-de-Silanes F, Shleifer A. The economic consequences of legal origins. *J Econ Lit* (2008) 46(2):285–332. doi:10.1257/jel.46.2.285
34. Michaels R. Comparative law by numbers? Legal origins thesis, doing business reports, and the silence of traditional comparative law. *Am J Comp Law* (2009) 57(4):765–95. doi:10.5131/ajcl.2008.0022
35. Klerman DM, Mahoney PG, Spamann H, Weinstein MI. Legal origin or colonial history?. *J Leg Anal* (2011) 3(2):379–409. doi:10.1093/jla/lar002
36. Spamann H. Empirical comparative law. *Annu Rev L Soc. Sci.* (2015) 11: 131–53. doi:10.1146/annurev-lawsocsci-110413-030807
37. Ruhl JB, Katz DM. Measuring, monitoring, and managing legal complexity. *Iowa L Rev* (2015) 101(1):191–244.
38. Schuck PH. Legal complexity: some causes, consequences, and cures. *Duke L J* (1992) 42(1):1–52. doi:10.2307/1372753
39. Tainter JA. *The collapse of complex societies*. Cambridge, United Kingdom: Cambridge University Press (1988).
40. Fonseca R, Lopez-Garcia P, Pissarides CA. Entrepreneurship, start-up costs and employment. *Eur Econ Rev* (2001) 45(4-6):692–705. doi:10.1016/s0014-2921(01)00131-3
41. Nicoletti G, Scarpetta S. Regulation, productivity and growth: OECD evidence. *Econ Policy* (2003) 18(36):9–72. doi:10.1111/1468-0327.00102
42. Ciccone A, Papaioannou E. Red tape and delayed entry. *J Eur Econ Assoc* (2007) 5(2-3):444–58. doi:10.1162/jeea.2007.5.2-3.444
43. Braunerhjelm P, Eklund JE. Taxes, tax administrative burdens and new firm formation. *Kyklos* (2014) 67(1):1–11. doi:10.1111/kykl.12040
44. Ellig J, McLaughlin PA. The regulatory determinants of railroad safety. *Rev Ind Organ* (2016) 49(2):371–98. doi:10.1007/s11151-016-9525-0
45. Chambers D, Collins CA, Krause A. How do federal regulations affect consumer prices? An analysis of the regressive effects of regulation. *Public Choice* (2017) 180(1):57–90. doi:10.1007/s11127-017-0479-z
46. U.S. Chamber of Commerce. The regulatory impact on small business: complex. cumbersome. costly (2017). U.S. Chamber of Commerce. Technical report.
47. Kleinberg JM. Authoritative sources in a hyperlinked environment. *J Acm* (1999) 46(5):604–32. doi:10.1145/324133.324140
48. Blei DM. Probabilistic topic models. *Commun ACM* (2012) 55(4):77–84. doi:10.1145/2133806.2133826
49. Newman MEJ. Mixing patterns in networks. *Phys Rev E* (2003) 67(2). doi:10.1103/physreve.67.026126
50. Colleoni E, Rozza A, Arvidsson A. Echo chamber or public sphere? Predicting political orientation and measuring political homophily in twitter using big data. *J Commun* (2014) 64(2):317–32. doi:10.1111/jcom.12084
51. McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. *Annu Rev Sociol* (2001) 27(1):415–44. doi:10.1146/annurev.soc.27.1.415
52. Whisner M. The United States code, prima facie evidence, and positive law. *L Libr Journal* (2009) 101(4):545–56.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 26 April 2021
doi: 10.3389/fphy.2021.650943

Using Complexity to Calibrate Legal Response to Covid-19

*Ofer Malcai and Michal Shur-Ofry**

Faculty of Law, Hebrew University of Jerusalem, Jerusalem, Israel

The global effort to fight the Covid-19 pandemic triggered the adoption of unusual legal measures that restrict individual freedoms and raise acute legal questions. Yet, the conventional legal tools available to analyze those questions—including legal notions such as proportionality, equality, or the requisite levels of evidence—implicitly presume stable equilibria, and fail to capture the nonlinear properties of the pandemic. Because the pandemic diffuses in a complex system, using complexity theory can help align the law with its dynamics and produce a more effective legal response. We demonstrate how insights from complexity concerning temporal and spatial diffusion patterns, or the structure of the social network, can provide counter-intuitive answers to a series of pandemic-related legal questions pertaining to limitations of movement, privacy, business and religious freedoms, or prioritizing access to vaccines. This analysis could further inform legal policies aspiring to handle additional phenomena that diffuse in accordance with the principles of complexity.

Keywords: law, complexity, COVID-19, exponential diffusion, fractal, proportionality, evidence, networks and privacy

Edited by:
Daniel Martin Katz, Illinois Institute of Technology, United States

Reviewed by:
Katherine Strandburg, New York University, United States
Daniel Rodriguez, Northwestern University, United States

***Correspondence:**
Michal Shur-Ofry
michalshur@mail.huji.ac.il
orcid.org/0000-0002-3623-7769

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *08 January 2021*
Accepted: *15 March 2021*
Published: *26 April 2021*

Citation:
Malcai O and Shur-Ofry M (2021) Using Complexity to Calibrate Legal Response to Covid-19. Front. Phys. 9:650943. doi: 10.3389/fphy.2021.650943

INTRODUCTION

Covid-19 is presenting unprecedented legal challenges to traditional legal policy. The effort to curb the pandemic entails the adoption of unparalleled measures that seriously compromise fundamental legal rights, ranging from free movement, to privacy, to the right to conduct business, or to carry out religious practices. The uneven spread of the pandemic within geographical areas triggers differentiated policy responses that ostensibly collide with legal notions of equality, whereas the high levels of uncertainty and unpredictability concerning the diffusion of the disease challenge the traditional legal stance that policy measures must rely on solid evidence.

Law and policy proposals have so far concentrated on insights from behavioral economics, which could assist policy makers in designing their response to the pandemic, for example, by effectively nudging people toward adopting desirable behaviors (e.g., [1]; cf. [2, 3]). This contribution wishes to draw the attention to an additional perspective for improving the legal response to Covid-19, that of complex systems.

The pandemic is a systemic phenomenon. While there are still indeterminacies as to its precise attributes, there is ample evidence that its general diffusion is consistent with temporal exponential dynamics and spatial fractal patterns that characterize diffusion in complex systems. Yet, many of the legal principles that are employed in handling the current challenges—including notions such as "proportionality," "equality," or "requisite levels of evidence" — were designed against implicit assumptions of stable equilibria and linear processes, and are therefore misaligned with the traits of the pandemic. We seek to mitigate this disconnect, and show how engaging with complexity theory can help calibrate legal policies to produce a more effective legal response to Covid-19.

Methodologically, our analysis relies both on specific research concerning the Covid-19 pandemic and its diffusion patterns, as well as on a large body of scholarship from the recent decades, in physics, network science and adjacent areas, that investigated and delineated the general traits of complex systems. While this literature increasingly influences numerous domains, its use in theoretical legal analysis, or in legal policy design is still relatively limited (cf. [4–6]). The following analysis demonstrates how engaging with insights from complexity theory can shed light on a series of pandemic-related legal questions, concentrating on four acute examples: (1) the legality of regulatory measures that limit individual liberties; (2) evidentiary questions concerning pandemic-related decision making; (3) the application of differentiated policies to various geographical areas; and (4) the use of network tools for prioritizing tests and vaccines. Our examples are illustrative and non-exhaustive. Rather, this perspective invites additional research that would assist in integrating the science of complex systems into legal policy, to improve the legal response to the pandemic, as well as to additional phenomena that diffuse in accordance with the principles of complexity, including, prominently, environmental challenges.

RESTRICTIVE MEASURES: PROPORTIONALITY AND EXPONENTIAL DYNAMICS

Many unusual measures that were adopted in the face of Covid-19 since March, 2020—including lockdowns, borders closures, the use of contact tracing technologies, the closure of businesses and restrictions on religious practices—were challenged before various courts worldwide (e.g., cases #4-21)[1]. In many jurisdictions, the legal scrutiny of those steps, and similar measures that restrict individual liberties, utilizes the notion of *proportionality*. Proportionality implies a balancing exercise: the steps imposed by governments or regulators must be weighted against the social harm they seek to prevent (e.g., [7, 8]). Relatedly, under prevalent legal doctrine, when the general social interest requires restricting or prejudicing individual rights, regulators must choose the "least restrictive means," namely only the necessary and most lenient measures available, which least interfere with individual liberties (e.g., [9], p. 464; [10, 11]). Applying these principles, courts in various countries ruled that measures such as lockdowns (cases #16, #19), restrictions on worship (case #4–10), or business closures (cases #11, #18) are disproportionate and illegal.

Both the proportionality principle and the "least restrictive means" principle envisage ordinary circumstances, where the restrictive steps are balanced against potential harms that are either relatively stable, or increase linearly with time. Consider, for example, the case of preventing a possible terrorist attack that could harm dozens of people. Under standard legal analysis, taking certain steps that compromise individual freedoms—such as instructing people to take off their shoes at airport security—in order to prevent this outcome would be considered proportionate, whereas adopting more extreme measures, such as a complete closure of borders would likely fail both the proportionality and the least-restrictive-means tests.

However, the assumptions underlying the proportionality test do not accurately reflect phenomena diffusing in complex systems. Typically, such diffusion displays temporal growth patterns that are non-linear. More specifically, while during the initial stages the rate of diffusion could be relatively stable, at a certain point, once contagion processes take off, it is expected to rise exponentially, leading to a sharp increase in the spread of the relevant phenomenon (e.g., [12–14]). The evidence concerning Covid-19 instructs that the pandemic indeed diffuses in accordance with these patterns. **Figure 1**, which describes the cumulative number of Covid-19 cases in the UK during February-March 2020 illustrates these dynamics[2].

As a result of the non-linear diffusion dynamics, the expected number of people contracting the disease is likely to grow exponentially, while the effectiveness of regulatory measures employed to restrain its diffusion is likely to significantly decline with time. To illustrate, let us consider three points in time in **Figure 1**. At T1 = Feb 24, the number of confirmed cases in the UK was still very small (13 cumulative cases), and the diffusion relatively slow. Two weeks later, at T2 = March 9, the number of cases was 677 and was growing exponentially, doubling itself every 3–4 days, so that at T3 = March 23, the disease was already widespread within the population, reaching 12,647 cumulative cases. Applying strict measures such as border closure at T1 can effectively curb a pandemic (provided, of course, it is accompanied by additional inter-state actions of contact tracing). Yet this same measure is much less effective if adopted only 2 weeks later (at T2), and almost insignificant at T3, when the disease has spread among the entire state's population so that a few more cases "imported" from other infected countries would not result in a substantial change.

Yet, from the perspective of law and policy making exponential growth may be hard to grasp: many serious problems that beg policy responses, from road accidents to the prevalence of cancer, do not exhibit these growth dynamics. Moreover, because diffusion starts slowly and steadily, it might be difficult, in early stages, to distinguish exponential growth from linear growth and realize the huge potential magnitude thereof, a phenomenon known as the "exponential growth bias" (e.g., [16, 17]). As a result, courts may be inclined to apply a conventional proportionality test, implicitly presuming that the pandemics' harms will grow in a stable and linear way.

The lens of complexity implies that, contrary to conventional legal analysis, the regulatory response to a pandemic may, in certain cases, warrant an "*inverse proportionality test.*" When the diffusion dynamics are nonlinear and the potential harm is likely to accumulate exponentially, strict measures to prevent it could

[1]We focus primarily on court cases and judge-made doctrines, rather than on administrative regulations that naturally vary significantly among jurisdictions.

[2]Similar growth patterns were documented in many other countries—see, for example, the growth of Covid-19 cases in the United States during March, 2020, World Health Organization, https://www.who.int/countries/usa/.

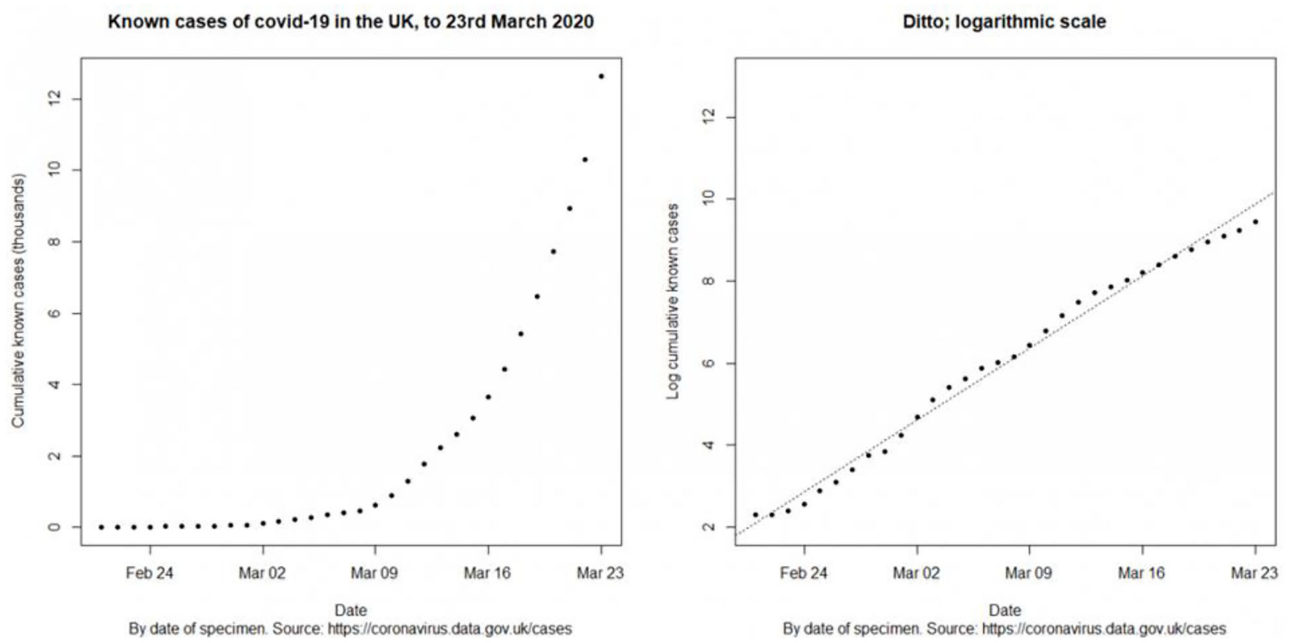


FIGURE 1 | The cumulative number of Covid-19 cases in the UK as a function of time in linear (left) and a log-log (right) representations (source: Stevens et al. [15]).

be considered proportionate at an early stage, when the actual harm is least apparent and least certain. Counterintuitively, those very same measures might be less defensible at a later stage when the large harms of the pandemic have already materialized. To illustrate, a legal proportionality test that is adapted to nonlinear dynamics would ratify China's decision to test nine million people in October 2020, in light of only a dozen detected cases of Covid-19 [18]. Similarly, New Zealand's "Go Hard, Go Early" policy, which entailed early border closure and a full lockdown before there was even a single death from Covid-19 in the country [19] would be considered proportionate. Notably, this policy, which very much aligns with the dynamics of complex systems, has also turned out to be extremely effective in controlling the pandemic. Conversely, the same "inverse proportionality test" might not endorse Israel's second border closure, imposed in September 2020, when the number of active cases in the country has already soared and reached tens of thousands of cases [20][3].

From a legal perspective, one might argue that instead of an "inverse proportionality test" our insights can be incorporated into the current proportionality doctrine, so that courts would simply take into account the complex systems properties of the relevant phenomenon, when assessing the proportionality of the state's response. Ideally, this approach would yield similar results. Yet from a practical perspective, framing the test as an "inverse proportionality test" has several advantages. First, it signals to courts and policy makers (who might not always be familiar with complexity theory) that certain phenomena warrant a qualitatively different response relative to "ordinary" situations. Secondly, it facilitates recognition that seemingly small current problems might justify major interventions. Third, it illuminates that in certain circumstances, late interventions, which may seem proportionate, might actually be no longer effective[4].

We do not delineate here a comprehensive set of circumstances that warrant an "inverse proportionality" test, beyond the case of Covid-19. As a rule of thumb, an inverse proportionality test could be appropriate when dealing with multiplicative (rather than additive) risks, that pose systemic threats, particularly when the relevant phenomenon diffuses in short time scales. These factors, and their further development in future research, could guide legal policy and prevent the misuse of the test in other, more ordinary circumstances.

Relatedly, a complexity-based analysis warrants the calibration of the legal requirement of "least restrictive means" to the stage of diffusion. The implicit assumptions underlying this legal doctrine—that the "second-best" and least restrictive measures are close-enough to the more restrictive measures, or that a gradual escalation of means is a feasible and reasonable policy—do not hold true in the face of exponential growth. Rather, under exponential diffusion dynamics adopting "very" or even "most" restrictive means at a very early stage may be orders of magnitude more effective than adopting "less" or "least" restrictive means.

A comparison of the measures imposed by Greece and Spain during the first wave of the pandemic, as described in **Figure 2**,

[3]Importantly, the aforesaid border closure preceded the emergence of new variants of the virus. The appearance of variants can set back the "diffusion clock" (with respect to the new variant) and in certain cases might create new justifications for closure.

[4]In that sense, our argument for an "inverse proportionality test" can be conceived as a *non-ideal* theory, aiming at guiding actions in the real world.

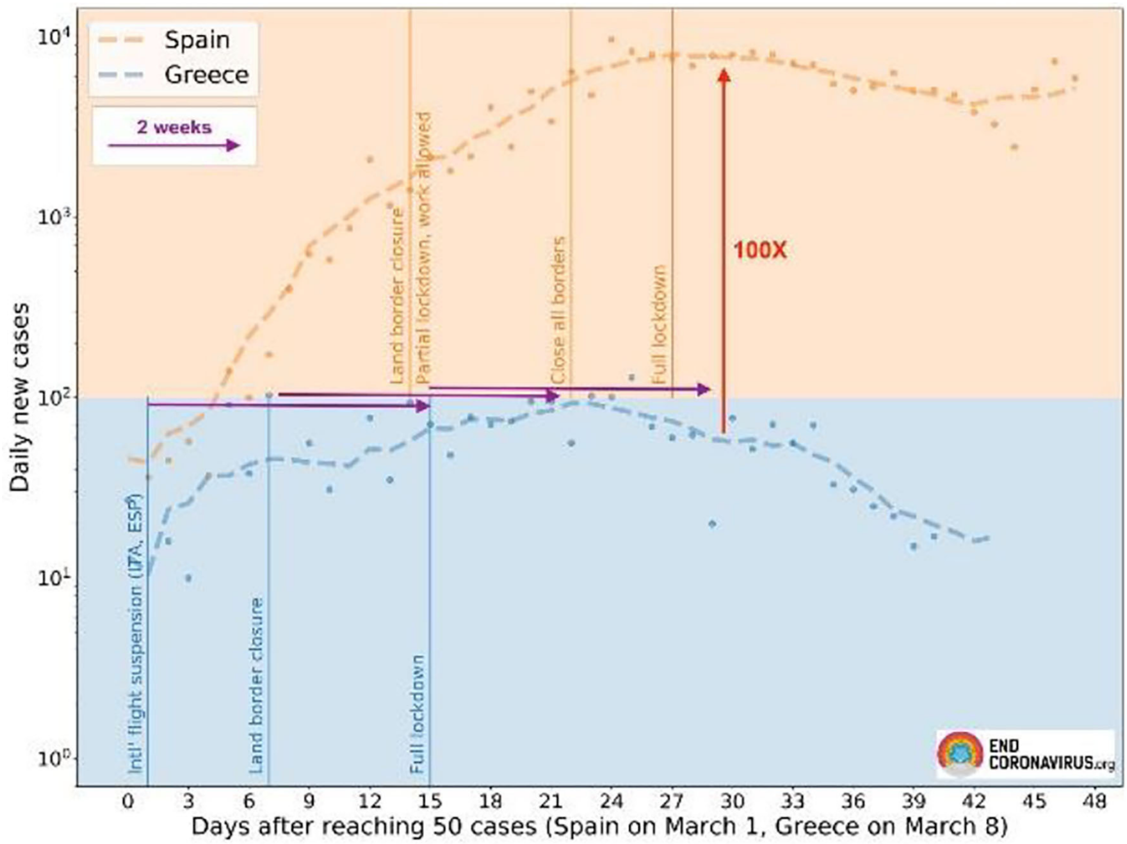


FIGURE 2 | Escalation of Measures: Spain v. Greece (Source: Yaneer Bar-Yam, EndCoronavirus.org).

is illustrative. The two countries started out with roughly the same number of cases. Within 2 weeks, Greece imposed "very restrictive measures," including border closure and a full lockdown, while Spain applied a "least restrictive measures" approach, gradually escalating from a partial lockdown to a full lockdown after 4 weeks. After 30 days, the number of daily cases in Greece was less than a hundred, while the number of cases in Spain was in the range of tens of thousands[5].

Adapting law to exponential diffusion dynamics therefore implies that, as part of the inverse proportionality test, the adoption of harsher (rather than least restrictive) means at an early stage of diffusion can be considered proportionate, and legal[6].

COMPLEXITY, EVIDENCE, AND DECISION MAKING

Understanding complex systems can further illuminate a series of legal questions concerning evidence and decision making during the pandemic.

First, *which evidence* is required before adopting pandemic-related decisions by states and regulators? Under conventional legal principles, common to numerous jurisdictions, government decision-making should be based on evidence, and informed by the best available data (e.g., [21]). These principles apply in particular force when state's policies infringe fundamental rights. In such circumstances the state needs to prove, on the basis of reliable evidence, that its actions are in fact necessary, and the standard of proof would generally be higher than in regular civil cases [e.g., [22]].

In the context of Covid-19, various scholars and policy makers advocated that the imposition of restrictive measures lacked sufficient evidentiary basis, and called for "more reliable data" before adopting "draconian countermeasures" (e.g., [23]).

[5]It should be clarified that while this comparison is illustrative and suggestive, it cannot provide a definitive explanation of the differences between the countries, since we do not consider additional factors which may have affected the diffusion dynamics.

[6]There are, of course, nuances among jurisdictions which we do not fully explore here. Some jurisdictions regard the "least restrictive means" test as a second step in the legal analysis, which comes into play only after it has been determined that an effective response is required. In such cases the state should still choose, among the range of *effective* means, those that impose the least restrictions on individual liberties, even under an "inverse proportionality test." However, in practice, the application of the two steps is often entangled. Moreover, as we discuss in the following Section, due to complex systems' dynamics there are significant uncertainties as to the precise effectiveness of particular measures. It is therefore important to recognize that very restrictive means may be necessary in order not to render the entire response ineffective.

Concomitantly, there have been continuous scientific efforts to produce single point forecasts predicting the pandemic's spread, through the use of various mathematical models (such as the logistic model, the SIR model, agent-based models, and variations thereof).

The lens of complexity instructs that the nonlinear properties of the pandemic render the ordinary legal expectation for accurate, specific, predictions largely unrealistic. Every mathematical model inevitably involves a degree of simplification and idealization. In linear systems certain simplification—which entails neglecting certain properties of the modeled phenomenon—is unlikely to bring about significant differences in the overall prediction [24]. Conversely, in complex nonlinear dynamic systems like the Covid-19 pandemic, slight differences in initial conditions assumed by mathematical models, can yield extreme differences in the total outcome. The precise interactions and interdependencies among the various components comprising the system, the system's structure, the sequence of interactions and other factors, that might be minuscule or random, especially during the initial stages of diffusion, can all result in vast differences in the system's overall response (e.g., [25–27]).

Research pertaining to Covid-19 indeed demonstrates how very small differences in the pandemic's growth rate, in the precise implementation details of the regulatory interventions, or in the population's degree of compliance with those regulations, may yield vast changes in the disease's trajectory [28]. These nonlinear dynamics make single point forecasts essentially impossible, and even led some scholars to maintain that the turning point of the epidemic, namely the point where the growth in the number of cases starts decreasing, cannot be predicted with any certainty before it actually occurs [29].

Moreover, pandemics are "fat-tailed" events, which implies that "frequent" (median) data observations do not provide a good indication of the average, or the magnitude, of the overall phenomenon (e.g. [30–32])[7]. This property further explains why Covid-19 does not lend itself to simple predictions. It also clarifies that additional data, especially observations coming from the bulk of the distribution–for example, data on the daily numbers of cases during the beginning of a pandemic–does not guarantee extra knowledge that will allow more accurate prediction of the overall phenomenon.

Due to these properties, waiting for more positive evidence will not necessarily produce more meaningful information for policy making. Furthermore, in the face of exponential diffusion postponing interventions in the hope of gathering "more data" is expected to make such interventions less effective and costlier. Rather than insist on precise predictions that are likely impossible, courts and policy makers may have to "satisfice" with evidence concerning the *general properties* of the Covid-19 pandemic as a non-linear, unstable and essentially collective phenomenon that diffuses exponentially. These properties are familiar characteristics of complex systems for which there is ample proof. The foregoing analysis also implies that the legal approach to the pandemic should be responsive, maintain flexibility, and take into account its inherent unpredictability (cf. [33]).

Nevertheless, courts applying judicial review of pandemic-related decisions sometimes tend to apply the conventional legal prescription and insist on more accurate data that cannot be realistically obtained. One illustration is a decision of the United States Court of Appeals for the 3rd Circuit in the case of *United States v. Raia* from April, 2020 (Case # 1). The court denied a criminal defendant's motion for compassionate release from prison given the Covid-19 pandemic, despite his numerous medical risk factors. The decision held that the applicant must exhaust administrative remedies prior to seeking relief in court. The "mere possibility" that the pandemic would spread in the prison system and harm him was insufficient for immediate compassionate release. The implicit requirement for more precise and concrete evidence may constitute an appropriate legal standard of review under ordinary circumstances. Yet, due to the nonlinear properties of Covid-19, satisficing with "mere possibility" might be inevitable.

Secondly, understanding the properties of complex phenomena should also influence the *weight and evaluation of evidence*. Both policy makers and legal professionals are trained to evaluate evidence according to past life-experiences, which are assumed to remain more or less stable. In the context of Covid-19, this conventional method may lead courts and policy makers to interpret *lack of evidence* that a certain activity triggers infections, as *positive evidence* that this activity is in fact "safe" (cf. [34]). For example, a research that examined over 3,000 people who trained in a gym in Oslo during May 2020 found no infected cases and concluded that "provided good hygiene and social distancing measures, there was no increased Covid-19 spread at training facilities." ([35], p. 2). This research was subsequently presented to the Israeli Parliament as positive evidence that opening gyms entails no added risk for infections ([36], p. 14). However, during the period of the study the pandemic in Oslo was in its early stages, and the number of new daily cases in the city was practically zero[8]. Taking into account the stage of diffusion, this absence of evidence about infections cannot be translated into meaningful knowledge about absence of risk for infections in gyms. The situation could be different in later stages of the pandemic's diffusion, where similar evidence, if obtained, may be more meaningful.

Finally, and relatedly, conventional legal analysis of evidence usually adopts a reductionist perspective that concentrates on an individual person or action (such as a particular crime, or a specific transaction) (cf. [37]) Therefore, policy makers adopting pandemic-related decisions, and courts reviewing such decisions, may wrongly apply a simple "additive" approach, assuming that because the risk to each individual and from

[7] The cumulative distribution of a random variable X is "fat-tailed" if its tail decays slowly; e.g., like power law: $\lim_{x\to\infty} P[X > x] \sim x^{-\alpha}$. For such distributions, the moment of order k exists [i.e., E(X) is finite] if and only if $\alpha > k$. Thus, if $\alpha \leq 1$, the mean does not exist. For $1 < \alpha \leq 2$ the variance is infinite. In this case the mean exists but the sample mean will converge very slowly with the true mean, and the standard statistical errors will understate the true uncertainty of the phenomenon [31, 32].

[8] See John Hopkins Corona Virus Resources Center. Available online at: https://coronavirus.jhu.edu/data.

each individual activity is low, the systemic risk caused by the pandemic is equally small ([31], p. 607). However, acquaintance with the properties of complex systems clarifies that, due to the interactions among individuals and the multiplicative dynamics of the pandemic, these small risks at the micro level can easily translate into a large systemic risk resulting in millions of death at the macro level, as indeed has been the case with Covid-19[9]. A complexity-informed approach therefore prescribes that the general evidentiary trajectory, which should guide judicial and regulatory decision-making in light of the uncertainties entailed in Covid-19 is to *err on the side of safety*.

The legal doctrine known as "the precautionary principle" echoes this logic. According to this principle "[w]hen an activity raises threats of harm to human health or the environment, precautionary measures should be taken even if some cause-and-effect relationships are not fully established scientifically. In this context the proponent of an activity, rather than the public, should bear the burden of proof" ([39]; cf. [40]). Originating in environmental and international law, the precautionary principle was adopted by courts and regulators in various jurisdictions (e.g., [34, 41]). Concomitantly, it has been subject of intense criticism, maintaining, essentially, that it is vague and fails to provide guidance sensitive to cost-benefit analysis (e.g., [42]). A complexity-based approach can alleviate some of these concerns by delineating several factors, which may serve as guidelines for the principle's application: phenomena that spread exponentially, in short time-scales, and pose systemic, existential, risks[10]. Under such circumstances, the risk is multiplicative while the costs of adopting precautionary measures are often additive–which indicates that under a cost-benefit analysis the overarching trajectory points toward precaution. Covid-19 may be a paradigmatic example, especially given its rapid diffusion, but it is not a single case. Legal scholarship has already observed that global environmental threats, particularly climate change, may pose unique systemic challenges to the legal system (e.g., [43, 44]). Increasing evidence suggests that phenomena related to climate change display exponential properties, albeit on a different time-scale, and similarly entail embedded uncertainty (e.g., [34, 45, 46]). The foregoing analysis indicates that courts and policy makers should favorably consider applying the precautionary principle in these and similar instances.

DIFFERENTIAL TREATMENT, EQUALITY, AND FRACTALS

In addition to the general legality of restricting basic freedoms, the response to the Covid-19 pandemic raises acute legal questions concerning *equality*. A central debate in this respect relates to the legitimacy of collective differential measures, primarily localized lockdowns implemented over a limited geographical area–ranging from small-scale units such as neighborhoods, to larger scales such as cities or regions.

Unlike nation-wide lockdowns, which limit basic rights equally, localized lockdowns create substantial inequalities in the limitations they impose on basic rights, and might be conceived as a violation of the right to be treated equally. Such differential lockdowns may seem particularly troublesome when the geographical region in question is mostly populated with ethnic minorities or disadvantaged groups, so that, even in the absence of intention to disadvantage the members of a particular group, the measures might be interpreted as "indirect discrimination" (cf. [47]).

Indeed, several lawsuits filed during the pandemic contested the legality of differential lockdowns, arguing discrimination and lack of evidence as to their effectiveness (e.g., Cases # 20–21). Another contentious question is the relevant *scale* for imposing differential measures (e.g., neighborhoods, cities, counties, etc.). For example, one lawsuit requested the court to order that a differential lockdown on a certain neighborhood be narrowed down to specific streets within that neighborhood (Case # 21).

The lens of complexity can shed light on questions pertaining to the legality of differential lockdowns, by providing a nuanced understanding of the spatiotemporal diffusion patterns of the pandemic.

The long-established SIR model for analyzing the spread of contagious diseases, traditionally assumed homogenous mixed populations [48]. However, complexity instructs that spatial diffusion in complex systems is usually non-homogenous. Rather, it often displays a spatial *fractal pattern* [49, 50]. In terms of visual display, systems whose spatial properties comprise a fractal tend to form *clusters* of high density, whereby objects concentrate in close spatial proximity to each other, surrounded by low-density areas (or "voids"), as displayed in **Figure 3a** (2). Fractal properties describe the spatial diffusion processes of a vast range of natural and social phenomena exhibiting contagious properties, including, for example, the growth of bacterial colonies [52], the evolution of cities [53], or the spread of local initiatives within the urban area [54]. The spatio-temporal diffusion of previous epidemics, such as the SARS-Cov in China and the MERS-Cov in the Middle East, displayed similar properties [55, 56].

Given these spatial properties of complex systems, it is not surprising that the spatial diffusion of Covid-19 in various countries is highly nonhomogeneous and tends to exhibit a fractal pattern. This pattern coincides with the spatial distribution of populations, and is characterized by clusters of cases, with local diffusion within the clusters [e.g., [57] (China); [51] (US)]. Some evidence suggests that this pattern also characterizes the global diffusion of the pandemic [58, 59][11].

Figures 3a,b, which describe the diffusion of Covid-19 in the United States, illustrate how this spatial pattern evolves over time.

As is apparent from **Figure 3**, during the initial period the spatial diffusion pattern is characterized by clusters of cases, which are geographically spread and relatively isolated from each

[9] As opposed to additive stochastic processes, which exhibit Gaussian distributions, multiplicative stochastic processes typically exhibit "fat tailed" power-law distributions of the form $P(x) \sim x^{-1-\alpha}$ (see, e.g., [38]), As noted above, if $\alpha \leq 1$ the mean of the distribution goes to infinity. In the context of an epidemic, such fat-tailed power law distribution implies that the expected harm is immense.

[10] These factors are similar to the guidelines we propose in the preceding Section, for applying the inverse proportionality test.

[11] For an attempt to integrate these properties into a modified SIR model, in order to predict the diffusion of Covid-19, see, e.g., [60].

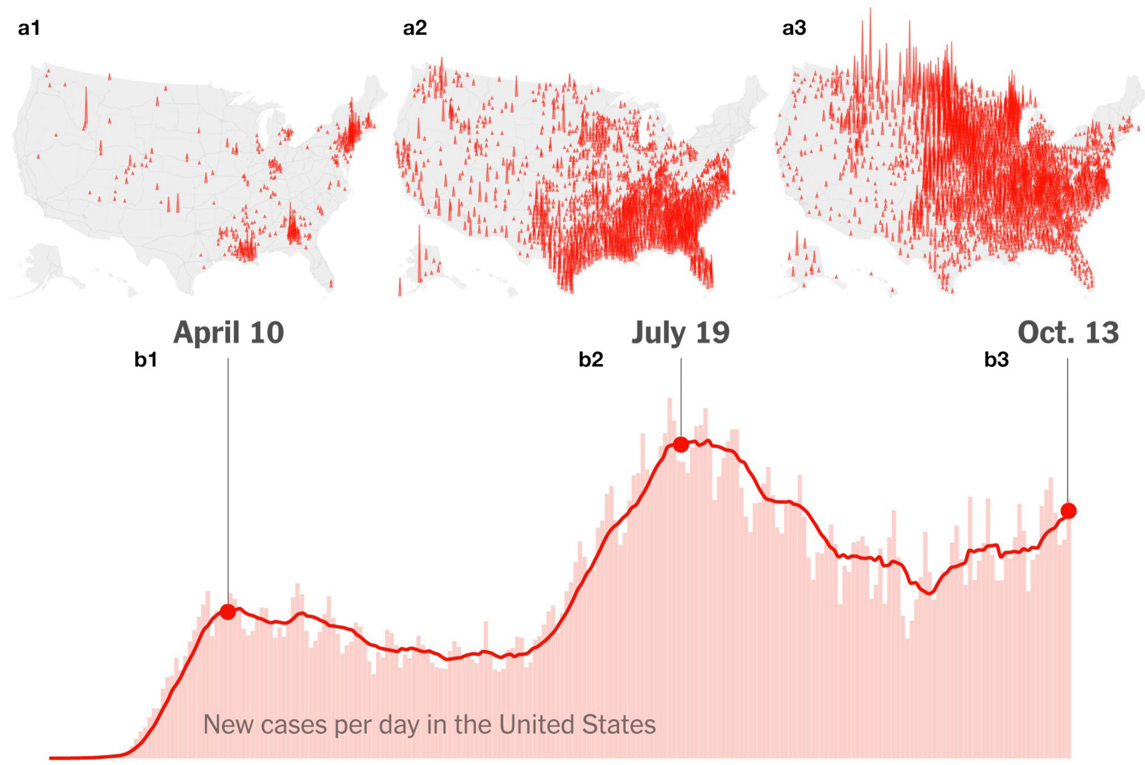


FIGURE 3 | (a) (1)–(3): Spatial Diffusion of Covid-19 cases in the United States, March-October 2020. **(b)** (1)–(3): Temporal Diffusion of Covid-19 the United States, March-October 2020 (Source: Leatherby, [51] New York Times, full figure available at: https://www.nytimes.com/interactive/2020/10/15/us/coronavirus-cases-us-surge.html).

other [**Figure 3a**(1)]. With the lapse of time and the increase in the numbers of cases the clusters expand from core to periphery and the "voids" among them narrow [**Figure 3a**(2)]. Eventually, when the pandemic is widespread, it covers most of the geographical area, and the "clustering" pattern is hardly observable [**Figure 3a**(3)][12].

Acquaintance with these spatiotemporal patterns has significant implications for the legality of differential policies. First, when planning interventions to curb a pandemic, regulators should assume that it will not spread in a homogenous way. Rather, they can rely on decades of complex systems research and expect, from early stages, that the pandemic's spatial diffusion will exhibit a fractal pattern (cf. [54]).

Secondly, the foregoing analysis implies that during particular stages differential geographical treatment can be highly efficient in curbing a pandemic. To illustrate, the Chinese policy that isolated Wuhan and its surrounding region, which suffered from large-scale infections, from other regions in which the number of Covid-19 cases was much lower, was successful in quickly containing the spatial diffusion of the pandemic in China [61]. Likewise, mathematical modeling of the spatial diffusion dynamics indicates that isolating infected regions ("clusters") while imposing social distancing measures shortly after experiencing community transmission, can lead to an exponential decrease in the number of infected regions, which may remain stable even after social distancing measures are lifted [62].

More specifically, understanding the spatiotemporal properties of the pandemic clarifies that the effectiveness of differential lockdowns, and the relevant scales for their implementation, are dynamic. During the very early stages of diffusion (or following a lockdown which brings down the number of infections), the fractal dimension of the entire pandemic (system) is practically zero [see, e.g., **Figure 3a**(1)][13]. This implies that measures for isolating infected clusters are highly effective, and the effective scale for such isolation can be rather small (e.g., a city, or even a neighborhood within a city). As the fractal dimension increases with the number of infected cases, the scale of isolated units must increase too in order to be effective. During that intermediate stage, the typical fractal dimension of the epidemic's spatial spread

[12]Notably, while **Figure 3a** is suggestive of the spatially inhomogeneous spread of a complex system, it is largely based on visual impression. Validating the fractal properties of the spatial diffusion pattern in the United States would require more ingrained spatio-temporal diffusion data, which we do not possess (cf. [54]).

[13]A fractal dimension is a non-integer number that expresses the dimension of objects exhibiting similar structures over a range of length scales. There are various methods to evaluate the fractal dimension of an empirical fractal, all based on multiple resolution analysis in which one measures a property P of the system (mass, volume, etc.) as a function of the yardstick used in measuring it (given by a yardstick of linear size r). Fractal objects are characterized by the power-law formula $P = kr^{-D}$, where D is the fractal dimension and k is a prefactor related to the lacunarity of the object (see, e.g., [63]).

would be $1 \leq D \leq 2$, and the spread would lack a natural scale [see, e.g., **Figure 3a**(2)]. Accordingly, the appropriate scale for differential measures is indeterminate and somewhat arbitrary. At that point the entire effectiveness of differential measures becomes doubtful, and global measures such as a nation-wide lockdown may be the more efficient choice. Finally, when the pandemic essentially spreads over the entire system and covers the relevant geographical area almost entirely, the fractal dimension roughly approaches the Euclidean (integer) dimension 2, [see, e.g., **Figure 3a**(3)][14]. During that stage differential geographical measures are ineffective and a general lockdown is likely inevitable.

Understanding the spatiotemporal diffusion of the pandemic can, therefore, provide regulators with tools to determine the *relevant timeframe* for implementing differential measures, and their *relevant scale*, and can be further used by courts reviewing the legality of such restrictions. To illustrate, a complexity based analysis supports the Israeli Supreme Court's rejection of a claim that a differential lockdown imposed on specific neighborhoods in Jerusalem during the first stages of the pandemic (in April 2020) was discriminatory and illegal. It also provides solid grounds to the Court's' refusal to narrow down the scale of the locked area to specific streets only, since the infected cases have already spread in various streets within that neighborhood (Case # 21).

NETWORK PRIORITIZING AND PRIVACY

The pandemic poses legal policy makers worldwide with acute challenges of prioritizing and dividing resources: how to allocate tests when testing capacity is limited? How to prioritize the distribution of vaccines among the population? Age, health risk factors, and medical-related occupation are frequently recognized as relevant considerations. Complexity provides policy makers with an additional tool for prioritizing, by using network analysis, which reveals, again, potential conflicts with legal doctrine.

Network analysis has become a prominent tool to describe and analyze complex systems of different kinds. By representing a relevant system as a network and mapping the links among the individuals comprising it ("nodes," in network parlance), network analysis allows to identify and describe various individual and systemic traits (e.g., [14, 64–66]). Two traits are particularly important for our purposes: "degree centrality," which describes the number of links that a particular node has (in our case: the number of in-person social ties a person has), and "betweenness centrality," which describes the extent to which a node—in our case: a person—"bridges" between different groups ("clusters") in the system [65][15].

Accumulating research of social networks from the past decades demonstrates that in-person networks typically exhibit power law distributions of social connections: a small number of people possess an exceptionally large number of social contacts, compared to the vast majority of the population (e.g., [13, 67])[16]. Studies further indicate that these "social hubs" play a crucial role in the effective diffusion of various objects through the network, from the flow of information to the spread of pandemics (e.g., [68–70]). To use an intuitive illustration, a person who meets hundreds of people per week is more likely to contract a contagious disease, and to spread it further, relative to someone whose network consists of only five weekly encounters. In the context of contagious diseases, then, high social connectivity can be regarded an additional risk factor.

More specifically, accumulating evidence on Covid-19 indicates that "super-spreading events" whereby "few individuals disproportionately infect a large number of secondary cases" are an extremely significant driver of the pandemic's diffusion (e.g., [71–73]). According to recent data, 10% of infectious individuals cause 80% of the Covid-19 infections [74]. Many super-spreading events have a spatial dimension, and occur in public gatherings, where large numbers of people concentrate in confined, typically indoor, spaces—weddings, places of worship, elderly homes, prisons, and meatpacking plants are a few prominent examples (e.g., [72, 73, 75]). Yet, significant spreading can also occur when an infected social hub contacts numerous people in different places during a short time span, even in the absence of large gatherings (cf. [72]). To illustrate, reports maintain that the first confirmed Covid-19 case of local transmission in the United States, also known as "patient zero," met with more than 800 people during a few days, prior to being diagnosed in February 2020. These numerous in-person encounters presumably triggered the subsequent spread of Covid-19 in the state of New York [76]. Similarly, studies suggest that musicians traveling among bars were a major driver in the spread of Covid-19 in Hong Kong between January and April 2020 [73].

Network studies from the past two decades further indicate that policies focused on locating, testing, and vaccinating highly connected individuals can significantly contribute to curbing pandemics. Christakis and Fowler [70], investigating the spread of flu among college students, found that tracking the health of more-connected individuals can provide more up-to-date information about the progress of the disease, relative to testing a random sample of people. Moreover, network models found that immunization schemes which prioritize inoculating highly-connected individuals can lower the entire network's vulnerability to a pandemic, and substantially decrease the threshold required for reaching herd immunity [67–69]. Notably, network prioritizing can also be used as part of the contemporary efforts to quickly identify and track new mutations of the virus.

[14] Notably, if the epidemic spreads freely, its eventual spatial diffusion will likely coincide (approximately) with the spatial distribution of the population, which is typically nonhomogeneous and exhibits fractal properties.

[15] A person's betweenness centrality is calculated by the number of shortest paths connecting all individuals in a network that pass through that person – (e.g.,[65], p. 334).

[16] We should clarify that, while multiplicative processes often generate power law distributions, our analysis is applicable to "fat tailed" distributions more generally, and does not depend on whether the empirically observed distribution is "truly" power law or other "fat-tailed" distribution that exhibits large skewness.

Research on network prioritizing has so far concentrated on people with high "degree centrality," namely people with a large number of in-person encounters. Our analysis suggests that an additional network metric that is relevant for network prioritizing in the context of Covid-19 is "betweenness centrality." Due to the spatial diffusion pattern of the disease, which we discussed in the previous Section, individuals with high betweenness centrality can significantly contribute to the transmission of the pandemic among clusters of different geographical locations, and among populations. This, in turn, might jeopardize differential-treatment policies. Overall, then, policies that take into account social connectivity metrics as a factor in prioritizing tests and vaccines can be an efficient tool in the effort to restrain the pandemic (cf. [67]).

From a law and policy perspective, implementing network prioritizing policies raises serious legal challenges, the most prominent of which concerns privacy. Such implementation obviously depends on the availability of connectivity data that enables to identify individuals with the highest degree centrality and betweenness centrality. However, the structure of the in-person social network is not readily observable. While information about highly connected individuals on online social networks ("influencers") is publicly available, this type of connectivity does not necessarily overlap with high numbers of in-person interactions, which are the relevant type of links in the case of pandemic transmission. Some relevant connectivity data can be inferred by identifying certain occupations, which entail a large number of in-person encounters (e.g., teachers, or salespersons), or transition between geographical areas (e.g., drivers of public transportation) (cf. [77]). Focusing on occupations can therefore provide a helpful prioritizing method, and is indeed adopted by some countries as part of their inoculation program. Nevertheless, it is insufficient for identifying other social hubs, such as New York's "patient zero."

Existing network studies extracted connectivity data from cellular companies, technology platforms (such as Facebook or Google) that collect mobility data, or epidemiological investigations (cf. [71, 73]). An alternative, indirect, strategy suggested by Cohen et al. [69] and applied by Christakis and Fowler [70] is the "friends" method. This strategy focuses on testing and treating the persons connected to a group of randomly selected individuals. This approach is based on the perception that, due to the structure of social networks, the "friends" group of randomly selected people possess, on average, more social contacts than the randomly selected group, so that focusing on the "friends" group increases the chances of reaching the social hubs [69]. While none of these methods is likely to yield perfect data about network structure, studies indicates that even incomplete data that identifies a sufficient number of hubs can make a significant difference in curbing a pandemic [67].

Nevertheless, the aforesaid methods for collecting and extracting information about individuals' connectivity in order to implement network prioritizing policies may conflict with legal rules protecting individual privacy. A prominent example is the European General Data Protection Regulations ("GDPR"), which limits the processing of "personal data," including, *inter alia*, "location data" and information related to a person's "social identity." The "repurposing" of data originally collected for a specific purpose (for example, targeted advertising), and using it for a different purpose (in our case: identifying social connectivity) is similarly restricted[17].

While privacy concerns should not be underestimated, the case of Covid-19 highlights the networked dimensions of privacy, and their theoretical and practical implications for privacy law. Traditional legal theory perceives privacy through a reductionist, individualistic lens—a person's "right to be let alone" [79], a right that "rests upon an individualist concept of society" ([80], p. 958). This reductionist paradigm frames privacy as a private good, implying that its costs and benefits are confined to a particular individual. Therefore, the principal legal regimes that protect privacy empower individual control, and place substantial weight on individual consent ([81], p. 390).

More recent theoretical accounts, however, recognize that framing privacy as a strictly individual interest "ignores the interconnected nature of human behavior and of human interests" [82]. Broadly, this strand of literature highlights that privacy has collective and social attributes: privacy-related decisions of individuals can produce "privacy externalities" and affect other people. It therefore suggests that in certain circumstances privacy protection should be regulated, rather than left to individual choice and consent [81, 83].

The discussion of privacy externalities in the legal literature often concentrates on cases where individuals' decision to renounce their privacy may cause harm to other individuals or to society at large. For example, sharing one's genetic information may harm the privacy of their relatives [84], while the sharing of data with social media platforms may reduce the overall level of privacy in a society [81, 83]. In the foregoing analysis privacy externalities operate in an opposite way. In other words, an individual decision to *maintain* their privacy (by not disclosing certain information about social connectivity) may impose significant social costs on the effort to curb a pandemic. And due to the structure of the social network and the pandemic diffusion dynamics that we discuss above, the costs entailed in maintaining connectivity information private accumulate in a multiplicative, nonlinear way.

An appropriate theoretical framework for addressing the networked dimensions of privacy in the context of Covid-19 could be found in the literature that understands privacy as a mechanism for encouraging socially beneficial flows of personal information. Most prominently, Nissenbaum's influential theory of privacy as "contextual integrity" [85, 86] perceives privacy not in terms of individual control or consent, but rather in terms of flows of personal information that are normatively appropriate in a particular social context. Contextual integrity theory thus avoids binary and dogmatic distinctions (e.g., between sensitive vs. insensitive information, or between prohibited vs. permitted uses). Rather, it recognizes

[17] Regulation (EU) [78], Art. 4(1), and Art. 5(1)(b), available at https://gdpr-info.eu/. Notably, alongside these restrictions, privacy legislation typically recognizes that public health concerns may justify certain exceptions to privacy protection – e.g., Art. 9(2)(i).

that transferring certain information (in our case: information about social connectivity) that may not be permitted in one context, may be permitted, and even required, in another context (e.g., to specific health authorities that need to prioritize Covid tests). This theory may therefore provide tools for incorporating the use of network science and its potential benefits in prioritizing Covid-19 interventions, into privacy analysis.

We do not suggest here a detailed regime for the use of connectivity data. However, understanding the significance of network prioritizing in responding to a pandemic, alongside the networked nature of privacy, can assist policy makers and judiciaries in devising such schemes, and in addressing ostensible tensions between privacy and (public and private) health in specific circumstances. More broadly, this analysis reveals that the law is not always sufficiently sensitive to *interactions* as a relevant factor in the design of legal policies. Complexity theory indicates that it should be.

CONCLUSIONS

The Covid-19 pandemic represents a great challenge for law and policy makers. The analysis above demonstrates that acquaintance with the properties of complex systems can provide regulators and judiciaries with an extremely valuable tool for devising and evaluating the response to Covid-19. Concomitantly, it exposes a gap between the principles governing complex systems, and extant legal doctrines, such as proportionality, equality, evidentiary requirements, and privacy, which implicitly presume stable equilibria, and fail to capture the nonlinear multiplicative properties of the pandemic. Embedding insights from complexity theory into legal analysis will thus help align legal policies with complex systems dynamics, and improve the legal response to the pandemic.

Our study is non-exhaustive, and is limited in various respects. Adopting a complexity perspective can likely illuminate additional questions related to the Covid-19 legal response. For example, while we concentrate on court cases and judicial doctrine, legal interventions through administrative regulation may be more flexible and therefore more apt to effectively adapt to the fast-changing realities of the pandemic (cf. [33, 87])[18].

More broadly, our present focus is confined to situations in which nonlinearities associated with complex systems can lead to instability and cascading disasters, and does not cover all traits associated with such systems. Furthermore, the details of our analysis are specific to the case of Covid-19, and the proposals we make may not be applicable "as such" to all other cases concerning complex phenomena, and may require adaptation[19]. Nevertheless, our analysis calls for applying the "lens of complexity" to additional legal policies related to complex nonlinear collective phenomena, from global environmental challenges to financial crises, and for detailed explorations of particular legal responses. More research in this vein will hopefully follow.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

Both authors contributed equally to the conceptualization, analysis, and drafting of this research, and approved it for publication.

FUNDING

This research was supported by ISF grant number 1322/18.

ACKNOWLEDGMENTS

We are grateful to the two referees of this paper for their extremely valuable suggestions. For helpful comments and discussions we thank Yaneer Bar-Yam, Doron Teichman, Eyal Zamir, and the participants of the Physics of Law conference (November 2020, online), and the Cyber Center Workshop at the Hebrew University of Jerusalem (January 2021, online). Stav Cohen and Roi Ohayon provided excellent research assistance.

[18]Obviously, such regulations, too, will eventually be subject to judicial scrutiny.
[19]To illustrate, not all complex phenomena diffuse at a pace similar to a pandemic (cf. [54]), a factor which may be significant for assessing proportionality.

REFERENCES

1. Bavel JJV, Baicker K, Boggio PS, Capraro V, Cichocka A, Cikara M, et al. Using social and behavioural science to support COVID-19 pandemic response. *Nat Hum Behav*. (2020) 4:460–71. doi: 10.1038/s41562-020-0884-z
2. Teichman D, Underhill K. Infected by bias: behavioral science and the legal response to COVID-19. *Am J Law Med*. Hebrew University of Jerusalem Legal Research. (2020). doi: 10.2139/ssrn.3691822
3. Zamir E, Teichman D. *Governmental Decision-Making Regarding the Covid-19 Pandemic—a Behavioral Perspective*. SSRN (2020). Available online at: https://ssrn.com/abstract=3698682 (accessed January 3, 2021).
4. Ruhl JB. Complexity theory as a paradigm for the dynamical law-and-society system: a wake-up call for legal reductionism and the modern administrative state. *Duke Law J*. (1996) 45:849–928. doi: 10.2307/1372975
5. Shur-Ofry M. IP and the lens of complexity. *IDEA*. (2013) 54:55–102. Available online at: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2268867
6. Ruhl JB, Katz D. Measuring, monitoring, and managing legal complexity. *Iowa Law Rev*. (2015) 101:191–244. Available online at: https://scholarship.kentlaw.iit.edu/fac_schol/865
7. Fallon RH. Strict judicial scrutiny. *UCLA Law Rev*. (2007) 54:1267–338. Available online at: https://www.uclalawreview.org/strict-judicial-scrutiny/
8. Jackson VC. Constitutional law in an age of proportionality. *Yale Law J*. (2015) 124:3094–196. Available online at: https://www.yalelawjournal.org/article/constitutional-law-in-an-age-of-proportionality
9. Note (1969). Less drastic means and the first amendment. *Yale Law J*. 78:464–74. doi: 10.2307/794878

10. Sykes AO. The least restrictive means. *U Chi Law Rev*. (2003) 70:403–19. doi: 10.2307/1600566
11. Brems E, Lavrysen L. 'Don't use a sledgehammer to crack a nut': less restrictive means in the case law of the European court of human rights. *Hum Rts Law Rev. (*2015) 15:139–68. doi: 10.1093/hrlr/ngu040
12. Bass F. A new product growth for model consumer durables. *Manag Sci*. (1969) 15:215–27. doi: 10.1287/mnsc.15.5.215
13. Rogers EM. *Diffusion of Innovations*. 5th. ed. New York, NY: Simon and Schuster (2003). 576 p.
14. Cohen R, Havlin S. Complex networks: structure, robustness and function. New York, NY: Cambridge University Press (2010). 238 p.
15. Stevens R, Perera R, Heneghan C, Hobbs R, Oke J. *Exponential Growth: What It Is, Why It Matters, and How to Spot It*. CEBM (2020). Available online at: https://www.cebm.net/wp-content/uploads/2020/09/exponential-growth (accessed January 3, 2021).
16. Stango V, Zinman J. Exponential growth bias and household finance. *J. Finance*. (2009) 64:2807–49. doi: 10.1111/j.1540-6261.2009.01518.x
17. Teichman D, Zamir E. *Mathematics, Psychology, and Law: The Legal Ramifications of the Exponential Growth Bias*. (2020). Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3804329
18. *Covid-19: China's Qingdao to Test Nine Million in Five Days*. BBC News (2020). Available online at: https://www.bbc.com/news/world-asia-54504785 (accessed January 3, 2021).
19. Jamieson T. "Go hard, go early": Preliminary lessons from New Zealand's response to COVID-19. *Am Rev Public Adm.* (2020) 50:598–605. doi: 10.1177/0275074020941721
20. *Israel Plans to Halt All International Flights as Part of Nationwide Shutdown News18.* (2020). Available online at: https://www.news18.com/news/auto/israel-plans-to-halt-all-international-flights-as-part-of-nationwide-shutdown-2875653.html. (accessed January 3, 2021).
21. Sanderson I. Evaluation, policy learning and evidence-based policy making. *Public Admin*. (2002) 80:1–22. doi: 10.1111/1467-9299.00292
22. Gunn TJ. Deconstructing proportionality in limitations analysis. *Emory Int. L. Rev*. (2005) 19:465–98.
23. Ioannidis JPA. *A Fiasco in the Making? As the Coronavirus Pandemic Takes Hold, We Are Making Decisions Without Reliable Data*. Statistics (2020). Available online at: https://www.statnews.com/2020/03/17/a-fiasco-in-the-making-as-the-coronavirus-pandemic-takes-hold-we-are-making-decisions-without-reliable-data/ (accessed January 3, 2021).
24. Shaw R. Modeling chaotic systems. In: Haken H, editor. *Chaos and Order In Nature*. Berlin, Heidelberg: Springer. (1981). p. 218–31.
25. Lorenz EN. Deterministic nonperiodic flow. *J Atmos Sci.* (1963) 20:130–141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2
26. Prigogine I. Exploring complexity. *Eur J Oper Res*. (1987) 30:97–103. doi: 10.1016/0377-2217(87)90085-3
27. Prigogine I, Allen P. The challenge of complexity. In: Schieve WC, Allen PM, editors. *Self-Organization and Dissipative Structures*. Texas: University of Texas Press (1982). p. 3–39.
28. Siegenfeld AF, Taleb NN, Bar-Yam Y. Opinion: what models can and cannot tell us about COVID-19. *PNAS*. (2020) 117:16092–5. doi: 10.1073/pnas.2011542117
29. Castro M, Ares S, Cuesta JA, Manrubia S. Predictability: can the turning point and end of an expanding epidemic be precisely forecast? *arXiv*. (2020). doi: 10.1073/pnas.2007868117
30. Taleb NN, Bar-Yam Y, Cirillo P. On single point forecasts for fat-tailed variables. *arXiv preprint arXiv:2004.08842.* (2020). doi: 10.1016/j.ijforecast.2020.08.008
31. Cirillo P, Taleb NN. Tail risk of contagious diseases. *Nat Phys*. (2020) 16:606–13. doi: 10.1038/s41567-020-0921-x
32. Fukui M, Furukawa C. Power laws in superspreading events: evidence from Coronavirus outbreaks and implications for SIR models. *medRxiv*. (2020). doi: 10.1101/2020.06.11.20128058
33. Ruhl JB. General design principles for resilience and adaptive capacity in legal systems-with applications to climate change adaptation. *North Carolina Law Rev*. (2011) 89:1373–401. Available online at: https://scholarship.law.unc.edu/cgi/viewcontent.cgi?article=4506&context=nclr
34. Read R, O'Riordan T. The precautionary principle under fire. *Environ Sci Pol Sustain Dev*. (2017) 59:4–15. doi: 10.1080/00139157.2017.1350005
35. Helsingen LM, Løberg M, Refsum E, Gjøstein DK, Wieszczy P, Olsvik O, et al. Randomized re-opening of training facilities during the COVID-19 pandemic. *medRxiv*. (2020). doi: 10.1101/2020.06.24.20138768
36. Teichman D, Zamir E. *How do you Cause Everyone to Start Wearing Facemasks: Nudges, Mandates, and the Legal Response to the Covid-19 Pandemic in Israel*. SSRN (2020). Available online at: https://ssrn.com/abstract=3698731 (accessed January 3, 2021).
37. Allen R. Taming complexity: rationality, the law of evidence and the nature of the legal system. *Law Probability Risk*. (2013) 12:99–113. doi: 10.1093/lpr/mgs008
38. Malcai O, Biham O, Solomon S. Power-law distributions and Levy-stable intermittent fluctuations in stochastic systems of many autocatalytic elements. *Phys Rev E*. (1999) 60:1299. doi: 10.1103/PhysRevE.60.1299
39. *The Wingspread Statement on the Precautionary Principle.* (1998). Available online at https://www.gdrc.org/u-gov/precaution-3.html (accessed January 3, 2021).
40. *Report of the United Nations Conference on Environment and Development*. (" Rio Declaration"), Principle 15 Rio de Janeiro (1992). Available online at: https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf (accessed January 3, 2021).
41. Nash JR. Standing and the precautionary principle. *Colum Law Rev*. (2006) 108:494–528. Available online at: https://chicagounbound.uchicago.edu/public_law_and_legal_theory/280/
42. Sunstein CR. Beyond the precautionary principle. *U Pa Law Rev*. (2003) 151:1003–58. doi: 10.2307/3312884
43. Heinzerling L. Reductionist regulatory reform. *Fordham Envtl Law J.* (1997) 8:459.
44. Ruhl JB. Working both (positivist) ends toward a new (pragmatist) middle in environmental law. *Geo Wash Law Rev*. (2000) 68:522. doi: 10.2139/ssrn.214528
45. Gleick PH. Climate change, exponential curves, water resources, and unprecedented threats to humanity. *Clim Change.* (2010) 100:125–9. doi: 10.1007/s10584-010-9831-8
46. Trisos CH, Merow C, Pigot AL. The projected timing of abrupt ecological disruption from climate change. *Nature*. (2020) 580:496–501. doi: 10.1038/s41586-020-2189-9
47. Delan D, Shannon G, Bhopal SS, Abubakar I. Racism and discrimination in COVID-19 responses. *Lancet*. (2020) 395:1194. doi: 10.1016/S0140-6736(20)30792-3
48. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. *Proc R Soc London.* (1927) 115:700–21. doi: 10.1098/rspa.1927.0118
49. Mandelbrot B. *The Fractal Geometry of Nature*. San Francisco, CA: W.H. Freeman (1982). 460 p.
50. Bak P, Tang C, Wiesenfeld K. Self-organized criticality. *Phys Rev A*. (1988) 38:364–74. doi: 10.1103/PhysRevA.38.364
51. Leatherby L. *U.S. Virus Cases Climb Toward a Third Peak*. NY Times (2020). Available online at: https://www.nytimes.com/interactive/2020/10/15/us/coronavirus-cases-us-surge.html (accessed January 3, 2021).
52. Ben-Jacob E, Cohen I, Shochet O, Aranson I, Levine H, Tsimring L. Complex bacterial patterns. *Nature*. (1995) 373:566–7. doi: 10.1038/373566a0
53. Batty M. *Cities and Complexity*. Cambridge: MIT Press (2007). 565 p.
54. Shur-Ofry M, Malcai O. Institutions for collective action and social contagion: community gardens as a case study. *Regul Govern.* (2021) 15:63–81. doi: 10.1111/rego.12256
55. Fang LQ, De Vlas SJ, Feng D, Liang S, Xu YF, Zhou JP, et al. Geographical spread of SARS in mainland China. *Trop Med Int Health*. (2009) 14:14–20. doi: 10.1111/j.1365-3156.2008.02189.x
56. Al-Ahmadi K, Alahmadi S, Al-Zahrani A. Spatiotemporal clustering of Middle East Respiratory Syndrome Coronavirus. (MERS-CoV) incidence in Saudi Arabia, 2012–2019. *Int J Environ Res Public Health*. (2019) 16:2520. doi: 10.3390/ijerph16142520
57. Kang D, Choi H, Kim JH, Choi J. Spatial epidemic dynamics of the COVID-19 outbreak in China. *Int J Infect Dis*. (2020) 94:96–102. doi: 10.1016/j.ijid.2020.03.076

58. Jiang B, de Rijke C. *A Power-Law-Based Approach to Mapping COVID-19 Cases in the United States.* (2020). Available online at: https://www.preprints.org/manuscript/202007.0607/v1 (accessed January 3, 2021).
59. Yeşilkanat C. M. Spatio-temporal estimation of the daily cases of COVID-19 in worldwide using random forest machine learning algorithm. *Chaos Solitons Fractals.* (2020) 140:110210. doi: 10.1016/j.chaos.2020.110210
60. Huang J, Zhang L, Liu X, Yun W, Liu C, Lian X, et al. Global prediction system for COVID-19 pandemic. *Sci Bull.* (2020) 65:1884–7. doi: 10.1016/j.scib.2020.08.002
61. Long Y, Chen Y, Li Y. Multifractal scaling analyses of the spatial diffusion pattern of COVID-19 pandemic in Chinese mainland. *arXiv:2010.02747* (2020).
62. Siegenfeld AF, Bar-Yam Y. Eliminating COVID-19: the impact of travel and timing. *arXiv:2003.10086* (2020).
63. Malcai O, Lidar DA, Biham O, Avnir D. Scaling range and cutoffs in empirical fractals. *Phys Rev E.* (1997) 56:2817–28. doi: 10.1103/PhysRevE.56.2817
64. Strandburg KJ, Csárdi G, Tobochnik J, Érdi P, Zalányi L. Law and the science of networks: an overview and an application to the "patent explosion". *Berkeley Tech Law J.* (2006) 21:1293. doi: 10.2139/ssrn.926354
65. Barabási AL. *Network Science*. Cambridge: Cambridge University Press (2016). 456 p.
66. Borgatti SP, Everett MG, Johnson JC. *Analyzing Social Networks*. 2nd ed. Los Angeles: SAGE Publications (2018). 363 p.
67. Dezso Z, Barabasi AL. Halting viruses in scale-free networks. *Phys Rev E.* (2002) 65:055103. doi: 10.1103/PhysRevE.65.055103
68. Pastor-Satorras R, Vespignani A. Immunization of complex networks. *Phys Rev E.* (2002) 65:036104. doi: 10.1103/PhysRevE.65.036104
69. Cohen R, Havlin S, Ben-Avraham D. Efficient immunization strategies for computer networks and populations. *Phys Rev Lett*. (2003) 91:247901. doi: 10.1103/PhysRevLett.91.247901
70. Christakis NA, Fowler JH. Social network sensors for early detection of contagious outbreaks. *PLoS ONE.* (2010) 5:e12948. doi: 10.1371/journal.pone.0012948
71. Lau MS, Grenfell B, Thomas M, Bryand M, Nelsone K, Lopmane B. Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. *PNAS.* (2020) 117:22430–5. doi: 10.1073/pnas.2011802117
72. Kuperschmidt K. Why do some COVID-19 patients infect many others, whereas most don't spread the virus at all? *Science*. (2020). doi: 10.1126/science.abc8931
73. Dillon CA, Wu P, Wong JY, Lau EHY, Tsang TK, Cauchemez S, et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat Med*. (2020) 26:1714–9. doi: 10.1038/s41591-020-1092-0
74. Endo A, Kucharski AJ, Funk S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res.* (2020) 5:67. doi: 10.12688/wellcomeopenres.15842.3
75. Leclerc QJ, Fuller NM, Knight LE, Funk S, Knight GM. What settings have been linked to SARS-CoV-2 transmission clusters? *Wellcome Open Res*. (2020) 5:83. doi: 10.12688/wellcomeopenres.15889.1
76. Cox C. The vulnerable can wait. *Vaccinate the Super-Spreaders First, WIRED*. (2020). Available online at: https://www.wired.com/story/covid-19-vaccine-super-spreaders/
77. Shur-Ofry M, Malcai O. *Using Network-Science Tools for Combatting Covid-19: Prioritizing Tests and Vaccines*. Memorandum to the Israeli Health Ministry (2020).
78. Regulation (EU) 2016/679 (General Data Protection Regulation) OJL 119, 04.05.2016; cor. OJL 127, 23.5.2018).
79. Warren SD, Brandeis LD. The right to privacy. *Harvard Law Rev*. (1890) 4:193–220. doi: 10.2307/1321160
80. Post RC. The social foundations of privacy, community and self in the common law tort. *Calif Law Rev*. (1989) 77:957–1010. doi: 10.2307/3480641
81. Fairfield JA, Engel C. Privacy as a public good. *Duke Law J*. (2015) 65:385–457. Available online at: https://scholarship.law.duke.edu/dlj/vol65/iss3/1
82. Bamberger KA, Evan Mayse A. *Privacy in Society: Jewish Law Insights for the Age of Big Data.* SSRN (2020). Available online at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3731770 (accessed January 3, 2021).
83. MacCarthy M. New directions in privacy: disclosure, unfairness and externalities. *ISJLP*. (2011) 6:425–512. Available online at: https://core.ac.uk/download/pdf/159597718.pdf
84. Ram N. DNA by the entirety. *Colum Law Rev*. (2015). 115:873–940. Available online at: https://columbialawreview.org/wp-content/uploads/2016/02/May-2015-14-Article-Ram_UPDATE-5.17.pdf
85. Nissenbaum H. Privacy as contextual integrity. *Washington Law Rev.* (2004) 79:119–58. Available online at: https://digitalcommons.law.uw.edu/wlr/vol79/iss1/10
86. Barth A, Datta A, Mitchell JC, Nissenbaum H. Privacy and contextual integrity: framework and applications. In: *2006 IEEE Symposium on Security and Privacy (S&P'06)*. Berkeley/Oakland, CA (2006). 198 p.
87. Arnold CA, Gunderson LH. Adaptive law and resilience. *Env Law Rep.* (2013) 43:10426. Available online at: https://www.researchgate.net/profile/Tony-Arnold-2/publication/257109831_Adaptive_Law_and_Resilience/links/53fe22680cf21edafd14f3f5/Adaptive-Law-and-Resilience.pdf

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX

Court Cases

1. United States v. Raia, No. 20-1033, 2020 WL 1647922 (3d Cir. Apr. 2, 2020) (US)
2. United States v. Kennedy, No. 18-20315, 2020 WL 1547878 (E.D. Mich. Apr. 1, 2020) (US)
3. United States v. Rivers, No. CR 2017-0023, 2020 WL 1676798 (D.V.I. Apr. 6, 2020) (US)
4. Legacy Church, Inc. v. Kunkel, No. CV 20-0327 JB\SCY, 2020 WL 3963764 (D.N.M. July 13, 2020)(US)
5. S. Bay United Pentecostal Church v. Newsom, 140 S. Ct. 1613, 207 L. Ed. 2d 154 (2020) (US)
6. S. Bay United Pentecostal Church v. Newsom, 592 U. S. <uscore> (2021)
7. Roman Catholic Diocese v. Cuomo, 208 L. Ed. 2d 206 (S. Ct., Nov. 25, 2020)(US)
8. Roman Catholic Diocese v. Cuomo, 2020 U.S. Dist. LEXIS 192292 (E.D.N.Y, Oct. 16, 2020)(US)
9. Calvary Chapel Dayton Valley v. Sisolak, 140 S. Ct. 2603 (July 24, 2020)(US)
10. Calvary Chapel Dayton Valley v. Sisolak, 2020 U.S. Dist. LEXIS 103234 (D. Nev., June 11, 2020)(US)
11. Rock House Fitness, Inc. v. Acton, No. 20CV000631, 2020 WL 3105522 (Ohio Com.Pl. May 20, 2020) (US)
12. LMV DEV SPE, LLC, DBA Kalahari Resorts & Conventions v. Amy Acton, No. 2020-CV-0201 (Erie Cty. Com.Pl. July 12, 2020) (US)
13. Sprague v. Her Majesty the Queen in right of Ontario, 2020 ONSC 2335, 317 A.C.W.S (3d) 352 (Can. Ont.)
14. HMQ v. A.C., 2020 ONSC 2870 (Can. Ont.)
15. Children's Aid Society of the Region of Halton v. R.O., 2020 ONCJ 209, 317 A.C.W.S (3d) 271 (Can. Ont.)
16. Reyno Dawid De Beer v. The Minister of Cooperative Governance and Traditional Affairs, No. 21542/2020, ZAGPPHC 184 (June 2, 2020) (S. Afr.)
17. Conseil d'Etat [CE] [Council of State] May 18, 2020, M.W et auters, No. 440366 (France)
18. VG Regensburg, Apr. 27, 2020, RO 14 E 20.687, https://www.vgh.bayern.de/media/vgregensburg/presse/20a00687b.pdf (Germany)
19. S.T.S.J. Madrid, Oct. 8, 2020 (R.O.J. No. 308/2020)(Spain)
20. HCJ 2435/20 Yedidya Loewenthal, Adv. v. Prime Minister Benjamin Nethanyahu (Apr. 7, 2020), available at Nevo Legal Database (by subscription, in Hebrew) (Isr.)
21. HCJ 2491/20 Ramot Alon Community Administration v. The Government of Israel, Prime Minister (Apr. 14, 2020), available at Nevo Legal Database (by subscription, in Hebrew) (Isr.)

frontiers in Physics

ORIGINAL RESEARCH
published: 12 May 2021
doi: 10.3389/fphy.2021.657607

Check for updates

Analysing EU Treaty-Making and Litigation With Network Analysis and Natural Language Processing

Michal Ovádek, Arthur Dyevre and Kyra Wigard*

KU Leuven Centre for Empirical Jurisprudence, Leuven, Belgium

We apply network analysis and topic modeling techniques to explore the evolution of the European Union's treaty making activity and the patterns of litigation they have given rise to. Our analysis reveals that, despite the expansion of the bloc's policy remit, its treaty-making activity retains a strong economic focus. Among the many agreements negotiated by EU institutions, the European Economic Agreement, the Ankara Agreement with Turkey and the World Trade Organization Agreement form the largest clusters of litigated cases. EU international agreements are disproportionately litigated in cases pertaining to residence rights and competition law.

Keywords: network analysis, natural language processing, topic modeling, international agreements, European union, litigation

OPEN ACCESS

Edited by:
J. B. Ruhl,
Vanderbilt University, United States

Reviewed by:
Satyam Mukherjee,
Indian Institute of Management Udaipur, India
Michele Bellingeri,
University of Parma, Italy

****Correspondence:***
Arthur Dyevre
arthur.dyevre@kuleuven.be

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *25 January 2021*
Accepted: *01 April 2021*
Published: *12 May 2021*

Citation:
Ovádek M, Dyevre A and Wigard K (2021) Analysing EU Treaty-Making and Litigation With Network Analysis and Natural Language Processing. Front. Phys. 9:657607. doi: 10.3389/fphy.2021.657607

1. INTRODUCTION

Negotiations with the United Kingdom over the post-Brexit deal have highlighted the role of the European Union (EU) as a global treaty-making powerhouse. Member states have delegated treaty-making powers to EU institutions over an expanding set of policy domains, starting with trade, and later matters of defense and security. While the EU is not a state, it is a major international actor and a full member of the World Trade Organization (WTO).

The Brexit negotiations further highlighted the sensitivity of the issue of judicial review. The Leave campaign, including Boris Johnson himself, explicitly referred to the European Court of Justice (ECJ), as a reason for withdrawing from the bloc and the UK government insisted on keeping the ECJ out of the trade deal [1]. (It was eventually decided that legal disputes would be entrusted to an *ad hoc* arbitration panel.) This reflects the broader phenomenon that litigation, by bringing judges into the picture, can decisively influence the effective operation of international treaties.

While treaty negotiation dynamics and aspects of the ECJ's case law have received attention from political scientists and legal scholars [2–4], it is difficult to get a general sense of the variety of agreements and the extent to which they have given rise to litigation.

The sheer number of agreements and EU court cases rules out manual analysis. So we attempt to provide such an overview using machine learning and network analysis methods. We use probabilistic topic modeling to analyse the contents of EU international agreements and network analysis to identify the main clusters of citations to international agreements.

What our data exploration reveals is that, despite the expansion of the bloc's policy remit, its treaty-making activity retains a strong economic focus. It also shows that, among the many agreements negotiated by EU institutions, the European Economic Agreement, the Ankara Agreement with Turkey and the WTO Agreement form the largest clusters of litigated cases. EU international agreements are disproportionately litigated in cases relating to residence rights and, more surprisingly, competition law, while the opposite is true in cases relating to internal market themes such as public procurement and VAT.

2. RELATED WORK

Our paper relates to the growing literature applying machine learning and Natural Language Processing (NLP) methods to the study of law and legal documents [5–7]. It also relates to the legal and political science literature applying network analysis methods to the analysis of case citation dynamics [8–11] as well as the evolution and structure of legislation and networks of judges and law professors [12–14].

3. DATA

First, we collected data on EU international agreements[1] (N $\approx 10,000$), including full texts where available in English, from the `EUR-Lex` website, the official EU legal database, using a dedicated data collection R package [15]. `EUR-Lex` is well-curated and the data can be assumed to be close to complete, if not so.

Although international agreements can take on various forms, from formal treaties to agreements made through letters, we distinguish two main categories of legal acts based on the metadata in the database. The first comprises agreements in the form of stand-alone documents, regardless of whether these take the form of a treaty, a formalized exchange of letters or of a new protocol to a preceding agreement. The second category is formed by "joint decisions," which are acts produced by a body which was itself set up under a pre-existing international agreement.

Figure 1 shows that joint decisions account for an increasing number of new international legal texts. Ovádek and Raina [16] explain that this trend is likely grounded in the heightened ambition and scope of EU international agreements, most notably exemplified by the Agreement on the European Economic Area which de facto extends the EU internal market to Norway, Iceland and Liechtenstein. Joint decisions then serve to deal with various technical issues arising from the operation of the legal relationship. This governance model has been applied in many EU international agreements. In subsequent analysis, we focus predominantly on the stand-alone agreements rather than joint decisions, as these constitute international agreements in the stricter sense.

To explore litigation patterns involving international agreements, we gathered the entire universe of rulings rendered by EU courts[2] up to 2020 (N $\approx 26,000$). The collected metadata included information on the legal acts cited, which we used to identify citations to EU international agreements.

[1]EU international agreements are those agreements to which the EU is a party in its own right. This excludes at present some well-known international documents such as the European Convention on Human Rights and Fundamental Freedoms.
[2]Historically, there have been three EU courts: the European Court of Justice (since 1953), the General Court (founded in 1988 as the Court of First Instance) and the Civil Service Tribunal (established in 2005 and dissolved in 2016).

Figure 2 shows the proportion of EU court cases that contain a reference to at least one international act. We see a steady rise until about 2010 when the trend reverses. Historically, cases referring to EU international agreements account for around 5% of the case law.

4. METHODOLOGY

4.1. Topic Modeling

Topic modeling is a suite of unsupervised document-clustering techniques developed to generate thematic annotations automatically, whereby topics are modeled as probability over words and documents as probability over topics [17, 18]. We use the structural topic model developed by [19] and implemented in the `stm` package for R to generate topics measuring issue attention in international agreements. The implementation builds on the Correlated Topic Model developed by [20].

This topic modeling approach is preferred over the more conventional Latent Dirichlet Allocation (LDA) for three reasons. First, we want to account for temporal variations in the number and thematic focus of agreements and legal disputes. Whereas, LDA is oblivious to the order in which documents appear in the corpus, we group documents by year, whereby topics in year t are assumed to have evolved from to topics prevalent in year $t-1$. Dynamic topic models have been shown to better fit temporal dynamics in issue attention than LDA variants [21]. Third, our approach ensures that the resulting topics are not overly skewed toward years with more documents, which can occur when there are significant variations in the number of documents over time—which is the case for both EU international agreements and EU court cases. Third, simultaneously with temporal changes, we want to compare topic prevalence in cases citing international agreements to cases where no such reference is made. Our approach allows to model this difference directly.

To allow an assessment of temporal dynamics, we specify a covariate interacting with topic prevalence:

$$\theta_{1:D}|t_{1:D}\gamma, \Sigma \sim LogisticNormal\mu = t_{1:D}\gamma, \Sigma). \qquad (1)$$

where t_d is the year in which document d was issued; γ is a $p \times (K-1)$ matrix of coefficients for topic proportion and Σ is a $(K-1) \times (K-1)$ covariance matrix.

To investigate the variance in topic proportion between cases citing EU international agreements and cases containing no such reference, we estimate a dynamic topic model of EU court cases in which also specify a dummy variable capturing reference to international agreements.

To calculate topic proportion conditional on covariates, the method originally implemented in the `stm` package relied on OLS regression. To constrain topic proportion within the $(0,1)$ interval, we model topic proportion conditional on covariates using quasi-binomial regression.

Our text-mining approach is based on the bag-of-words paradigm. Accordingly, punctuation, numbers, html tags, rare words, and words common to many documents (including stop-words) were removed from the raw texts—all these are standard pre-processing steps in bag-of-words studies.

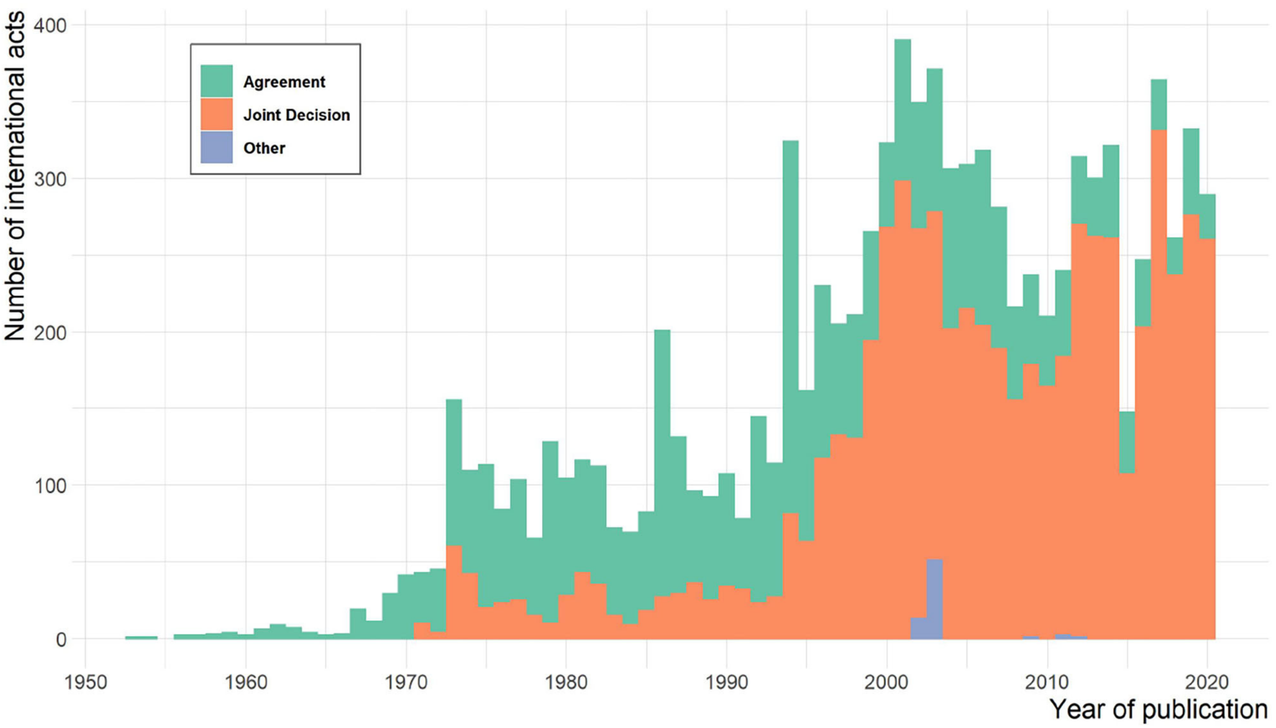


FIGURE 1 | Categories of international acts in the dataset (1953–2020).

4.2. Network Analysis

We use network analysis to model patterns of citations in EU court cases. Recent years have seen network analysis proliferate in social sciences and legal studies [9, 10, 13, 22]. In network analysis, a network consists of nodes (also known as "vertices") and edges (or links). In our analysis, a node is either an EU court case or an EU international agreement while edges represent citations either to an agreement or to another case. As with applications of network analysis to citation patterns in judicial opinions, we model cases, agreements and citations as directed networks. The directed nature of our legal citation networks results from the fact that agreements do not cite cases while a new case can only cite an older case[3].

To generate our citation networks, we construct an adjacency matrix of court rulings and EU international agreements. A node is adjacent to another if an edge connects them. Formally, if U is the set of all nodes $u_1, ..., u_n$ in a network, the adjacency matrix A_{ij} is a square $n \times n$ matrix connecting nodes u_i and u_j. The elements of the matrix take on value one if two nodes are adjacent and zero otherwise.

Our analysis is primarily concerned with the network centrality of agreements and cases citing agreements. A basic measure of the importance of a treaty or precedent in judicial opinions is in-degree centrality [8, 9, 23]. In-degree centrality simply measures the number of inward citations. In-degree centrality gives equal weight to all inward citations, regardless of the position of the citing node in the network. An alternative measure of importance is eigenvector centrality. Unlike degree centrality, eigenvector centrality takes into account the position occupied by citing cases in the network. The metric assigns greater weight to inward citations cases that are themselves cited more frequently [8].

Finally, we use the fast-greedy community detection algorithm to identify clusters of densely connected cases and agreements [24]. The underlying intuition behind this community-detection algorithm is that cases form a community if they refer more to cases (and agreements) inside the community than to cases (and agreements) outside the community [9]. In mathematical terms, Newman [24] defines the problem of community detection in networks as one of optimizing the value Q in the following function:

$$Q = \sum_i (e_{ii} - \sum_j e_{ij}^2) \quad (2)$$

where e_{ij} is the proportion of edges between nodes in community i and community j. Values of $Q \neq 0$ have the interpretation of indicating a network division (into communities) where some degree of community structure is present. Calculating Q for all possible network divisions is computationally expensive, however, even with just a few dozen nodes. The fast-greedy algorithm uses hierarchical clustering to solve the problem approximately. Using this technique we obtain a classification of

[3]Our analysis ignores references to other documents and agreements which occasionally occur in EU international agreements.

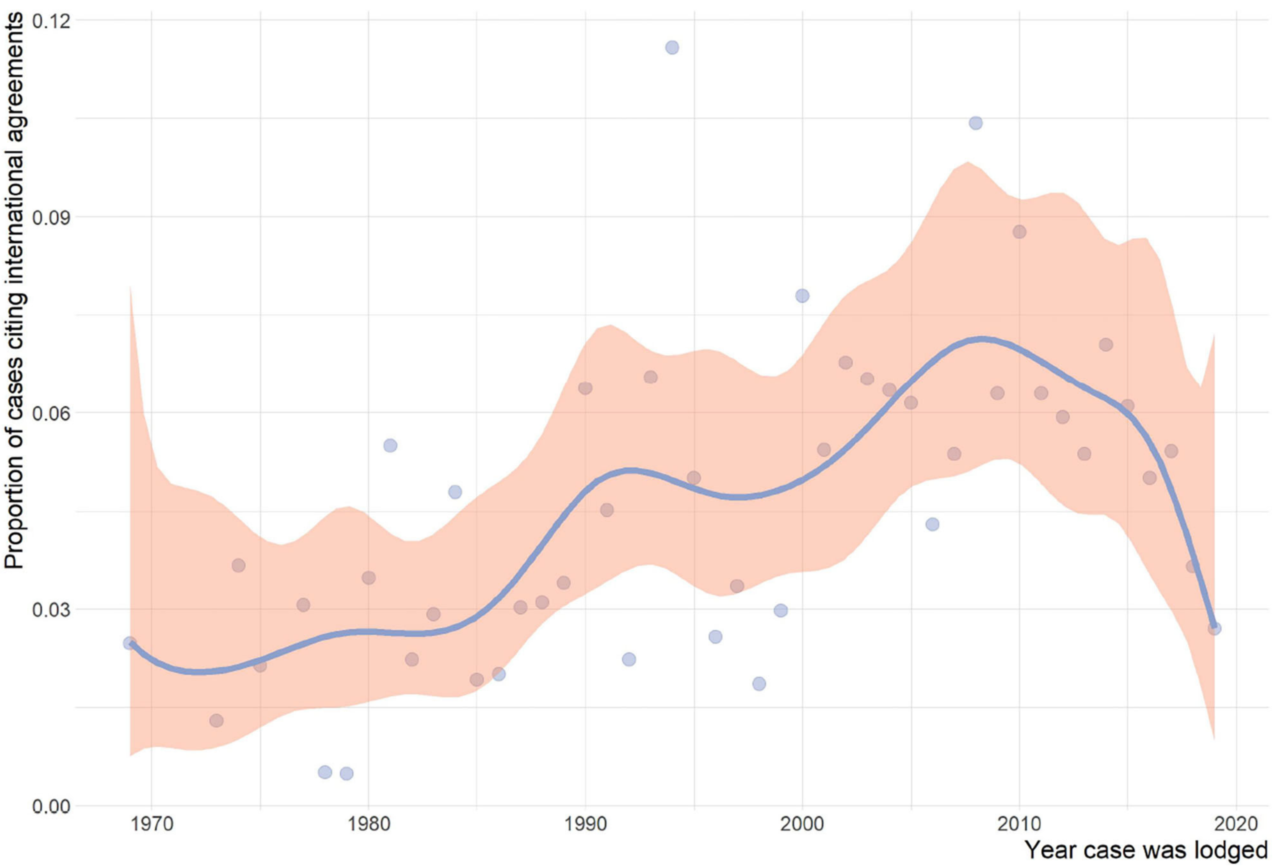


FIGURE 2 | Proportion of EU court cases that contain a reference to at least one international act (1969–2019). We estimate a smoothed trend line with 95 per cent confidence intervals using quasi-binomial regression.

all nodes in our network as belonging to one of S communities where S is determined computationally by finding the maximal value of Q.

5. RESULTS

5.1. Topics in EU International Agreements

To set k, which determines the number of topics, we relied primarily on interpretability and our domain knowledge. We found that $K = 9$ resulted in the most interpretable model. Because agreements sometimes contain technical nomenclatures with numerous acronyms resulting in less interpretable models, we excluded terms with fewer than three characters.

Plotted in **Figure 3** are the nine topics (summarized by their most characteristic words) and their proportion in the corpus over time. EU international agreements largely pertain to trade in goods ("products," "materials," "textile," "wine") and arrangements relating to their shipment and labeling ("originating," "weight," "pdo"). Three product categories can be discerned from the topmost characteristic words: textiles, fish, and other agricultural products (such as wine), including the protection of geographical indications ("pdo" stands for "Protected designation of origin").

Except for wine and protection of geographical indications, topics relating to products and product shipment (topic "originating, product, materials, value" and topic "exceeding, materials, weight, heading") have seen their relative importance decline over time. Textile ("textile, quantitative,limits, export") experienced a surge in attention in the 1980s and 1990s but later reverted to relative obscurity. Topics relating to services and air transport ("services, cpc, public, law" and "air, authority, services, persons") have steadily grown in importance.

Fishing—an issue that featured prominently in Brexit negotiations—saw a blip in the 1970s. After that, topic proportion remained more or less constant at around ten per cent.

Of great historical importance is the EU's relationship with its member states' former African, Caribbean and Pacific (referred to as "ACP") colonies (topic "programmes, acp, research, projects"). Development cooperation between the EU and ACP countries has given rise to a succession of agreements, starting with the Yaoundé Agreements (1969), followed by the Lomé Conventions (1974) and the Cotonou Agreement (2000).

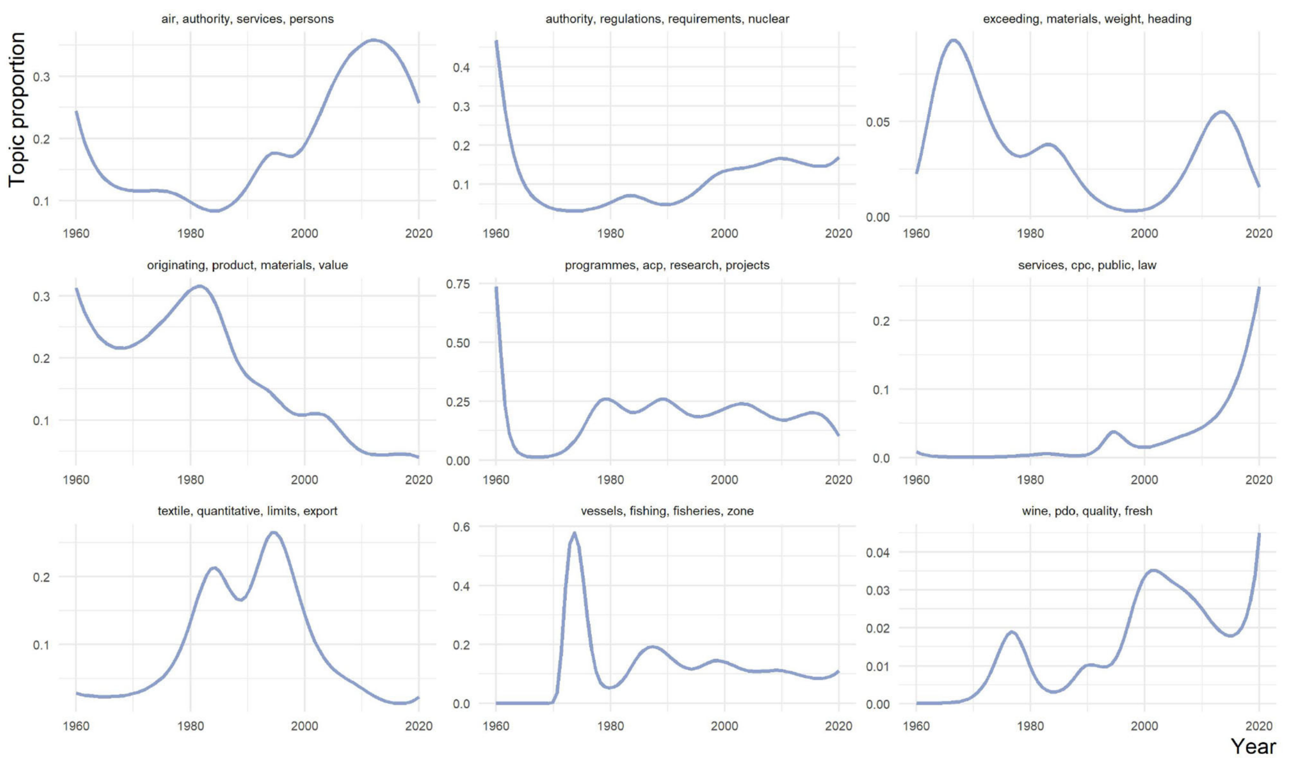


FIGURE 3 | Dynamic topic model of EU international agreements. Topics are summarized by their four most-characteristic words.

The temporal shifts in topic proportion visible in **Figure 3** reflect, for a part, the evolution of market integration and the growing emphasis on services in later stages of the construction of the internal market. But, while new treaties from the Single European Act (1987) to Lisbon (2009) have granted the EU competences in new policy areas, such environmental protection, immigration and security, our topic model indicate that EU treaty-making continues to concentrate on trade.

5.2. Citation Patterns in Litigation

After examining issue attention in the agreements, we now consider references to these agreements in EU court cases.

References to EU international agreements in EU court cases exhibit marked disparities. Just seven agreements—the Agreement on the European Economic Area (EEA), the WTO Agreements, the Ankara Agreement, the 1970 Additional Protocol to the Ankara Agreement, the International Convention on the Harmonized Commodity Description and Coding System (HS Convention), the Aarhus Convention and the EU-Switzerland Agreement on free movement of persons—are cited in more than 20 rulings (see the **Supplementary Material**). Most cited is the EEA Agreement (mentioned in 288 rulings); followed by WTO Agreements (mentioned in 139 rulings), the Ankara Agreement (mentioned in 66 rulings) and its 1970 Additional Protocol (41 rulings).

Figure 4 conveys the same point more systematically. The network is restricted to edges representing direct citations to the agreement. The communities, as identified by the fast-greedy algorithm, are largely isolated from each other, which signifies that two distinct treaties rarely have direct legal bearing on the same case. We employ the in-degree node centrality metric to show the importance of a node in the citation network.

Interestingly, even thematically proximate agreements, such as the WTO agreements and the HS Convention, both of which address international trade in goods, form case clusters that are almost completely separate. The main exception in the network is the EEC-Turkey Ankara Agreement, which is accompanied by the 1970 Additional Protocol. However, given the explicit legal connection between these two international agreements, we should expect that both will often be relevant to the same legal dispute. Similarly, a sparser chain of rulings connects the successive Lomé Conventions between the EEC and the ACP countries, though curiously these are not connected to the Cotonou Agreement which succeeded the Lomé framework in 2000.

Our second network goes beyond direct citations and considers the centrality of both citing and cited rulings. Properties of these networks—including number of nodes, edges, diameter, average degree, modularity, connectance, and transitivity—are reported in **Table 1**. **Figure 5** displays the most prominent agreements and rulings along with the main clusters identified via our Newman's [24] fast-greedy algorithm.

To facilitate data visualization, we use eigen centrality to reduce the size of the network. As explained section 4, eigen

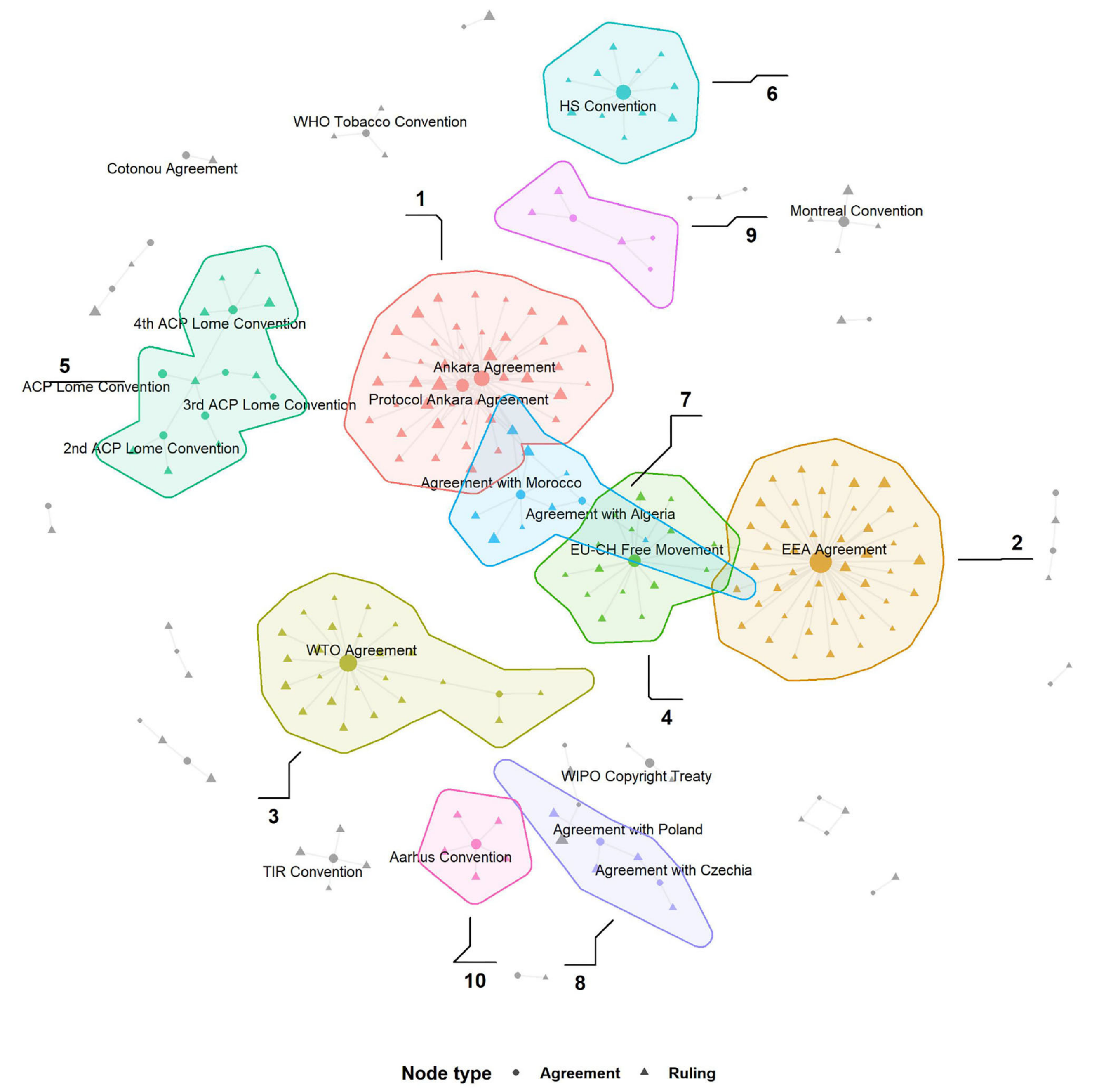


FIGURE 4 | Most prominent EU international agreements in the case law of the European Court of Justice. Network is restricted to direct citations to international agreements. Node size for international agreements reflects in-degree centrality. The more edges point toward a node, the larger its size. Colored areas, labeled from 1 to 10, represent network communities identified via Newman's [24] fast-greedy algorithm.

centrality captures the importance of cases from which citations originate[4]. The network plotted in **Figure 5** reflect this definition of case importance.

Compared to **Figure 4**, there is overall greater overlap between communities, although communities two (EEA Agreement), three (Ankara Agreement and Protocol to Ankara Agreement), five, seven, and eight clearly stand somewhat apart from the overlapping core formed by, in particular communities one (WTO Agreement) and four. The cases in the community clustered around the Ankara Agreement (community 3), such as

[4]Let x_i be the eigen centrality of node i in network Q. Then $x_i = \frac{1}{\lambda} \sum_{j \in M(i)} x_j$ where $M(i)$ is a set of all neighbors of i and λ is a constant.

TABLE 1 | Summary metrics of citation network.

Property	Value
Number of nodes	3,863
Number of edges	8,198
Average node degree	4.24
Average node in-degree	2.12
Average node out-degree	2.12
Network diameter	11
Connectance	0.0005495041
Modularity coefficient	0.725
Transitivity	0.071

*Metrics correspond to the network displayed in **Figure 5**. Diameter measures the shortest path between the two most distant nodes. Connectance measures the ratio of realized to possible links (computed as $\frac{L}{N*(N-1)}$). The modulatority coefficient is calculated from the vector of community membership generated via Newmans's [24] fast-greedy algorithm for community detection. Transivity measures the probability that adjacent nodes are connected.*

C-561/14 relate predominantly to the rights of Turkish citizens to live and work in the EU. The WTO community (community 1) encompass some landmark cases concerning the interpretation and application of international law in the EU legal order, such as *ATAA*, *Rosneft*, and *Krizan*, the latter dealing specifically with access to information and justice in environmental matters as regulated by the Aarhus Convention. For the first two of these cases, the network seems to reflect well the general nature of the international law questions they raised, as they stand in the center of the entire network with connections spanning across communities.

5.3. Litigation Topic and Incidence of References to EU International Agreements

To explore how cases citing international agreements may differ from cases that do not, we estimated a topic model of EU court cases with covariates for time and reference to international acts as explained in section 4. It is important to bear in mind that EU cases far outnumber EU international agreements. So, with a substantially larger set of documents, we found that, for this task, $K = 30$ produced the most interpretable model.

Figure 6 illustrates the comparative evolution in topic proportion for the two case categories. Some topics indicate a clear divergence between cases with and cases without reference to international acts. For example, cases concerning residence and family reunion rights (topic "residence, country, family, nationals") have an obvious international dimension, which has emerged early in litigation. The number of cases not referencing international agreements in this area has been catching up, however, possibly spurred by the creation of important EU rules such as Directive 2004/38/EC on the right of citizens of the Union and their family members to move and reside freely within the territory of the Member States and the ECJ's interpretation of the notion of EU citizenship [25].

Surprisingly, international agreements seem to be disproportionately invoked in competition law cases ("fine, undertakings, cartel, fines"). This may reflect the globalization of antitrust regulation promoted by the European Commission, which has resulted in the insertion of competition provisions in several agreements [26].

International agreements seem to have become increasingly less relevant in cases pertaining to public procurement ("contracts, tender, award, consumer") and indirect taxation ("vat, tax, sixth, taxable"). These legal areas, along with trademarks ("mark, trademark, euipo, board") and road safety ("insurance, vehicles, vehicle, freedom") have seen increasing regulatory harmonization at EU level; a development that seems to been accompanied by intensifying litigation [7]. To the extent that these topics are highly prevalent in recent years, they may explain the pattern seen in **Figure 2**, which shows a declining proportion of EU court citing EU international agreements.

That staff cases ("staff, officials, competition, post"), most of which employment disputes between EU civil servants and EU institutions, almost never cite international law appears banal, although it provides face validity for our methodological approach.

5.4. Case Clusters and Topic Proportion

Finally, we combine network analysis and topic modeling to assess variations in thematic focus in cases belonging to distinct communities. For this purpose, we averaged topic proportion across cases belonging to the same community. Depicted in **Figure 7** is a radial plot comparing topical distribution in the EEA and Ankara Agreement clusters, corresponding to, respectively, community two and three in the network illustrated in **Figure 5**. Average topic proportion for other communities is reported in **Supplementary Material**.

Whereas cases in community two, which are centered around the EEA treaty, concern primarily the free circulation of products and capital, cases in community three, which are clustered around agreements with Turkey, deal mostly with rights of Turkish citizens to live and work in the EU. Combining the two results provides further validation of both approaches. Given how relatively little overlap there is between the two communities, we would expect the topical content of the cases to be quite different, which is precisely what we observe in the topic model.

6. DISCUSSION

What the law is in a given domain or on a given question is typically the expression of information scattered across a large web of texts connected in complex ways [13, 27]. The larger the web, the more difficult it becomes to comprehend its general structure and dynamics using the tools lawyers and legal academics have traditionally applied to study and research the law—manual parsing of documents. While they still require domain knowledge, network analysis and NLP methods provide a scalable alternative to explore the complexity of law.

We applied these two techniques to examine three aspects of the large body of international agreements concluded by EU institutions: (1) their dominant theme, (2) their comparative

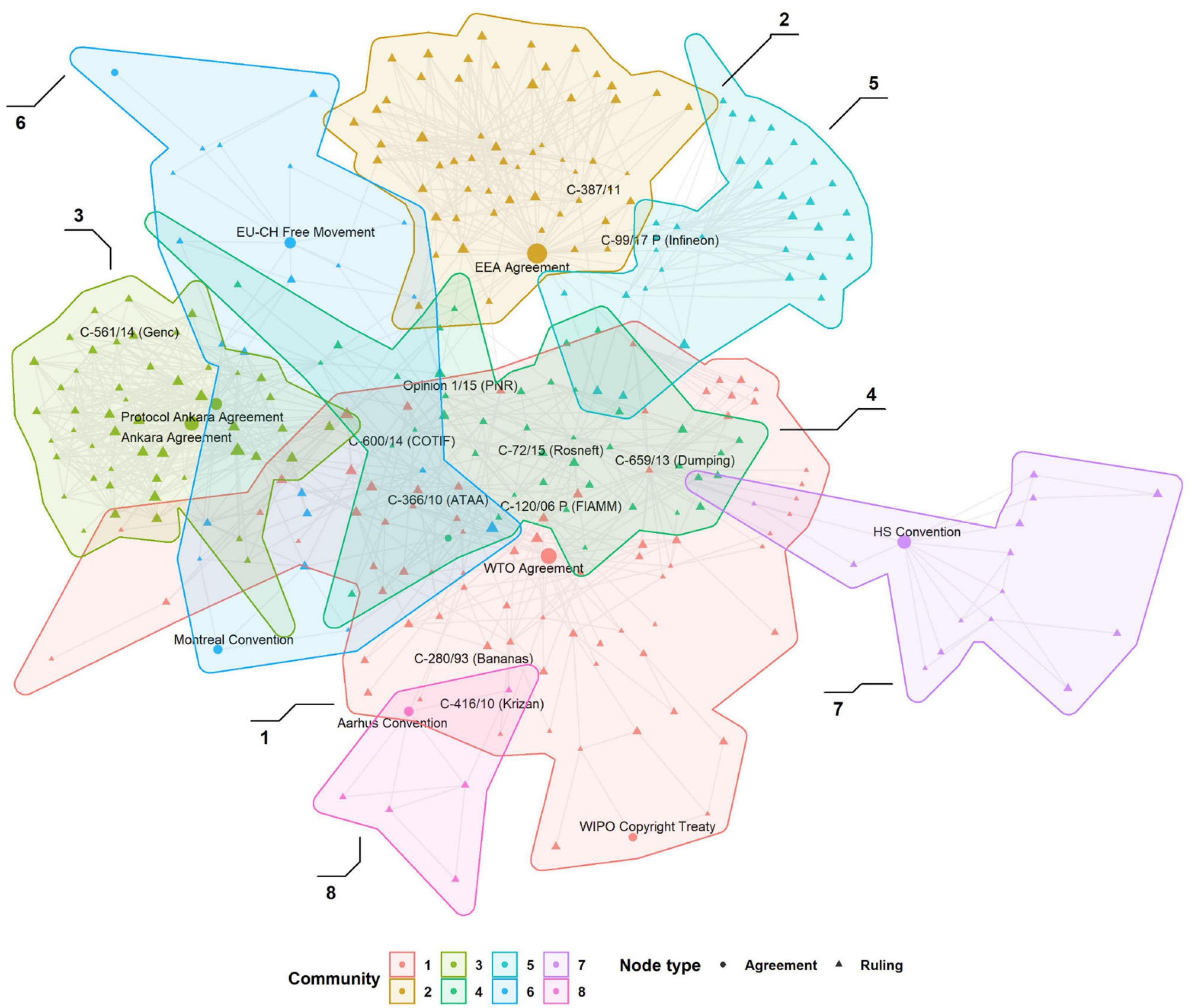


FIGURE 5 | Most influential agreements and rulings. The network shows only nodes with above average Eigen centrality. Node size reflects in-degree centrality. Colored areas, labeled from 1 to 8, represent network communities as identified by Newman's [24] fast-greedy algorithm.

centrality in EU court cases, and (3) in the area of litigation in which they are more likely to be involved. We found that economic issues continue to dominate the EU's treaty-making activity; that the EEA, Ankara Agreement, and the WTO Agreement form the largest litigation clusters; and that references to international agreements is proportionally higher in disputes pertaining to antitrust and residence and family reunion rights.

The particular salience of international agreements in residence and family reunion rights speaks directly to the British government's insistence on excluding both mobility rights and the ECJ's jurisdiction from a post-Brexit free trade agreement [1]. Private litigant's standing combined with justiciable mobility rights seem to operate as a powerful litigation catalyst, inviting judges to step in.

Our analysis is primarily conceived as exploratory, but we are confident that it achieves its goal of providing an overview of the EU's treaty-making activity and litigation. Still, we point out two limitations which similar studies may seek to address in the future. The first is that our analysis of litigation is restricted to treaties and agreements to which the EU is formally party. However, international agreements to which the EU is not party—such as the Vienna Convention on the Law of Treaties and the United Nations Charter—have been invoked in ECJ decisions. Future work may seek to map these dynamics. Second, our text-mining procedure follows a bag-of-word approach, which disregards synonymy as well as polysemy and co-reference resolution. Future research may seek to apply distributed semantic and transformer models, which implement

FIGURE 6 | Topic proportion in EU court cases (1980–2019) with and without reference to EU international agreements. Topics are summarized by their four most-characteristic words.

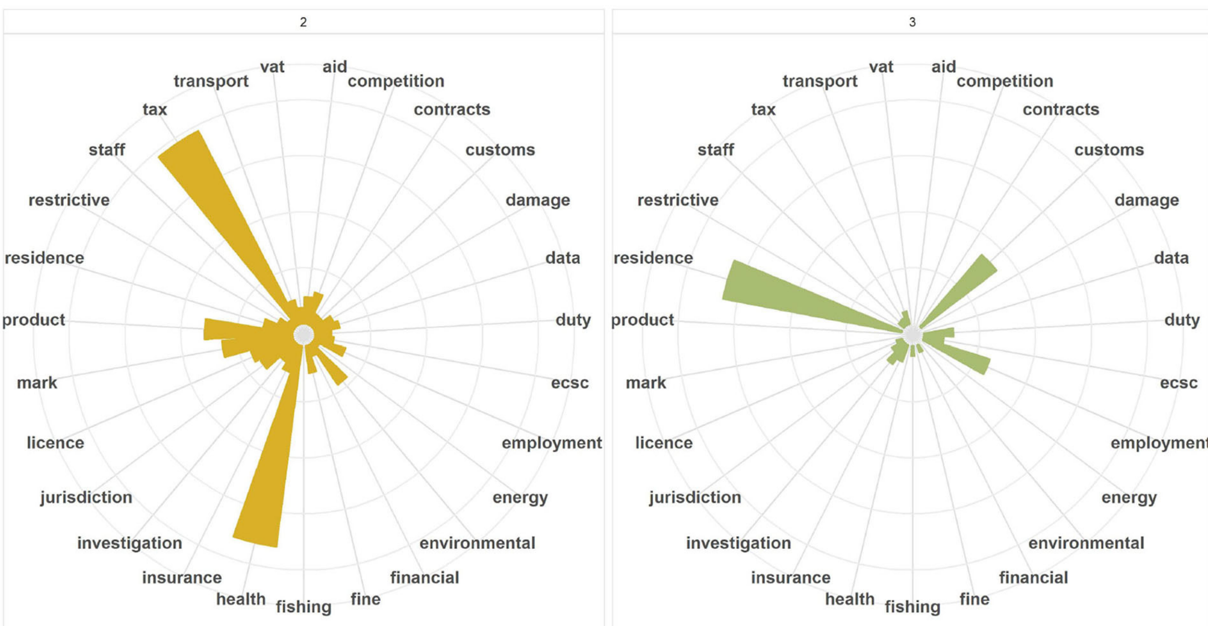


FIGURE 7 | Average topic proportion in network community two (EEA Agreement) and three (Ankara Agreement) identified via fast-greedy algorithm [24]. Values closer to the origin of the circle indicate lower topic proportion. Topics are labeled using the topmost characteristic word.

word embeddings capturing more of the context in which words and even sentences occur [28, 29].

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at: www.eur-lex.eu.

AUTHOR CONTRIBUTIONS

MO conducted the data analysis with input and supervision from AD. AD and MO wrote most of the article. KW contributed suggestions and syntactical improvements on the last draft. All authors contributed to the article and approved the submitted version.

FUNDING

The authors gratefully acknowledge financial support from European Research Council Grant No. 638154 (EUTHORITY).

ACKNOWLEDGMENTS

We are grateful to Federico Benetti for research assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2021.657607/full#supplementary-material

REFERENCES

1. Leinarte E, Barnard C. From constitutional adjudication to trade arbitration enforcing mobility rights post-brexit. *Eur Foreign Affairs Rev*. (2020) 25: 589–612. doi: 10.17863/CAM.58501
2. Young AR. European trade policy in interesting times. *J Eur Integr*. (2017) 39:909–23. doi: 10.1080/07036337.2017.1371705
3. Lenaerts K, De Smijter E. The European community's treaty-making competence. *Yearb Eur Law*. (1996) 16:1. doi: 10.1093/yel/16.1.1
4. Teodosiu MJ. The treaty-making powers of the european economic community. *J Int'l L Prac*. (1993) 2:501.
5. Frankenreiter J, Livermore MA. Computational methods in legal analysis. *Annu Rev Law Soc Sci*. (2020) 16:39–57. doi: 10.1146/annurev-lawsocsci-052720-121843
6. Ash E, Chen DL. Case vectors: spatial representations of the law using document embeddings. In: *Law as Data*. Livermore M, Rockmore D, editors. Santa Fe: Santa Fe Institute Press (2019). p. 313–37. doi: 10.2139/ssrn.204926
7. Dyevre A, Lampach N. Issue attention on international courts: evidence from the European court of justice. *Rev Int Organ*. (2020) 15:1–23. doi: 10.1007/s11558-020-09391-0
8. Fowler JH, Johnson TR, Spriggs JF, Jeon S, Wahlbeck PJ. Network analysis and the law: measuring the legal importance of precedents at the US Supreme Court. *Polit Anal*. (2007) 15:324–46. doi: 10.1093/pan/mpm011
9. Lupu Y, Voeten E. Precedent in international courts: a network analysis of case citations by the european court of human rights. *Br J Polit Sci*. (2012) 42:413–39. doi: 10.1017/S0007123411000433
10. Larsson O, Naurin D, Derlén M, Lindholm J. Speaking law to power: the strategic use of precedent of the court of justice of the European Union. *Compar Polit Stud*. (2017) 50:879–907. doi: 10.1177/0010414016639709
11. Sadl U, Olsen HP. Can quantitative methods complement doctrinal legal studies: using citation network and corpus linguistic analysis to understand international courts. *LJIL*. (2017) 30:327. doi: 10.1017/S0922156517000085
12. Ruhl J, Katz DM. Measuring, monitoring, and managing legal complexity. *Iowa L Rev*. (2015) 101:191.
13. Katz DM, Coupette C, Beckedorf J, Hartung D. Complex societies and the growth of the law. *Sci Rep*. (2020) 10:1–14. doi: 10.1038/s41598-020-73623-x
14. Katz DM, Stafford DK. Hustle and flow: a social network analysis of the American federal judiciary. *Ohio St LJ*. (2010) 71:457.
15. Ovádek M. Facilitating access to data on European Union laws. *Polit Res. Exchange*. (2021) 4:e1870150. doi: 10.1080/2474736X.2020.1870150
16. Ovádek M, Raina A. The evolution of EU trade law through the prism of competence: a quantitative, longitudinal perspective. *J World Trade*. (2019) 53:489–508.
17. Blei DM, Ng AY, Jordan MI. Latent dirichlet allocation. *J Mach Learn Res*. (2003) 3:993–1022. doi: 10.1162/jmlr.2003.3.4-5.993
18. Blei DM. Probabilistic topic models. *Commun ACM*. (2012) 55:77–84. doi: 10.1145/2133806.2133826
19. Roberts ME, Stewart BM, Tingley D, Airoldi EM. The structural topic model and applied social science. In: *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*. Vol. 4. Lake Tahoe: Harrahs and Harveys (2013).
20. Blei DM, Lafferty JD. A correlated topic model of science. *Ann Appl Stat*. (2007) 1:17–35. doi: 10.1214/07-AOAS136
21. Blei DM, Lafferty JD. Dynamic topic models. In: *Proceedings of the 23rd International Conference on Machine Learning*. Pittsburgh, PA (2006). p. 113–20. doi: 10.1145/1143844.1143859
22. Borgatti SP, Mehra A, Brass DJ, Labianca G. Network analysis in the social sciences. *Science*. (2009) 323:892–5. doi: 10.1126/science.1165821
23. Freeman LC, Roeder D, Mulholland RR. Centrality in social networks: II. Experimental results. *Soc Netw*. (1979) 2:119–41. doi: 10.1016/0378-8733(79)90002-9
24. Newman ME. Fast algorithm for detecting community structure in networks. *Phys Rev E*. (2004) 69:066133. doi: 10.1103/PhysRevE.69.066133
25. Šadl U, Madsen MR. Did the financial crisis change European citizenship law? An analysis of citizenship rights adjudication before and after the financial crisis. *Eur Law J*. (2016) 22:40–60. doi: 10.1111/eulj.12167
26. Aydin U. Promoting competition: European union and the global competition order. *J Eur Integr*. (2012) 34:663–81. doi: 10.1080/07036337.2012.707366
27. Ruhl J, Katz DM, Bommarito MJ. Harnessing legal complexity. *Science*. (2017) 355:1377–8. doi: 10.1126/science.aag3013
28. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. *arXiv [Preprint]. arXiv:13013781*. (2013).
29. Devlin J, Chang MW, Lee K, Toutanova K. Bert: pre-training of deep bidirectional transformers for language understanding. *arXiv [Preprint]. arXiv:181004805*. (2018).

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 17 May 2021
doi: 10.3389/fphy.2021.664621

Check for updates

Conceptual Graphs and Terminological Idiosyncrasy in UNCLOS and CBD

Pierre Mazzega [1,2]*

[1] *Solidarity, Societies, Territories, Interdisciplinary Laboratory LISST UMR 5193, CNRS–University of Toulouse Jean-Jaurès, Toulouse, France,* [2] *Strathclyde Centre for Environmental Law and Governance (SCELG), University of Strathclyde, Glasgow, United Kingdom*

OPEN ACCESS

Edited by:
Daniel Martin Katz, Illinois Institute of Technology, United States

Reviewed by:
Michele Bellingeri, University of Parma, Italy
Michael Bommarito, Michigan State University College of Law, United States

***Correspondence:**
Pierre Mazzega
pmazzega@gmail.com

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *05 February 2021*
Accepted: *09 April 2021*
Published: *17 May 2021*

Citation:
Mazzega P (2021) Conceptual Graphs and Terminological Idiosyncrasy in UNCLOS and CBD. Front. Phys. 9:664621. doi: 10.3389/fphy.2021.664621

Do two conventions of international environmental law necessarily endow the same word with the same meaning? A single counterexample is enough to answer in the negative: this is the case of the term "*resource*" in the United Nations Convention on the Law of the Sea (UNCLOS) and the Convention on Biological Diversity (CBD). Beyond this result, we tackle the questions, raised by the method of analysis implemented, about the semantics of legal texts, a source of interpretative flexibility but also of cognitive amalgamations and confusions of various types. A conceptual graph is associated with each proposition or sentence comprising the term "*resource.*" Some expressions, especially those of a deontic nature and noun phrases naming a group of interrelated entities or a fact, are encoded in nested graphs. The scope of a term is revealed by the neighbourhood of its uses. Neighbouring expressions, positioned along the paths of conceptual graphs, are ranked owing to their distance from the target expression. Then the neighbours the most contributing to the distributional meaning of the targets are classified in a coarse taxonomy, providing basic ontological traits to "*resource*" and related expressions in each convention. Although the two conventions rely on the same language, the weak overlap of their respective neighbourhoods of the term "*resource*" and associated expressions and their contrasted ontological anchorages highlight idiosyncratic meanings and, consequently, divergent orientations and understandings regarding the protection and conservation of resources, especially of living resources. Thus, the complexity of legal texts operates both in the gap between language semantics and cognitive understanding of the concepts used, and in the interpretative flexibility and opportunities for confusion that the texts offer but that the elementary operations of formalisation allow to deconstruct and clarify.

Keywords: environmental law, conceptual graph, lexicon, distributional semantics, lattice, idiosyncrasy

INTRODUCTION

Fifteen years ago, in the continuation of the debate of ideas on the complexity of law and inflation of normative production [1–6] the question arose of defining legal complexity in a way that could be related to the exploration of large corpus of texts available on various online platforms[1]. The rise in power of network analysis and graph theory [7–11] quite naturally oriented research towards a relatively simple, intuitive and tractable approach which consists in identifying and analysing the networks induced by various types of citations or referencing between textual segments–article, laws, etc. [12, 13]. Physics (and ecology) shows that the structural properties of a complex system depend on the scale at which the system is observed, and thus conditions the choice of the relevant paradigm of analysis. Testing this idea, we highlighted the differentiated statistical properties of texts at the intra-article, article [14] and code level [15], and in a network of tens of codes of French law [16, 17].

Secondly, it became useful to analyse what simple lexicometric indicators can reveal about the emergence of a theme in international environmental law conventions [18]. The current international health situation gives particular relief to our analysis of the emergence of the health-environment theme, from the 90s, and more recently (roughly from 2010) of "One Health"–which links human health, animal health and environmental health, in the Rio conventions [19, 20].

However, these analyses remain far from the centre of the legal forge–at least in its literary expression: the meaning of the texts. Certainly the difficulty of semantic analysis, its disciplinary specificities and the diversity of existing approaches (in particular the logicist vs. distributional currents [21]), or even its links with syntactic analysis and discourse analysis, are *a priori* discouraging initiatives that would go in such direction[2]. Yet it is an entire continent that the complex systems approach now sets out to explore. Indeed, the considerable progress made over the past 30 years by Natural Language Processing [23–25] allows rapid and reliable access to the identification and characterisation of various grammatical units that make up lexemes, phrases, sentences, and even texts. Analyses of legal texts based on linguistic concepts have already been proposed ([26–29]; see also the special issue introduced by Robaldo et al. [30]) which open up perspectives that have not yet been explored. The multitude of relationships that it is possible to build between components of texts occurring at various levels of grammatical organisation of utterances deploys network structures. Graph theory, with its algebraic ramifications, offers a battery of more or less standard concepts and tools for analysing these networks.

At one of the most elementary levels of compositional semantics, the analysis proposed here sets out to answer a simple question: Do two conventions of international environmental law necessarily endow the same word with the same meaning? Let us note provisionally that if the answer is positive, then the textual glosses and the dialogues established in various collective arena can be conducted without change. No lexical ambiguity fosters risk of misunderstanding. A negative answer would mean that various legal streams make differentiated uses of the same lexicon. The understanding and use of legal texts then depend on the context of interpretation, or of the intention behind them.

Simple question in the sense that it only touches the lexical layer of texts, contributing to their meaning. This is a non-trivial question, however, insofar as the conventions use the same language, here English, and even belong to the same genre of discourse. Moreover, in addition to the production of appropriate evidence to support an answer, two additional objectives are pursued here: (a) to design an approach capable of highlighting idiosyncratic uses of terms from a restricted textual corpus; (b) identify a first essential property of legal texts that an analysis in context of the linguistic material can reveal, and outline the consequences on the normative level.

After setting the legal context of the study in Section Introduction, Section 3 UNCLOS, CBD and the Resource Issue presents the approach developed. It is inspired by the central hypothesis of distributional semantics (initially proposed by Firth [31]) according to which the meaning of a term or expression emerges from its use in context. By capturing very large textual corpora (composed of billions of words), this approach can legitimately claim to be statistically driven. With two texts in our pocket and at most a few dozen occurrences of the same term, we will not be able to avail ourselves of this advantage. The use of an expression in context amounts to identifying its immediate neighbourhoods in the corpus. We also propose to use the construction of conceptual graphs rather than raw sentences or syntax trees to identify these neighbourhoods. The use of these neighbourhoods constituting the meaning of a term in context is exposed in Section Encoding the Conventions Textual Data.

The neighbourhoods of the target term "*resources*" (and of associated expressions) as taken from the UNCLOS and CBD conventions, are described in Section Lexical Neighbourhoods and formalised as a lattice relying on a coarse taxonomy. Their comparison leads to favour the hypothesis of an idiosyncratic use of the term "*resource*" and its associated expressions in these two conventions. Section Meanings of "Resource" in UNCLOS and CBD discusses the limitations, advantages and avenues opened up by the analytical method. Then follow the implications which seem to emerge at the normative level from these results. Section Discussion concludes this exploratory study.

UNCLOS, CBD, AND THE RESOURCE ISSUE

The question addressed by this study corresponds to a testable hypothesis. It would suffice to find a single term or expression

[1]Like the European Eur Lex platform https://eur-lex.europa.eu/homepage.html, the French Legifrance https://www.legifrance.gouv.fr/, US codes at the Legal Information Institute of Cornell Law School https://www.law.cornell.edu/uscode/text, the IUCN gateway to environmental law ECOLEX https://www.iucn.org/theme/environmental-law/resources/ecolex, the UN Treaty collection https://treaties.un.org/, WTO legal texts https://www.wto.org/english/docs_e/legal_e/legal_e.htm, just to cite a few of them.

[2]"*The little existing research on legal language suggests that, more than by a specialized vocabulary, it is characterized by overly complex sentences, the overuse of passives, whiz-deletion and unclear pronoun reference, archaic and misplaced prepositional phrases, and its own set of articles and demonstrative pronouns. The historical development of legal language is unique, paralleling but independent of the development of the rest of English. Legal language is both the medium of communication and the primary tool of the legal profession, and is powerful because it carries the force of law.*," cited from Crandall and Charrow [22].

that does not have the same meaning in two conventions to answer in the negative. Our candidate is the term "*resource*" and associated frozen expressions such as "*natural resource,*" "*living resource,*" etc. This term has a relatively high number of occurrences in several international environmental law conventions that regulate the management of resources (here "management" includes in particular access to, sharing of, proper management, conservation and protection of resources), but according to the perspectives and objectives specific to each of them. Resource management is a major subject of past, present and likely future tensions in our societies and their differentiated developments, for which the key players on the international scene and for international environmental law are States.

Regarding the conventions, we consider the United Nations Convention on the Law of the Sea (hereafter UNCLOS) and the Convention on Biological Diversity (CBD). There are several reasons for this choice. The UNCLOS defines the levels of territorial jurisdiction over the seas and oceans, regulates the passage of vessels, establishes the rules for access to marine resources as well as the conditions for the conduct of activities using these resources, in particular with regard to their impact on the marine environment and its non-living and living resources. The convention stipulates the duties of states in the conservation and management of the resources of the high seas, and establishes the architecture of international ocean governance that has prevailed in recent decades. The CBD regulates the rights and duties of States regarding the use, management, preservation, and conservation of biological resources, including genetic resources. It establishes the rules for cooperation between nations and for sharing the benefits derived from all forms of exploitation of these resources. The marine environment is regulated by CBD, just like other natural environments.

However, marine biodiversity and the environment of the high seas are exposed to increasing threats–physical and chemical modifications of water linked to climate change, various pollution, overexploitation of resources, loss of habitats, etc. directly or indirectly linked to human activities at sea but also on land. The growing needs for resources on a planetary scale, in particular mineral resources (like metals, rare earths) and in connexion with the energy transition [32], on the contrary, are pushing us to increasingly turn to the oceans [33–36] perceived as a kind of immense marine continent, relatively little explored, promise of largely unexploited stocks of mineral and living resources [37, 38]. At the same time, a growing harvest of scientific results warns about the risks that projects of large-scale deep sea mineral resource exploitation pose to living resources and marine biodiversity, and to ecosystems specific to the deep seabed, ocean ridges and seamounts [39]. Therefore, though the UNCLOS and CBD obey the same principles of public international law, the areas of marine activity that the two conventions cover differ, and their goals diverge.

After 10 years of discussions, and the mixed observation regarding the success of the conventions in achieving some objectives for the completion of which they had been designed and implemented, the international community initiated negotiations in 2018 aimed at establish a new binding treaty of the sea[3] Under the aegis of the United Nations and under the UNCLOS, this treaty aims at "the conservation and sustainable use of marine biodiversity in areas beyond national jurisdiction." Resource assessments, environmental baseline studies and assessment studies of the impact of deep seabed exploration and resources exploitation (still underdeveloped) on marine biology, are still limited [40], so that the negotiations around the new treaty must come to terms within a context of high scientific uncertainty [41], or even ignorance as to the vulnerability of the ecosystems that shelter the marine life.

We are therefore at a pivotal moment in the future of the oceans, at least as envisaged by international environmental law, which combines, and potentially, contradicts (at least) two aims of management of marine resources which will have major impacts both on the development of nations and on the marine environment, ecosystems, and life that resides there. In this context it is interesting to return to the source of these regulations and to analyze what these inaugural texts say about resources and how they approach them. Their original versions have shaped the contemporary form of regulation of their respective domains and the initial direction of their developments through Conferences of Parties or works of *ad hoc* scientific groups to this day.

ENCODING THE CONVENTIONS TEXTUAL DATA

UNCLOS and CBD Conventions

The UNCLOS (Montego Bay, 10 December 1982) entered into force on 16 November 1994, and the CBD (Rio de Janeiro, 5 June 1992) on 29 December 1993. Both conventions count today 168 Parties[4] The UNCLOS text includes a preamble, 320 articles divided into 17 Parts, plus 9 annexes. The CBD is shorter, with a preamble, 42 articles and 2 annexes. The annexes of the two conventions are excluded from our analysis.

Both conventions present sets of definitions gathered under "*use of terms*" titles. Only two of these sets introduce definitions including the term "*resource,*" say the UNCLOS Art. 1 of Part I *Introduction*, and Art. 133 of Part XI *The Area*; in the CBD, Art. 2 gives definitions related to the "*resource.*" Lexicology classically distinguishes definitions of the kind "*x is a y*" from those of the kind "*x means y*": the first one is targeting the entity of the world designed by *x*, while the second one provides information on the term *x* and about the lexical environment in which the term is inserted, in relation to the elements from which it is distinguished. "*Use of terms*" sections of legal conventions provide definitions of the second kind. These definitions are worth being included in our analysis, but they are by no way able to render the richness of the meaning of the defined terms, to capture their relations with other concepts or notions, and do not allow a fine distinction of their denotative and connotative uses in context. In addition, they are based on a sort of latent ontology, neither explicit nor explicated in the conventions (this is not their role) but which can be postulated as a minimum representation

[3]See Available online at: https://www.un.org/bbnj/content/background.

[4]See the UN Treaty Collection, chap. XXI for UNCLOS and chap. XVII for CBD at https://treaties.un.org/.

common to the drafters of legal texts, and linking all the concepts and notions used. Finally, these definitions only concern a very limited number of terms, so that a broader approach must be deployed to look for possible idiosyncratic uses of terms.

For this purpose, we use all 105 occurrences of the term "*resource*" (singular or plural) in UNCLOS articles, and the 49 occurrences in CBD articles. We do not formalise, by means of a conceptual graph, each complete article in which one or more occurrences of "*resource*" are inserted, but only the sentences concerned, or even the only propositions having an autonomous meaning (for example when these propositions form a list of options or cases). These restrictions are justified insofar as they reduce the formalisation workload without having any impact on the way in which we define the neighbourhood of a term or of an expression as will be seen.

Conceptual Graphs

Conceptual graphs [42, 43] were designed to represent knowledge (assertions, rules or constraints established on domains, queries and answers, etc.) and to translate the various useful manipulations of knowledge in terms of rigorous mathematical operations on graphs [44]. Here we only use–and introduce–the basic properties of conceptual graphs[5]. Two types of vertices are distinguished–those representing concepts or notions [45] and those representing *n-ary* relations–connected by edges. In CGs, an edge reifies the link between a conceptual-type node and a relation-type node. For the purposes of our analysis it is enough to define a basic conceptual graph G as a 4-uple (C, R, E, λ) where (C, R, E) is a finite, undirected and bipartite multigraph (that is the possibility to have edges with the same end nodes[6]), C being the set of concept nodes, R the set of relation nodes, and E the set of edges; λ is a labelling function of the nodes and edges of G. The vocabulary $V(G)$ of graph G is here defined as the set of labels of the only vertices representing entities or concepts (set C).

Each sentence or proposition where the word "*resource*" occurs is encoded in an elementary conceptual graph. G_{UNCLOS} (respectively, G_{CBD}) is the set of all disconnected elementary graphs formed from the concerned sentences or propositions of the UNCLOS (resp. CBD) convention, and V_{UNCLOS} (resp. V_{CBD}) its vocabulary. A sub-graph of an elementary graph is shown in **Figure 1** as illustration. An elementary graph can be as simple as in **Figure 1** or count tens of nodes, some of them connected along looping paths (the larger elementary graph we built encodes UNCLOS Art. 150 *Policies relating to activities in the Area*, and has 67 nodes and 79 edges).

On **Figure 1** concept nodes are represented in rectangles and relation nodes in ellipses. Conceptual nodes can only be connected with relation nodes, and relation nodes only with conceptual nodes. It is often useful to identify some words or expressions as being attributes (relation with label "*ATTR*"; the label of the attribute is in a hexagonal vertex) of concepts. In such case, the attribute and the concept it modifies are interpreted as a single concept ("*due notice*" in our example). Several expressions that are more or less "frozen" (linguistic stasis) have this form, whether they include the word "*resource*" (e. g., "*living resources,*" "*natural resources*") or not (e g., "*coastal state,*" "*country in development*").

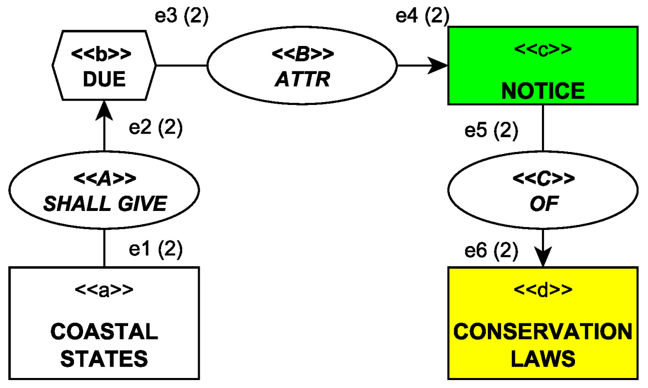


FIGURE 1 | Example of conceptual graph encoding the sentence "coastal States shall give due notice of conservation laws" (extracted from UNCLOS Art.62 §5) with concept node set $C = \{a, b, c, d\}$, relation node set $R = \{A, B, C\}$, edge set $E = \{e1, .., e6\}$, and labelling function $\lambda(a) =$ *'coastal states'* $\lambda(b) =$ *'due'*, ..., $\lambda(A) =$ *'shall give'*, etc. The edge labels indicate their identifier and arity in parenthesis. $V = \{$*coastal states*, *due notice*, *conservation laws*$\}$ (see text).

When encoding a set of sentences or propositions, it appears that many concepts are occurring several times while the expressions of relations in natural language are much more variable and diverse. Because they are linking two or more concepts, relations are not really contributing to the meaning of a concept. Therefore, the meaning of a target concept like "*resource*" will be captured only from the neighbouring concepts (a notion to be defined more precisely in Section Lexical Neighbourhoods).

Encoding Sentences in CGs

Conceptual graphs (CGs) represent knowledge in the form of a structure, articulating concepts and their relationships. In a textual corpus these articulations are expressed with the resources of natural language. CGs then make it possible to extract and represent the knowledge carried by the text–which is organised in a structure often qualified as "deep"–without depending on the singular linguistic form chosen to express this knowledge–form designated as "surface structure."

For the purposes of our study, here we represent only sentences separated from each other. In doing so, a few regularities are observed which guide this rewriting of sentences in a graph, without however making them rigid rules. In CGs, concepts and relationships are represented by a distinct type of node. Two nodes of the same type cannot be directly linked.

Each simple noun phrase is represented by a concept type node. Complex noun phrases, formed by several nested noun

[5] In particular, we will adopt a definition of a "vocabulary" that does not conform to that used for conceptual graphs, but simpler and better suited to our needs; where appropriate we will also use directed graphs (oriented edges).

[6] Since the CGs are multigraphs some application may require modelling them as weighted graph (see e g., Bellingeri et al. [46]). However, the sentence-by-sentence analysis performed here does not require resorting to it.

phrases, are generally representable by a linear succession of conceptual and relational nodes. From a syntactic point of view, the notion of "resource" which is our target appears in noun phrases. The nature or quality of the resource is specified by the most general addition of an adjective, as in "biological resource" or "mineral resource." The adjectival modification is represented via the relation "attribute" or in an equivalent way, by making the adjective the component of a frozen expression in a single node. Relationship nodes are most often occupied by verbs, possibly accompanied by an adverb treated as an attribute of the verb. We consider that the distance between an attribute and what it relates to–noun or verb–is zero.

However, sentences do not always have the simple syntactic structure as shown in **Figure 1**, due in particular to the use of anaphoric relations or of subordinate clauses in complex sentences. Expression in natural language makes frequent use of anaphoric relations and co-references, the resolution of which is essential for a good understanding of the text. This is for example the use of a pronoun which replaces a nominal antecedent, or the use of the referring word (pronoun, verb) which takes the place of a non-nominal antecedent, a source of more complex syntactic structures [47, 48]. Anaphora resolution is performed here by setting a relation node between the antecedent (the replaced term) and the other term syntactically associated to the pronoun. A frequent case is when a possessive pronoun backed by a concept (as in "its exclusive economic zone") refers to the entity which includes the entity designated by this concept (in our example the "coastal state"). Whatever the overall syntactic structure of the sentence in which these two components appear, a relation is established between them with the label "of" (the result reads "exclusive economic zone of coastal state"). This procedure naturally modifies the paths linking concept nodes and their distance. Authorised here by the small volume of our textual corpus, this simple approach produces reliable results without resorting to complex NLP procedures.

Most expressions of natural language can be attached either to the type of relations, or to a subtype of entities, such as "*actors,*" "*material resources,*" "*cognitive resources*" or "*norms.*" But let us consider the term "*pollution*" defined in Art. 1 of UNCLOS as follows:

"*'pollution of the marine environment' means the introduction by man, directly or indirectly, of substances or energy into the marine environment, including estuaries, which results or is likely to result in such deleterious effects as harm to living resources and marine life, hazards to human health, hindrance to marine activities, including fishing and other legitimate uses of the sea, impairment of quality for use of sea water and reduction of amenities.*"

"*Pollution*" is explicitly defined as an action ("*introduction of*") linking some actor ("*man*") with material entities ("*substances,*" "*energy*", "*marine environment*"), through several relations ("*result in,*" etc.). From this perspective, "*pollution*" should be represented as a relation node. Elsewhere, "*pollution*" can design the polluting substances, that is the concept of a material entities, not a relation, and should be an element of the concept set C. In fact, the definition above shows that the concept of pollution is indissolubly a subsystem composed of several other interrelated concepts. In such situation, we encapsulate the fully encoded subsystem in a concept-type node nested in the conceptual graph. The larger node with label "*pollution*" is in relation with some of its internal conceptual nodes ("*man,*" "*substances,*" etc.) if such description is in the sentences.

This way to proceed is relevant because legal norms establish deontic relations ("*oblige,*" "*permit,*" "*prohibit,*" etc.) with other actions (belonging to relation node set) or with concepts derived from actions (like pollution, activities, development, growth, conservation, etc.). Several other expressions—like "*growth of international trade,*" "*fisheries industries,*" "*development of all countries*"—obviously designate complex systems which precise components and relations are neither assignable nor specified. They are represented as an empty concept-like node (with the corresponding label) and nested in the conceptual graph.

It is worth commenting on the anaphoric relations carried by pronouns. Special care must be taken to encode pronouns as they replace a term, an expression or a noun phrase. A relation node is set between the antecedent (the replaced term) and the other term syntactically associated to the pronoun. A frequent case is when a possessive pronoun backed by a concept (as in "*its exclusive economic zone*") refers to the entity which includes the entity designated by this concept (in our example the "*coastal state*"). Whatever the overall syntactic structure of the sentence in which these two components appear, a relation is established between them with the label "*of*" (the result reads "*exclusive economic zone of coastal state*"). This procedure naturally modifies the paths linking concept nodes and their distance.

At the end, each conceptual graph represents a sentence or proposition as it is interpreted, in the sense that: (a) words or expressions are classified as relations or concepts (a property that will ease the determination of any conceptual neighbourhood); and (b) all syntactic ambiguities are resolved (sentences being often decomposable in several distinct syntactic trees). The encoding of sentence in conceptual graphs is a task of knowledge extraction from natural language that is reputedly difficult to perform automatically (with in particular low recall performances; [49, 50]. Doing it manually provides the required data for achieving our objective and incidentally establishes some kind of standard reference for further work on computer-based knowledge extraction for texts.

Warned of these difficulties and equipped with the procedures described above, we choose to encode by hand each sentence mentioning the target term x as a conceptual graph $g^k_{CONV}[x]$. These elementary graphs are disconnected from each other as results from the building procedure (which is sentence-based). Their interconnection would be possible, for example based on the concept nodes that they have in common, but this would provide no additional information on the sought neighbourhoods of the target term.

We define the graph of a convention $CONV$ related to term x, $G_{CONV}[x]$, as the set of the $k = 1..K$ elementary graphs $g^k_{CONV}[x]$. These graphs encode the deep structure of the knowledge carried by the sentences in tree form (without cycle), and sometimes include cycles (for e.g., induced by anaphoric relations and co-references). In all cases, it is possible to follow the paths which pass through each noun phrase which includes the word "resource" or an associated frozen expression, and which, through verbs–or more generally the nodes of relations, link them to other concepts. The vocabulary of $G_{CONV}[x]$, $V(G_{CONV}[x])$, is defined as the union of the vocabularies of its elementary graphs. $G_{UNCLOS}[resource]$ and $G_{CBD}[resource]$ gather 66 and 20 elementary graphs[7], respectively. As we shall see, only a subset of their vocabularies are involved in defining the neighbourhood of the "*resource*" term in the UNCLOS and CBD conventions.

LEXICAL NEIGHBOURHOODS

Paths

We define the neighbourhood of the target term x in a given convention as the set of labels of the conceptual nodes belonging to all paths passing through x in the elementary graphs associated with this convention. This definition clearly excludes labels of relation nodes from lexical neighbourhoods. It also leads to discard those terms that are not linked to the target term x through the knowledge representation. In particular terms that are not in the same proposition are not included in the lexical neighbourhood. An example is given in the sentence below, where terms neighbouring target term "*resource*" are in bold:

*"Each **Contracting Party** [-3] shall take legislative, administrative or policy measures, as appropriate, with the **aim** [-2] that **Contracting Parties** [-1], in particular those that are developing countries, which provide genetic resources are provided **access** [-2] to and **transfer** [-2] of **technology** [-1] which makes use of those resources, on mutually agreed terms, including technology protected by patents and other intellectual property rights, where necessary, through the provisions of Articles 20 and 21 and in accordance with international law and consistent with paragraphs 4 and 5 below." (CBD, Art. 16 Access to and Transfer of Technology,* §*2).*

The number in brackets after each term indicates the distance to the occurrence of the target term ("*resource*") to which it relates (a negative sign indicates a predecessor, a single plus a successor). The distance or rank only counts the conceptual nodes of the graph which separates the neighbouring term from the target term along the path. In other words, in a relation of type yRx where x and y are concepts and R a relation, the distance from y to x is -1, or y has rank -1 with regard to the target x.

The example illustrates an important aspect of our method: the neighbours of an expression are identified along paths in the conceptual graph, not as neighbours in the raw sentence as is often done. The neighbours are selected following knowledge representation links captured in conceptual graphs, not positional information provided by the sentence. For this reason, though "*developing countries*" is a conceptual node of the graph, and near from "*genetic resources*" in the sentence, it is not in its neighbourhood as it refers to "*contracting parties*" and as such, occurs on another path of the graph. "*access,*" "*transfer*" and "*technology*" logically refers to "*those resources*" not directly to "*genetic resources,*" hence the negative ranks (even if the demonstrative "*those*" indicates that these resources are genetic resources).

[7]Graphs containing only the expressions "human resources" or "financial resources" are omitted.

This approach conforms to the distributional hypothesis of semantics that assumes that terms occurring in similar contexts have similar meaning, but the underlying topology we use is defined from conceptual graphs representing knowledge embedded in a sentence or proposition. It is beyond the scope of this study to decide whether such distributional neighbourhoods authentically define the meaning of a word or phrase, or ultimately only allow the assessment of similarities of meanings (see Sahlgren [51], and references within). But in any case, the comparison of lexical neighbourhoods should allow us to detect possible idiosyncratic uses of the same term.

Neighbours and Ranks

The above example also shows the importance of distinguishing between frozen expressions. The term "*genetic resources*" supposedly does not mean the same thing as "*living resources,*" "*natural resources,*" or "*mineral resource.*" An ontology could relate all these terms to the generic class of resources. However, it is obvious that the contexts of use of each of these expressions will differ greatly, depending on the uses that are made of these resources or on the measures and regulations implemented for their management. We can only compare neighbourhoods attached to the same expression or to expressions supposedly referring to the same concept.

Moreover, whether or not to use an expression in a convention is already informative on the field covered by the legal instrument. Likewise, and more significantly, the number of occurrences of an expression provides a first indicator of the lexical–and therefore conceptual–landscape in which the text constructs and moves. Optionally, this number of occurrences can be normalised by the length of the text (evaluated in number of words), then making it possible to compare occurrence densities (remember that the text of the UNCLOS is much longer than that of the CBD). We will therefore compare sets of target terms or expressions (those using the word "*resource*") in order to better define the regulated domain, and sets of neighbouring expressions relating to each target term in each convention to detect possible idiosyncratic uses of terminology. We will also use information taken from the rank matrices which values indicate the number of occurrences of a neighbour expression at a given rank.

The use of a neighbour's rank (its distance from the target expression along the path) is justified with the idea that the more distant a term is, the less it contributes to the (distributional) meaning of the target. Another possible use is to identify frozen or semi-frozen expressions (such as for example "*resources of the exclusive economic zone*") which appear frequently, or that themselves include frozen expressions, the phrases being often nested (as in "*areas beyond national jurisdiction*" which already has the acronym ABNJ in use, and now BBNJ for "*biodiversity in ABNJ*"). Indeed, the occurrence of an expression at a preferential rank from a target (which is a statistically detectable behaviour) suggests the presence of a frozen expression, at least in the analysed corpus.

Let F_{CONV} be the set of the frozen expressions that include the term "*resource*" found in convention *CONV* (*UNCLOS* or *CBD*), with cardinality $|F_{CONV}|$. $N_{CONV}[x]$ denotes the set of terms or expressions (labels of conceptual nodes) found in the neighbourhood of target expression x in the convention. We limit the set to terms with rank in the interval $[-4, +4]$. By construction, $N_{CONV}[x]$ is a subset of the vocabulary $V(G_{CONV}[x])$ (see Section Encoding Sentences in CGs). From $N_{CONV}[x]$ we derive the set $\tilde{N}_{CONV}[x]$ by substituting, when necessary, each single word or word entering in an expression of the set for its lemmatized form (e.g., "states parties" –> "sate party") using the NLTK lemmatizer [24] based on WordNet [52].

A rank/occurrence index $I^x_{CONV}(y)$ is associated to each neighbour y of target x in convention *CONV* as given by:

$$I^x_{CONV}\left(y\right) = 100 \times n^{-1}(x) \sum_{j=1}^{n(y)} \left|r_j(y)\right|^{-1} \quad (1)$$

where $n(x)$ [resp. $n(y)$] is the number of occurrence of target term x (resp. neighbour term y) and $r_j(y)$ the rank of the jth occurrence of y. The index is built such that if y occurs only with rank +1 or−1 and whenever x occurs, then $I^x_{CONV}\left(y\right) = 1$. The contribution of each occurrence of y to its rank/occurrence index is inversely proportional to its rank or distance to x: on the average, more distant terms have a lower rank/occurrence index than nearer terms. The rank/occurrence index provides an easy way to compare the contribution of each neighbour expression y to the distributional meaning of target expression x. It is used to identify the most import terms contributing to the meaning of x as it is used in the context of a given convention.

MEANINGS OF "RESOURCE" IN UNCLOS AND CBD

Target and Neighbour Expressions

CBD and UNCLOS have four and seven expressions, respectively, using the word "resource"[8], forming the following target expression sets:

T_{CBD} ={biological resource, genetic resource, natural resource, resource}

T_{UNCLOS} ={living resource, marine resource, mineral resource, natural resource, non-living resource, resource, resource deposit}

The difference in these two sets results from the difference in the domains covered by the two conventions, as could be expected. But it also suggests that the links or interactions between activities in one domain and the resources of the other domain, are not considered in the conventions. In particular, unless the "*living resource*" of UNCLOS can be interpreted as an expression synonym to the "*biological resource*" of CBD, the exploration and exploitation of mineral and/or non-living resource is not considered in relation to the "*biological resources*" in the UNCLOS; and reversely the protection or conservation of biological diversity is not envisioned in the CBD in relation with the activities regulated by UNCLOS.

For each target expression of the sets T_{CBD} and T_{UNCLOS}, we find the five nearest expressions defined as the neighbour expressions with highest rank/occurrence indexes (see equation 1) in a given convention. These nearest neighbours are listed in **Table 1**.

The vocabulary formed by all expressions close to the targets, found in CBD (resp. in UNCLOS), comprises 81 (resp. 203) expressions or terms, forming 145 (resp. 488) pairs[9] with one of the four (resp. seven) target-expressions. Thirty-seven of them appear in **Table 1**, indicating some partial overlap of the sets of neighbour expressions. "*Genetic resource*" and "*biological resource*" are used in conjunction with the most varied sets of neighbour expressions in the CBD (with, respectively, 51 and 27 neighbours). In this aspect, the expressions "*resource*" and "*living resource*" occupy the first places in the UNCLOS (with, respectively, 82 and 80 neighbours).

Most terms in **Table 1** refers to actors (State, coastal State, country, Party, etc.), to the geographical zones or territories delimited on a jurisdictional basis (exclusive economic zone, the Area, seabed or subsoil–implied "*of the Area*" or "*of/in the EEZ,*" but also explicitly "*jurisdiction*" and "*limit of national jurisdiction*") and to their rights (sovereign right, access). Most of the other expressions concern activities and capabilities, or some resources (polymetallic nodule, natural resource, mineral). These features indicate quite clearly that resources, whatever their type, are well-understood from the angle of law, in the legal genre of discourse.

Now consider the two target expressions shared by the two conventions. The term "*natural resource*" is very little used in CBD. Its only two neighbours are roughly the same as the two closest neighbours in UNCLOS (although 42 related expressions are identified in this convention): "*sovereign right,*" and "*Sate*" in CBD *vs.* "*coastal State*" in UNCLOS. The convergence of the distributional meaning of the expression "*natural resource*" between the two conventions is plausible, even if statistically poorly documented. The alignment, at least partial, of these meanings probably corresponds to an ontologically generic use of this term. In fact, CBD Article 15 §1 indirectly states that genetic resources are natural resources. For its part, UNCLOS Art. 56 §1 includes living and non-living resources under the natural resources, and Article 77 §4 adds mineral resources in the context of Part VI of the convention.

The sets of terms close to the target "*resource*" found in CBD and UNCLOS are disjointed. No similarities seem to emerge. The conceptual landscape built by the CBD around the term "*resource*" is based on the notion of actor (the State), of his role and powers. UNCLOS rather stresses on the activities, resources and the location where they both are or take place.

For the reasons explained at the beginning of this section, it is also important to see whether the expressions "*biological*

[8]The two expressions ≪*financial resource*≫ and ≪*human resource*≫ are discarded.

[9]For example, the pair "natural resource" (target expression) and "exclusive economic zone" (rank +2) appearing from UNCLOS Art. 56 where are mentioned the "natural resources (...) of seabed"–(under understood) "of the exclusive economic zone," the latter indication coming from the beginning of the article.

TABLE 1 | Target expressions (1st column), convention (2d column), and neighbour expressions ranked (1 to 5) by decreasing rank/occurrence index (given above each expression). In bracket after the convention acronym, the number of neighbour expressions from which the nearest expressions are found.

Target expression	CONV	1	2	3	4	5
Biological resource	CBD [27]	8.3 {AC} Contracting party	7.4 {AC} State	7.4 {PA} Method for sustainable use	7.4 {MR} Biological diversity	5.6 {NO} Sovereign right
Genetic resource	CBD [51]	24.3 {AC} Contracting party	11.8 {PA} Access	4.9 {AC} Country	3.9 {MR} Technology	3.9 {AC} Party
Living resource	UNCLOS [80]	33.1 {MR} EEZ	16.6 {AC} Coastal state	8.6 {AC} State	6.2 {MR} Region	6.2 {MR} Subregion
Marine resource	UNCLOS [10]	10.0 {CR} Technol. capacity	10.0 {PA} Technol. assistance	10.0 {CR} National capability	10.0 {MR} Economic benefit	10.0 {AC} Developing coastal state
Mineral resource	UNCLOS [13]	15.4 {MR} Polymetallic nodule	11.5 {AC} Developing state	11.5 {CR} Qualification	11.5 {MR} Continental shelf	7.7 {MR} The Area
Non-living resource	UNCLOS [17]	11.8 {MR} Natural Resource	7.8 {AC} Coastal state	5.9 {MR} Subsoil	5.9 {MR} Seabed	5.9 {MR} Continental shelf
	CBD [2]	50.0 {AC} Sovereign right	50.0 {NO} State	–	–	–
Natural resource	UNCLOS [42]	8.1 {NO} Sovereign right	4.8 {AC} Coastal state	4.7 {CR} Significance	4.7 {MR} Seabed	4.7 {MR} Non-living resource
Resource	CBD [9]	16.7 {AC} State	16.7 {NO} Sovereign right	7.4 {NO} Environmental policy	6.5 {NO} Jurisdiction	6.5 {PA} Control
	UNCLOS [82]	12.8 {MR} The Area	2.9 {PA} Activity in the Area	2.4 {PA} System of exploration & exploitation of resource	2.4 {NO} Sovereignty	2.4 {MR} Mineral
Resource deposit	UNCLOS [5]	20.0 {MR} The Area	20.0 {NO} Limit of national jurisdiction	20.0 {AC} Coastal state	20.0 {PA} Activity in the Area	10.0 {NO} Jurisdiction

In braces, the type of the designated entity (see text). EEZ, "exclusive economic zone"; "qualification" relates to some competencies of actors; "significance" relates to some resource of activity with regard to some actor.

resource" (CBD) and "living *resource*" (UNCLOS) designate the same concept or not. The main features of the conceptual landscape of the first expression concerns actors ("*contracting party,*" "*state*") then some cognitive resource, the biological diversity (interpretable as a material resource) and norm ("*sovereign right*"). The notion of "*living resource*" in UNCLOS is centred on geographical sets ("*EEZ,*" "*region,*" "*subregion*") and actors ("*coastal State,*" "*State*"). These sets are to be related to the Sates' jurisdiction or location. This comparison shows the importance of State actors in relation with biological and living resource, but diverge on the other determinants, UNCLOS insisting on a geographical mapping of resource locations or of actors' cooperation, while CBD focuses on actor's rights and compatibility of resource uses and preservation of biological diversity. However, this analysis mostly relies on an interpretation of the full sentences.

Lattices and Distributional Meaning

Is there a vector space where this kind of analysis can be done from the sole information of **Table 1**? A simple approach to word embedding is to associate a dimension of vector space with each term (here, each neighbouring expression). In such configuration, the target expression is represented by a vector in the subspace spanned by its neighbouring expressions, each value of the rank / occurrence index being a coordinate. The cardinality of the union of the sets of neighbouring lemmatized expressions of the two conventions (see **Table 1**) is given by:

$$\left|\cup_{CONV=CBD,UNCLOS}\left(\cup_{x\epsilon F_{CONV}}\tilde{N}_{CONV}[x]\right)\right| = 37 \qquad (2)$$

Thus, the set of all neighbouring expressions defines a vector space of dimension 37, populated by 11 vectors, each representing a target expression (the same expression considered in two conventions is represented by two vectors). Since none of these vectors occupy exactly the same subspace, they are two-by-two orthogonal. Calculating the cosine of the angle between two vectors then provides no information on the similarity of the (distributional) meaning of the expressions they represent. Vector space reduction techniques[10] that would allow these angles between word or expression vectors to be calculated are of little interest here: they are relevant for large sets of vectors.

[10]Like singular value decomposition.

Moreover, after reduction, the new dimensions of the embedding space do not correspond to lexical elements and are therefore no longer interpretable.

We propose the following approach. Each neighbouring expression is attached to a generic ontological class. All of these classes is a kind of coarse taxonomy, which induces a partition of all 37 related expressions. We distinguish the following classes[11]: actor ("AC" label), material resource ("MR" label), cognitive resource ("CR" label) and process or activity ("PA" label). As the analysis relates to international law conventions, all entities relating to a type of legal norm (understood in a loose sense) are attached to a fifth class "norm" (label "NO"). The class relevant to each expression is given in **Table 1**.

No longer considering neighbouring expressions but their classes, **Table 1** corresponds to another matrix: each row is a target expression in the context of a convention, and each column indicates the class of an expression of its lexical neighbourhood. The corresponding mathematical structure is a lattice, similar to those used in Formal Concept Analysis (FCA [54–56]). The lattice presents a double nested hierarchy established between the 11 target expressions (or "objects" in the FCA language) and the five classes ("attributes" in FCA) to which the neighbouring expressions belong. Some line diagrams representing this lattice are presented in **Figure 2**. Any target expression of a vertex which is on a descending path from a vertex with an attribute (class) have this attribute (and is an element of the extent of the attribute). Conversely, any attribute of a vertex lying on an ascending path starting from a target expression, is the class of one of the target's neighbour expressions (and is an element of the intent of the expression). To simplify the figures, the reduced representation of the lattice is used here: class (resp. target expression) labels are given only in the node occupying the higher (resp. lower) position they appear in the hierarchy (therefore, some nodes do not have an apparent label). The lattice immediately shows that none of the four attributes is an attribute of all target expressions, and that none of the target expressions collect all the attributes in its profile (set of its attributes).

Let us go back to the comparison of the terms "*biological resource*" of CBD and "*living resource*" of UNCLOS. The two sub-lattices linked to each of these targets are highlighted in **Figure 2**. Note first that the living resource and non-living resource of UNCLOS have, in this rough taxonomy, the same profile (they occupy the same vertex). This profile (actor AC and material resource MR) is a subset of the biological resource profile of the CBD, the latter also requesting the classes "norm" (NO) and "process and activity" (PA) in its lexical neighbourhood. Considering that we have restricted these neighbourhoods to the only 5 expressions with highest rank/occurrence indices, this difference between profiles is a significant feature of the semantic difference of the two expressions: the biological resources of the CBD cannot be assimilated. to the living resources of UNCLOS. Moreover, under the aspect of this taxonomy of neighbours, biological resources dominate the

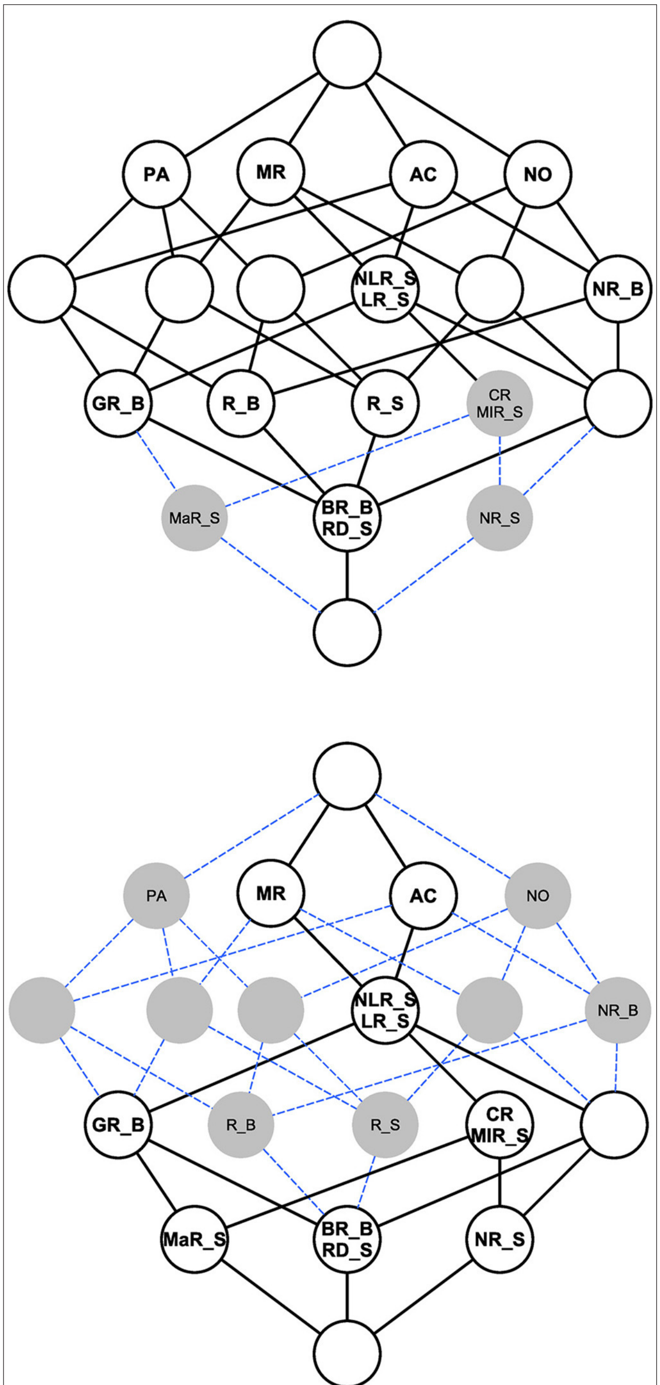


FIGURE 2 | Reduced representation (see text) of the expressions/classes lattice. (Top) Part of the lattice linked with the CBD "biological resource" expression (label BR_B); (bottom) Part of the lattice linked with the UNCLOS "living resource" expression (label LR_S). Labels combine the final B (from "biodiversity") for CBD or S (from "sea") for UNCLOS (sea) and the following sub-labels: GR, genetic resource; MaR, marine resource; MiR, mineral resource; NLR, non-living resource; NR, natural resource; R, resource; RD, resource deposit. Generic ontological classes are "actor" (AC label), "material resource" (MR label), "cognitive resource" (CR label), "process or activity" (PA label), and "norm" (NO label). Grey nodes and blue dashed links do not belong to the sub-lattice (but to the overall lattice) (figures built with the free software Concept Explorer 1.2; [57]).

[11]This taxonomy is inspired by the meta-model developed for the modelling of socio-ecological systems by Sibertin-B1anc et al. [53] to which we add here the "norm" class.

three other kinds of resources ("*genetic resources,*" "*natural resources,*" and "*resources*") regulated by the CBD. Surprisingly, it is the "*resource deposit*" of UNCLOS which presents the same profile as "*biological resources*" of CBD while dominating "*non-living resource*" (and its paired expression "*living resource*") and "*resource,*" expressions used in UNCLOS.

The sub-lattices corresponding to CBD "resource" (label R_B) and UNCLOS "resource" (label R_S) are exhibited on **Figure 3**. On the CBD diagram, "*resource*" is subsumed by "*biological resource,*" and have a larger profile (subsumes) than "*natural resource.*" In UNCLOS, "*resource deposit*" subsumes "*resource.*" Both CBD and UNCLOS "*resource*" have the attributes "*process and activity*" and "*norm.*" CBD "*resource*" adds the "*actor*" attribute while UNCLOS "*resource*" adds "*material resource.*" It must be kept in mind that this kind of subsumption relationships (or hyponym-hypernym relationship) is only valid in the context set by the coarse taxonomy used in this study. These differences emerging from the analysis of the distributional meaning established on the basis of a somewhat minimalist neighbourhood of expressions, they apparently express very distinct cognitive orientations as to what are resources, for the one or the other of these conventions.

The sub-lattice associated to the attribute "norm" is shown in **Figure 4**. The extent of "norm" is the union of the sets {"*resource,*" "*resource deposit,*" "*natural resource*"} of UNCLOS expressions and {"*natural resource,*" "*resource,*" "*biological resource*"} of CBD expressions. All other target expressions, "*genetic resource*" from CBD, "*living resource,*" "*non-living resource,*" "*mineral resource,*" and "*marine resource*" from UNCLOS, are not connected to the "*norm*" class in this simple taxonomy.

The link with the concepts attached to normativity is made in relation to the relatively general expressions involving the resources, rather than with their derivations of a more technical or specialised character.

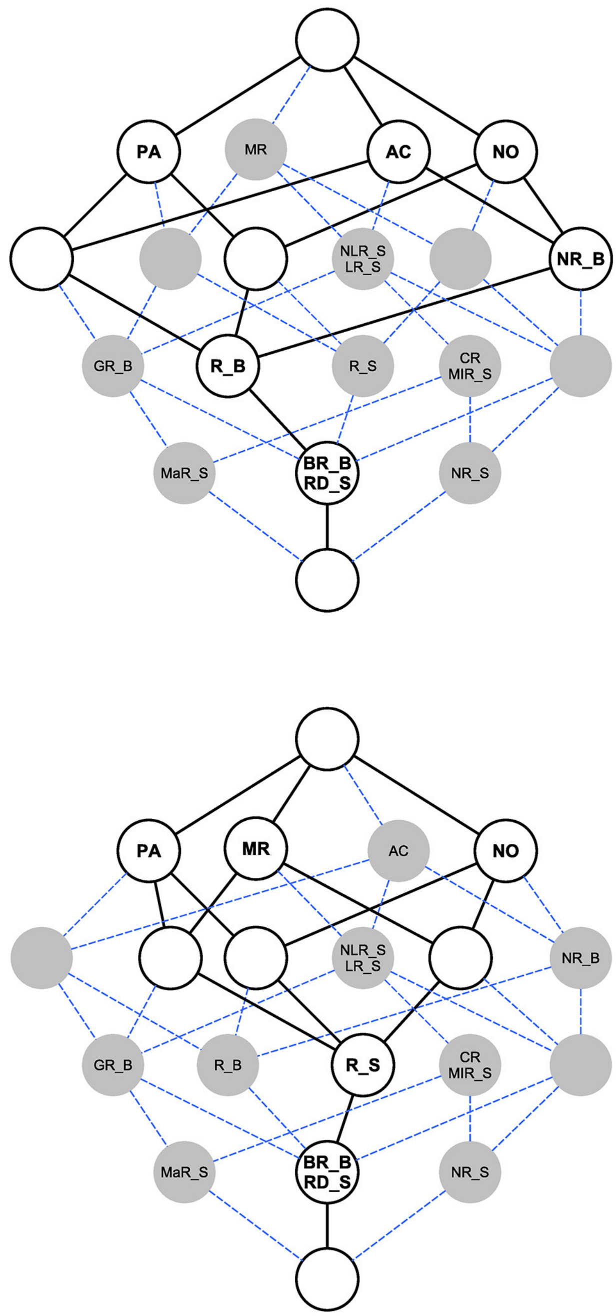


FIGURE 3 | (Top) Part of the lattice linked with the CBD "resource" expression (label R_B); (bottom) Part of the lattice linked with the UNCLOS "resource" expression (label LR_S). Reduced representation, see **Figure 2** (figures built with the free software Concept Explorer 1.2; [57]).

DISCUSSION

Where has the search for a plausible answer to the original question led us? The experimental study argues for a notable and observable difference in the meaning of the same expression in two conventions of international environmental law. Legal language is not free from internal lexical idiosyncrasies, within the legal genre itself. Definitions are not sufficient to contain the meaning of legal terms or expressions. Distributional semantics provides complementary analysis tools, sensitive to the different contexts of use of these expressions. This result could moreover constitute only the first cog in a progression showing that the emergence of meaning, going up the levels of segmentation of texts until the constitution of a legal discourse, accumulates epistemological divergences between distinct legal currents.

Certainly this first conclusion needs to be confirmed on the basis of a more extensive textual corpus and diversified sources of rights. The analytical method outlined here can be replicated and supported by the use of various NLP and knowledge extraction tools. However, some features and some consequences of our analysis deserve further discussion.

Lexical Idiosyncrasy and Statistical Measures

Objectively, the description of the distributional meanings of the term "*resource*" and associated expressions rests on a statistically

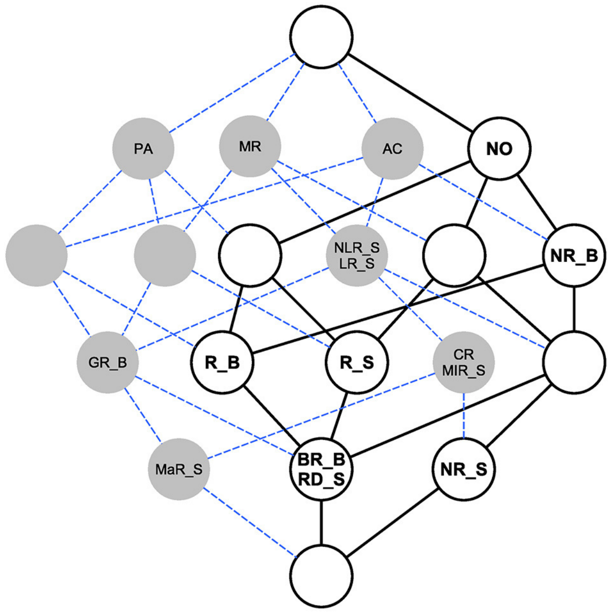


FIGURE 4 | Part of the lattice linked with the attribute "norm" (label NO). Reduced representation, see **Figure 2** (figure built with the free software Concept Explorer 1.2; [57]).

fragile basis. This situation is not a flaw in the approach but rather reflects the essential condition for analysing small corpora. The study made it possible to explicitly describe the contexts of use of each target expression and their contributions to its distributional meaning. As such, the results are likely to warn against any approach to normative texts which would dispense with a reflection on the variability of the meaning of words and on the dependence of this meaning on the regulatory context, at the risk of misinterpreting texts and depart from the intentions of legislators. The differentiated meanings of the same word used in different conventions reveal blind spots in international environmental law that need to be addressed.

These observations also have implications for the methodological aspect of the analyses. Any physical or informational measure (e.g., posterior probabilities, information functions, entropy) based on the frequency of occurrence of phrases or words is potentially affected. Indeed, counting of occurrences implicitly presupposes that the meaning attributed to an expression does not vary in the corpus. This hypothesis is not always valid. In these shifts in meaning, legal and political cultures specific to the regulated field and to the agents who design and promote legal instruments are expressed. The estimate of the frequencies of occurrences and the derived statistical measures remain valid, but on condition that they are applied to sets of semantically or ontologically homogeneous utterances, sets whose limits are traceable via the analysis of the deep structures of language and identification of the professional or "epistemological affiliation" of the authors.

Conceptual Graphs and Legal Text Formalisation

On a technical level, the purpose of using conceptual graphs is generally to build up a knowledge base that can then be queried (to answer questions or produce new knowledge) via machines. Our posture is different: the work of formalising legal proposals, sentences or articles via conceptual graphs creates the conditions for an interrogation in direct contact with the legal matter (data), on mechanisms, artifices and techniques–implicit or explicit, intentional or unconscious, known or hidden–used by "the legislator" in the production of normative texts. Even if the theory of conceptual graphs cannot claim the universality of its capacities to transcribe any text into natural language and therefore presents limits of applicability, the formalisation exercise offers the opportunity to explain a part of the latent cognitive options which govern the choice of expressions in natural language and their conceptual underpinning. In this process, the nature of these revealed choices makes it possible to question the clarity and distinction of the concepts used and, admittedly a more adventurous steps, to try to understand the consequences of these choices.

The Same Language but Different Lexical Meanings

The language used by the various international conventions is the same: for example, the English versions of the textual corpus. The lexicon, except for any technical terms specific to each legal (or related scientific) field, is also the same. But the use of certain key terms is differentiated according to conventions. A term appears in a convention in one or two types of occurrence: (a) as a word or part of a phrase in sentences or clauses; (b) as an entry of a definition. A definition is a short text, usually presented in intention, and positions itself between hypo and hyper-specificity. Thus, regardless of the method used, the comparison of the definitions by two conventions of the same given term, provides little insight into the analysis. On the other hand, the notional context in which the term is inserted through its uses is rich in lessons. If the notional contexts of use of a term in two conventions differ significantly as we have seen, the hypothesis of idiosyncratic conceptions attached to each convention is necessary. Such variations of the meaning of a term convey a semantic meaning specific to a given text.

But the analysis shows yet another thing. The terms of the notional neighbourhood suitable for the use of a target term in a convention can be related to a taxonomy, as we have done. However, a comparison shows a disparity in the ontological anchoring of expressions more or less frozen around the same central term (here the term "*resource*"). Belonging to a community of lexical expressions therefore in no way guarantees an alignment of the ontological bases for the design of the signified things. The ontological anchors of terms are established implicitly and along with a quasi-fortuitous lexical choice of particular terms in the development of legal discourse [12].

[12] It is not in the scope of this study to know if this state of affairs is a good or a bad, a defect or a quality.

Terminological Idiosyncrasy and Normativity

The text of a convention finalises a process of consultation and negotiation between actors (delegations) duly mandated and having not only different aims, but also different implicit knowledge or backgrounds (this also between members of the same delegation or group). However, this same text is also the starting point for new phases of negotiations, amendments or extensions. It constitutes the linguistic and cognitive reference for these developments carried out during conferences of the Parties or brought by *ad hoc* working groups. Even if, as we can see, the corpus of texts produced during these developments and attached to the source convention is enriched with new notions, the base on which these notional expansions necessarily rest, remains the inaugural notional and relational landscape established by the convention.

While staying mostly confined to the ”cognitive cone“ projected by the convention, the subsequent work carried out under its aegis reinforces (in the sense of learning) the significance of this landscape, freezing its contours and internal structuring. The use of a particular expression that was specific in the beginning, in a distinctive context, becomes idiosyncratic. Its fictitious neighbourhood is fully functioning, at the cost of increasing relegation in an implicit and unthinkable context, therefore sheltered from possible questioning. Thus, the normativity of the convention is intentionally expressed in the legal instruments that it establishes, but also, and in a latent and perhaps more profound way, in the structuring of the interrelationships between concepts that it explicitly invokes and whose idiosyncratic meanings are made to be reproduced and to persist over time.

Back to UNCLOS and CBD

The follow-up to the negotiations on the new sea treaty shows the difficulty of reaching a consensus capable of both meeting the contrasting expectations of States and of building a new governance of the oceans [58] while preserving some older institutional structures, and to produce an effectively binding and efficient instrument, responding to the urgent need for BBNJ regulation [59]. On the level of words and concepts alone, the terminology of the treatise has also been debated [60], but the most remarkable fact is the shift towards a new lexical set linked to “*resources.*” Indeed, in the “Revised draught text of an agreement under the United Nations Convention on the Law of the Sea on the conservation and sustainable use of marine biological diversity of areas beyond national jurisdiction”[13], after exclusion of expressions “*financial resources*” and “*human resources,*” “*marine genetic resources*” glean about 90% of the occurrences (more than 70 occurrences), against about 10% for the term “*resources,*” and a single occurrence of the term “*biological resources.*”

It will be understood: the new treaty regulates marine genetic resources in the ABNJ. All aspects of the management of other types of resources, especially living or biological resources–not even to speak more prosaically of marine life–are left to the discretion of earlier treaties. No reinforcement of an ontological anchoring of the concepts used is made explicit[14]. The regulatory framework for these other resources and for activities having an impact on these resources remains that set by the UNCLOS and CBD and their Conferences of the Parties, with their terminological and conceptual dissonances, or blind spots.

The abandonment of the development of a rigorous conceptual framework that can support the normative discourse, in favour of lexical choices emerging in a fortuitous way from the development of the text may constitute the price to pay for obtaining a soft consensus around a treaty, an additional piece to a kind of “diplomatic” law [61]. However, the deleterious and irreversible effects on living resources, marine biodiversity and marine life, which an ineffective law would allow to slip through its nets, should not be added to this bill.

CONCLUSION

The analysis of expressions including the term “*resource*” in the CBD and UNCLOS, shows that the two conventions do not use these expressions in the same conceptual landscape. In this sense, they associate them with different meanings. The meaning of an expression is established according to the distributional hypothesis that the meaning of a word mainly emerges from the lexical environment in which it is inserted, from its use in a particular context. Rather than considering the simple alignment of words, we go here through a formalisation of sentences or propositions in the form of conceptual graphs, a step which imposes, among other things, to remove syntactic ambiguities. The neighbourhood of a target expression is then extracted along the paths of the graph which pass through the vertex whose label is this expression. We also take into consideration the distance of each neighbouring concept or notion to the target, in order to penalise the contribution of the most distant expressions in the distributional meaning of the target expressions.

The comparison of the rank/occurrence matrices of the neighbours then makes it possible to evaluate the similarity or disparity of the distributional meanings of the expressions frozen with the word “*resource.*” For this purpose, only the most contributory expressions to the target meaning are retained. The diversity of neighbouring expressions then requires their classification in a partition induced by a coarse taxonomy (with only five classes: actors, material resources, cognitive resources, norms, processes and activities). The neighbourhood comparison structure is a lattice that reveals relationships of subsumption or non-comparability between target expressions considered in the context of each convention.

Beyond highlighting idiosyncratic uses of expressions linked to the notion of resource in international law conventions, the developed method is potentially applicable to a large set of lexical entries. It also shows the disparity in the ontological

[13]See Available online at: https://digitallibrary.un.org/record/3847798?ln=en (accessed January 15, 2021).

[14]Except for ”marine genetic resources“ which are the subject of a definition in the draft Article 1 ”use of terms" of the treaty (with the reservations made in Section UNCLOS, CBD, and the Resource Issue on any definition).

anchors of lexically similar expressions in legal texts, anchors that are both implicit and constructed along with a random lexical choice of particular terms in the development of legal discourse. The normativity of conventions is then expressed at the lexical level in the reinforcement, reproduction and persistence of these distributional meanings and of their fortuitous ontological basis.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analysed in this study. The data include the texts of the two conventions available on their respective sites (CBD: https://www.cbd.int/; UNCLOS: https://www.un.org/depts/los/convention_agreements/texts/unclos/unclos_e.pdf).

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

FUNDING

With funding from United Kingdom Research and Innovation (UKRI) Global Challenges Research Fund (GCRF) One Ocean Hub (Grant Ref: NE/S008950/1).

ACKNOWLEDGMENTS

I would like to thank professors P. Vivo, D. M. Katz, and J. B. Ruhl for giving me the opportunity to present this work within the framework of the conference The Physics of Law, November 12–13, 2020. The manuscript has been greatly improved thanks to discussions with Dr C. Lajaunie (environmental lawyer, INSERM - France), for which I am very grateful. My thanks also to Prof. E. Morgera (SCELG, University of Strathclyde) PI of the One Ocean Hub UKRI project, for the possibility of taking part into an international project on the oceans. I thank the two anonymous reviewers for allowing me to improve this manuscript through comments which represented a challenge that I had to take up.

REFERENCES

1. Schuck PH. Legal complexity: some causes, consequences, and cures. *Duke Law J.* (1992) 42:1–52. doi: 10.2307/1372753
2. White M. Legal complexity and lawyers' benefit from litigation. *Int Rev Law Econ.* (1992) 12:381–3. doi: 10.1016/0144-8188(92)90016-K
3. Epstein RA. *Simple Rules for a Complex World*. Cambridge, MA: Harvard University Press (1995).
4. Conseil d'Etat. *Rapport Public 2006 - Sécurité Juridique et Complexité du Droit, Jurisprudence et Avis de 2005.* Paris: Etudes et Documents du Conseil d'Etat, *n.* 57 (2006).
5. Doat M, Le Goff J, Pédrot PH. *Droit et Complexité – Pour une Nouvelle Intelligence du Droit Vivant*. Rennes: Presse University (2007).
6. Ruhl JB. Law's complexity: a primer. *Geor State Univ Law Rev.* (2008) 24:885–912. Available online at: https://readingroom.law.gsu.edu/gsulr/vol24/iss4/9
7. Jungnickel D. *Graphs, Networks and Algorithms. Algorithms and Computation in Math. Vol. 5.* Berlin: Springer (1999). doi: 10.1007/978-3-662-03822-2
8. Watts DJ. *Small Worlds: The Dynamics of Networks between Order and Randomness.* Princeton, NY: Princeton University Press (2003).
9. Newman M, Barabási AL, Watts JD (eds). *The Structure and Dynamics of Networks*. Princeton, NJ: Princeton University Press (2006).
10. Brandes U, Erlebach T. *Network Analysis - Methodological Foundations*. Berlin: Springer (2005). doi: 10.1007/b106453
11. Cox DJ, Holder LB. *Mining Graph Data.* New York: NJ: Wiley & Sons (2007).
12. Bourcier D, Mazzega P. Toward measures of legal complexity. In: *Artificial Intelligence and Law, Stanford Law School.* New York, NY: ACM Press (2007). p. 211–5.
13. Bommarito MJ, Katz DM. A mathematical approach to the study of the United States code (March 25, 2010). *Phys A.* (2010) 389:4195–200. doi: 10.1016/j.physa.2010.05.057
14. Bourcier D, Mazzega P. Codification law article graphs. In: Lodder AR, Mommers L, editors. *Legal Knowledge Information Systems*. Amsterdam: IOS Press (2007). p. 29–38.
15. Boulet R, Mazzega P, Bourcier D. Network analysis of the french environmental code. In: Casanovas P, editor. *AI Approches to the Complexity of Legal Systems*. Heidelberg: LNASpringer (2010). doi: 10.1007/978-3-642-16524-5_4
16. Boulet R, Mazzega P, Bourcier D. A network approach to the french system of legal codes- part i: analysis of a dense network. *Artif Intell Law.* (2011) 19:333–55. doi: 10.1007/s10506-011-9116-1
17. Boulet R, Mazzega P, Bourcier D. Network approach to the French system of legal codes. Part II: the role of the weights in a network. *Artif Intell Law.* (2018) 26:23–47. doi: 10.1007/s10506-017-9204-y
18. Lajaunie C, Mazzega P. Mining CBD. *Braz J Int Law.* (2016) 13:277–92. doi: 10.5102/rdi.v13i2.4058
19. Lajaunie C, Mazzega P. One health and biodiversity conventions. The emergence of health issues in biodiversity conventions. *IUCN Acad Environ Law J.* (2016) 7:105–21. Available online at: http://www.iucnael.org/en/e-journal/current-issue-
20. Lajaunie C, Mazzega P, Boulet R. Health in biodiversity-related conventions: analysis of a multiplex terminological network (1973-2016). In: Chen SH, editor. *Big Data in Computational Social Science and Humanities.* Springer (2018). p. 165–82. doi: 10.1007/978-3-319-95465-3_7
21. Boleda G, Herbelot A. Formal distributional semantics: introduction to the special issue. *Comput Linguist.* (2016) 42:619–35. doi: 10.1162/COLI_a_00261
22. Crandall J, Charrow VR. *Linguistic Aspects of Legal Language*. (1990). Available online at: https://eric.ed.gov/?id=ED318248
23. Feldman R, Sanger J. *The Text Mining Handbook: Advanced Approaches in Analyzing Unstructured Data.* Cambridge: Cambridge University Press (2007). doi: 10.1017/CBO9780511546914
24. Bird S, Loper E, Klein E. *Natural Language Processing with Python*. Sebastopol, CA: O'Reilly Media Inc (2009).
25. Maynard D, Bontcheva K, Augenstein I. *Natural Language Processing for the Semantic Web*. San Rafael, CA: Morgan & Claypool Publishers (2017).
26. Francesconi E, Montemagni S, Peters W, Tiscornia D. Integrating a bottom-up top down methodology for building semantic resources for a multilingual legal domain. In Francesconi E, Montemagni S, Peters W, Tiscornia D, editors. *Semantic Processing of Legal Texts. Where the Language of Law Meets the Law of Language*. Vol. 6036. Berlin: Springer (2010). p. 95–121. doi: 10.1007/978-3-642-12837-0_6
27. Wagh RS. Knowledge discovery from legal documents dataset using text mining techniques. *Intern J Comput Applicat*. (2013) 66:32–34. doi: 10.5120/11258-6501
28. Wyner AZ, Peters W, Price D. Argument discovery and extraction with the argument workbench. In: *Proceedings of the 2nd Workshop on Argumentation Mining*. New York, NY (2015). doi: 10.3115/v1/W15-0510

29. Peters W, Wyner A. *Legal Text Interpretation: Identifying Hohfeldian Relations From Text*. In: *Proceeding of LREC* (2016). Available online at: http://www.lrec-conf.org/proceedings/lrec2016/summaries/253.html
30. Robaldo L, Villata S, Wyner AZ, And M. Grabmair. Introduction for artificial intelligence and law: special issue “natural language processing for legal texts”. *Artif Intell Law*. (2019) 27:113–5. doi: 10.1007/s10506-019-09251-2
31. Firth JR. A synopsis of linguistic theory 1930-1955. In: Palmer FR, editor. *Selected Papers of JR*. Bloomingtaon: Indiana University Press(1957) 168–205.
32. World Bank Group. *The Growing Role of Minerals and Metals for a Low Carbon Future.* Washington, DC: World Bank Group/EGPS (2017).
33. Jones N. Sea holds treasure trove of rare-earth elements. *Nature*. (2011). doi: 10.1038/news.2011.393
34. Hein JR, Mizell K, Koschinsky A, Conrad T. Deep-ocean mineral deposits as a source of critical metals for high- and green-technology applications: comparison with land-based resources. *Ore Geol Rev.* (2013) 51:1–14. doi: 10.1016/j.oregeorev.2012.12.001
35. McLellan BC, Yamasue E, Tezuka T, Corder G, Golev A, Giurco D. Critical minerals and energy–impacts and limitations of moving to unconventional resources. *Resources.* (2016) 5:19. doi: 10.3390/resources5020019
36. Sharma R. Deep-Sea mining: current status future considerations. In: Sharma R, editors. *Deep-Sea Mining*. Berlin: Springer (2017). doi: 10.1007/978-3-319-52557-0
37. Ifremer. *Les Ressources Minerals Marines Profondes. Synthèse d'une Etude Prospective à L'horizon 2030.* Issy les Moulineaux (2011).
38. World Ocean Review. *Marine Resources – Opportunities and Risks*. *Future Ocean.* Mare: The International Ocean Institute (2014).
39. Jones DOB, Kaiser S, Sweetman AK, Smith CR, Menot L, Vink A, et al. Biological responses to disturbance from simulated deep-sea polymetallic nodule mining. *PLoS ONE.* (2017) 12:e0171750. doi: 10.1371/journal.pone.0171750
40. Madureira P, Brekke H, Cherkashov G, Rovere M. Exploration of polymetallic nodules in the area: reporting practices, data management and transparency. *Marine Policy.* (2016) 70:101–7. doi: 10.1016/j.marpol.2016.04.051
41. Miller KA, Thompson KF, Johnston P, Santillo D. An overview of seabed mining including the current state of development, environmental impacts, knowledge gaps. *Front Marine Sci*. (2018) 4:418. doi: 10.3389/fmars.2017.00418
42. Sowa JF. Conceptual graphs. *IBM J Res Dev*. (1976) 20:336–57. doi: 10.1147/rd.204.0336
43. Sowa JF. *Conceptual Structures: Information Processing in Mind and Machine*. Boston, MA: Addison-Wesley (1984).
44. Chein M, Mugnier ML. *Graph-Based Knowledge Representation*. London: Springer (2010).
45. Khairullina D. Concept vs notion and lexical meaning: what is the difference? *J Hist Cult Art Res.* (2018) 7:303–13. doi: 10.7596/taksad.v7i5.1929
46. Bellingeri M, Bevacqua D, Scotognella F, Alfieri R, Nguyen Q, Montepietra D, et al. Link and node removal in real social networks: a review. *Front Phys.* (2020) 8:228. doi: 10.3389/fphy.2020.00228
47. Sukthanker R, Poria S, Cambria E, Thirunavukarasu R. Anaphora and coreference resolution: a review. *Inform Fus.* (2020) 59:139–62. doi: 10.1016/j.inffus.2020.01.010
48. Kolhatkar V, Roussel A, Dipper S, Zinsmeister H. Anaphora with non-nominal antecedents in computational linguistics: a survey. *Comput Linguist.* (2018) 44:547–612. doi: 10.1162/coli_a_00327
49. Khodak M, Risteski A, Fellbaum C, Arora S. AutomatedWordNet construction usingword embeddings. In: *Proceedings of the 1st Workshop on Sense, Concept and Entity Representations and their Applications*. Valencia: Association for Computational Linguistics (2017). doi: 10.18653/v1/W17-1902
50. Neale S. A survey on automatically-constructed wordnets and their evaluation: lexical and word embedding-based approaches. In: *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).* Miyazaki (2018). Available online at: https://www.aclweb.org/anthology/L18-1270.pdf
51. Sahlgren M. The distributional hypothesis. *Rivis Linguist.* (2008) 20:33–53. Available online at: http://www.italian-journal-linguistics.com/wp-content/uploads/Sahlgren-1.pdf
52. Fellbaum C, editor. WordNet: an electronic lexical database. In: *Bradford Books*. Cambridge, MA: MIT press (1998). doi: 10.7551/mitpress/7287.001.0001
53. Sibertin-Blanc CH, Therond O, Monteil C, Mazzega P. The entity-process framework for integrated agent-based modeling of social-ecological systems. In: Boulet R, Lajaunie C, Mazzega P, editors. *Law, Public Policies and Complex Systems: Networks in Action, Law, Governance and Technology Series*. Berlin: Springer (2019). p. 57–87. doi: 10.1007/978-3-030-11506-7_4
54. Ganter B, Wille R. *Formal Concept Analysis*: *Mathematical Foundations*. Berlin: Springer (1999).
55. Poelmans J, Elzinga P, Viaene S, Dedene G. Formal concept analysis in knowledge discovery: a survey. In: Croitoru M, Ferré S, and Lukose D, editors. *Conceptual Structures: From Information to Intelligence, LNCS. Vol. 6208*. Berlin: Springer (2010). p. 139–153. doi: 10.1007/978-3-642-14197-3_15
56. Codocedo V, Napoli A. Formal concept analysis and information retrieval—a survey. In: Sacarea JB, Ojeda-Aciego M. *International Conference in Formal Concept Analysis.* Nerja: Springer; LNCS. (2015). p. 61–77. doi: 10.1007/978-3-319-19545-2_4
57. Yevtushenko SA. System of data analysis “concept explorer.” In: *Proceedings of the 7th National Conference on Artificial Intelligence KII-2000*. (2000). p. 127–34.
58. Clark NA. Institutional arrangements for the new BBNJ agreement: moving global, regional and hybrid. *Marine Policy.* (2020) 122:104143. doi: 10.1016/j.marpol.2020.104143
59. Mendenhall E, Elizabeth De Santo E, Elizabeth Nyman E, Tiller R. A soft treaty, hard to reach: the second inter-governmental conference for biodiversity beyond national jurisdiction. *Marine Policy.* (2019) 108:103664. doi: 10.1016/j.marpol.2019.103664
60. Leary D. Agreeing to disagree on what we have or have not agreed on: the current state of play of the BBNJ negotiations on the status of marine genetic resources in areas beyond national jurisdiction. *Marine Policy.* (2019) 99:21–9. doi: 10.1016/j.marpol.2018.10.031
61. Flückiger A. *Le principe de Clarté de la loi ou L'ambiguïté d'un Ideal. Cahiers du Conseil Constitutionnel No 21 Dossier: la Normativité (Janvier 2007).* (2007). Available online at: https://www.conseil-constitutionnel.fr/nouveaux-cahiers-du-conseil-constitutionnel/cahiers-du-conseil-constitutionnel-n-21-janvier-2007 (accessed January 15, 2021).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 28 May 2021
doi: 10.3389/fphy.2021.658463

Measuring Law Over Time: A Network Analytical Framework with an Application to Statutes and Regulations in the United States and Germany

Corinna Coupette[1†], *Janis Beckedorf*[2†], *Dirk Hartung*[3,4*], *Michael Bommarito*[5] and *Daniel Martin Katz*[3,4,5]

[1]*Max Planck Institute for Informatics, Saarbrücken, Germany,* [2]*Ruprecht-Karls-Universität Heidelberg, Heidelberg, Germany,* [3]*Center for Legal Technology and Data Science, Bucerius Law School, Hamburg, Germany,* [4]*CodeX – the Stanford Center for Legal Informatics, Stanford Law School, Stanford, CA, United States,* [5]*Illinois Tech – Chicago Kent College of Law, Chicago, IL, United States*

OPEN ACCESS

Edited by:
Sitabhra Sinha,
Institute of Mathematical Sciences, India

Reviewed by:
Marios Koniaris,
University of Thessaly, Greece
Massimiliano Zanin,
Polytechnic University of Madrid, Spain
Chico Camargo,
University of Exeter, United Kingdom

***Correspondence:**
Dirk Hartung
dirk.hartung@law-school.de

[†]*These authors have contributed equally to this work and share first authorship*

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *25 January 2021*
Accepted: *29 April 2021*
Published: *28 May 2021*

Citation:
Coupette C, Beckedorf J, Hartung D, Bommarito M and Katz DM (2021) Measuring Law Over Time: A Network Analytical Framework with an Application to Statutes and Regulations in the United States and Germany. Front. Phys. 9:658463. doi: 10.3389/fphy.2021.658463

How do complex social systems evolve in the modern world? This question lies at the heart of social physics, and network analysis has proven critical in providing answers to it. In recent years, network analysis has also been used to gain a quantitative understanding of law as a complex adaptive system, but most research has focused on legal documents of a single type, and there exists no unified framework for quantitative legal document analysis using network analytical tools. Against this background, we present a comprehensive framework for analyzing legal documents as multi-dimensional, dynamic document networks. We demonstrate the utility of this framework by applying it to an original dataset of statutes and regulations from two different countries, the United States and Germany, spanning more than twenty years (1998–2019). Our framework provides tools for assessing the size and connectivity of the legal system as viewed through the lens of specific document collections as well as for tracking the evolution of individual legal documents over time. Implementing the framework for our dataset, we find that at the federal level, the United States legal system is increasingly dominated by regulations, whereas the German legal system remains governed by statutes. This holds regardless of whether we measure the systems at the macro, the meso, or the micro level.

Keywords: legal complexity, evolution of law, quantitative legal studies, empirical legal studies, legal data science, network analysis, natural language processing, complex systems

1 INTRODUCTION

Originating from mathematics and physics, complexity science has been successfully applied in the study of social phenomena [1, 2]. More recently, it was introduced as an approach to gain a quantitative understanding of the structure and evolution of law [3]. While legal scholars have long used concepts and terminology from complexity science in legal theory [4–6], research has also called for the development of computational models, methods, and metrics to describe how law evolves in practice [7].

Network analysis, a critical tool for understanding many complex systems [8–10], has proven particularly useful for scientific work answering this call. It has been used, inter alia, to analyze network data derived from decisions by national courts [11–17], international courts [18–24], and international tribunals [25, 26], as well as from statutes (i.e., rules promulgated by the legislative branch of government) [27–33], constitutions [34–36], and international treaties [37–41]. Relevant work in this context explored, for example, which characteristics of complex systems occur in statutory law [27, 30, 31], how references to judicial decisions are used to shape legal arguments [13, 14, 20], or where social dynamics exist between judges or international arbitrators [26, 42]. The network analytical methods employed include centrality measures, clustering, and degree distributions [11, 12, 27, 34, 38]. However, while all studies examine network representations of legal document collections, the data models and methods employed vary widely, which makes it hard to assess the quality of individual results and compare findings across studies. Furthermore, most of this research considers one legal document type only, and some important categories of legal documents, most prominently regulations (i.e., rules promulgated by the executive branch of government with authorization of the legislative branch of government), have—to the best of our knowledge—not received any network analytic attention.

This points to two gaps in the literature: First, on the methodological side, there exists no comprehensive framework for quantitative legal document analysis using network analytical tools. Such a framework should be flexible in three ways: It should (1) produce sensible results for different document types, countries, and time periods, (2) allow us to explore document collections of vastly different sizes, and (3) offer insights on the global (*macro*), intermediate (*meso*), and local (*micro*) level of analysis. Second, on the empirical side, there is a lack of studies that combine multiple legal document types or include regulations.

In this article, we take a step toward filling both gaps. We offer a comprehensive framework for analyzing legal documents as multi-dimensional, dynamic document networks and demonstrate its utility by applying it to an original dataset of statutes and regulations from two different countries, the United States and Germany, that spans more than twenty years (1998–2019). Our framework provides tools for assessing the size and connectivity of the legal system as viewed through the lens of specific document collections as well as for profiling individual legal documents over time. It goes beyond the existing literature, inter alia, by adapting network analytical methods to the peculiarities of legal documents, allowing the joint examination of multiple document types, and enabling temporal analysis. Implementing the framework for our dataset, we find that the United States legal system is increasingly dominated by regulations, whereas the German legal system remains governed by statutes, regardless of whether we measure the systems at the macro, the meso, or the micro level.

The remainder of the paper is structured as follows. In **Section 2**, we specify our network model of legal documents and detail how we instantiate it to analyze statutes and regulations in the United States and Germany. **Section 3** describes our methodological framework, and the results of applying this framework to our original dataset are presented in **Section 4**. We conclude by discussing the strengths and weaknesses of our approach in **Section 5**, where we also identify avenues for future research. Our exposition uses the basic terminology of graphs and networks; for textbook introductions, see [43–45].

2 DATA

In this section, we introduce our network model of legal documents (2.1) and instantiate it for our original dataset (2.2).

2.1 Data Model Specification

As visualized in **Figure 1A**, the legal system consists of multiple levels: the local (e.g., municipal) level, the intermediate (e.g., state or province) level, the national (e.g., federal) level, and the supranational (including international) level. Horizontally, it is usually subdivided into the legislative, executive, and judicial branches of government. These public parts are framed by the private sector, which operates on all levels, and the research community, which studies all parts of the legal system (including itself [46–50]). In all parts of the legal system, agents of varying sizes produce different types of outputs that create, modify, delete, apply, debate, or evaluate legal rules. These agents and their typical outputs are summarized in **Table 1**.

As the agents interact, they consciously interconnect their outputs. For example, court decisions regularly contain references to statutes, regulations, contracts, and other court decisions. **Figure 1B** gives an overview of the classic dependencies between the typical outputs of agents in the legal system. It illustrates that the *documented* part of the legal system constitutes a multilayered document network, which is changing over time as the agents continue producing or amending their outputs. Since the connections between the legal documents are placed deliberately by the agents, they encode valuable information about the content and the context of these documents. A lot of this information cannot be inferred from the documents' language alone (reliably or at all). Therefore, investigating the dynamic document network representation of a legal system using network analytical tools promises insights into its structure and evolution that would be hard or impossible to obtain via other methods.

To perform network analysis of a dynamic network of legal documents, we need to represent it as a series of graphs. Here, we build on a generalizable network model of statutory materials [27] and exploit the fact that the typical outputs listed in **Table 1** have three common features (beyond the obvious characteristic that they all contain *text*):

1. They are hierarchically structured (*hierarchy*).
2. Their text is placed in containers that are sequentially ordered and can be sequentially labeled (*sequence*).
3. Their text may contain explicit citations or cross-references (henceforth: references) to the text in other legal documents or in other parts of the same document (*reference*).

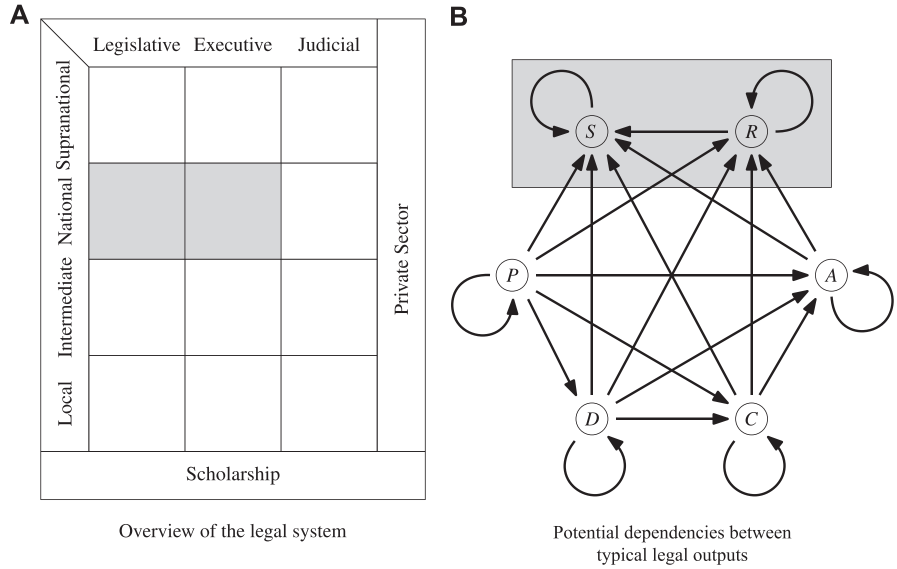


FIGURE 1 | Two-dimensional overview of the legal system **(left)** and dependencies between its typical outputs **(right)**, with the areas covered by the dataset introduced in this work highlighted in grey. The legal outputs in Panel **(B)** are (clockwise from the top left and in reverse topological order) Statutes, Regulations, Administrative Acts, Contracts, (Court) Decisions, and (Scholarly) Publications. The arrows illustrate typical dependencies between the document types, e.g., through explicit references from arrow tails to arrow heads (although in reality, most dependencies can be bidirectional, e.g., some courts also cite legal scholarship). Note that the sources covered by our dataset lie at the end of all typical dependency chains (i.e., statutes and regulations are last in the topological order of the dependency graph).

TABLE 1 | Overview of agents and outputs in the legal system.

	Branch of government			Private sector	Scholarship
	Legislative	**Executive**	**Judicial**		
Typical Large Agent	Parliament	Agency	Court	Firm	Institute
Typical Small Agent	Parliamentarian	Bureaucrat	Judge	Individual	Scholar
Typical Output(s)	Statute	Regulation Administrative Act	Decision	Contract	Publication

Therefore, each document at a given point in time (henceforth: *snapshot*) is represented as a (sub)graph, with its *hierarchy* modeled as a tree using *hierarchy edges*. We capture a document's *reference* using *reference edges* at the level corresponding to the document's *sequence*, which, inter alia, prevents the graph induced by these references from becoming too sparse (thereby eliminating some noise in the data and facilitating its analysis). The result is a directed multigraph, as illustrated in **Figure 2** for documents that are statutes or regulations (whose sequence level is the *section* level). Depending on the analytical focus, other edge types can be included (e.g., *authority edges* pointing from regulations to statutes can indicate which statutes delegated the rule-making power used to create which regulations), and depending on the document types considered, different types of edit operations are possible (e.g., court decisions and scholarly articles are only seldom changed after their initial publication), but the general model applies to all outputs listed in **Table 1**.

2.2 Data Model Instantiation

To illustrate the power of the methodological framework laid out in **Section 3** and produce the results presented in **Section 4**, we instantiate the data model described in **Section 2.1** with the outputs of the legislative and executive branches of government at the national level in the United States and Germany, i.e., federal statutes and regulations, over the 22 years from 1998 to 2019 (inclusive). The rules contained in these sources are universally binding and directly enforceable through public authority (the combination of which distinguishes them from the other outputs listed in **Table 1**). In the United States and Germany, they are arranged into edited collections: the United States Code (USC) and the Code of Federal Regulations (CFR) in the United States, and the federal statutes and federal regulations (which have no

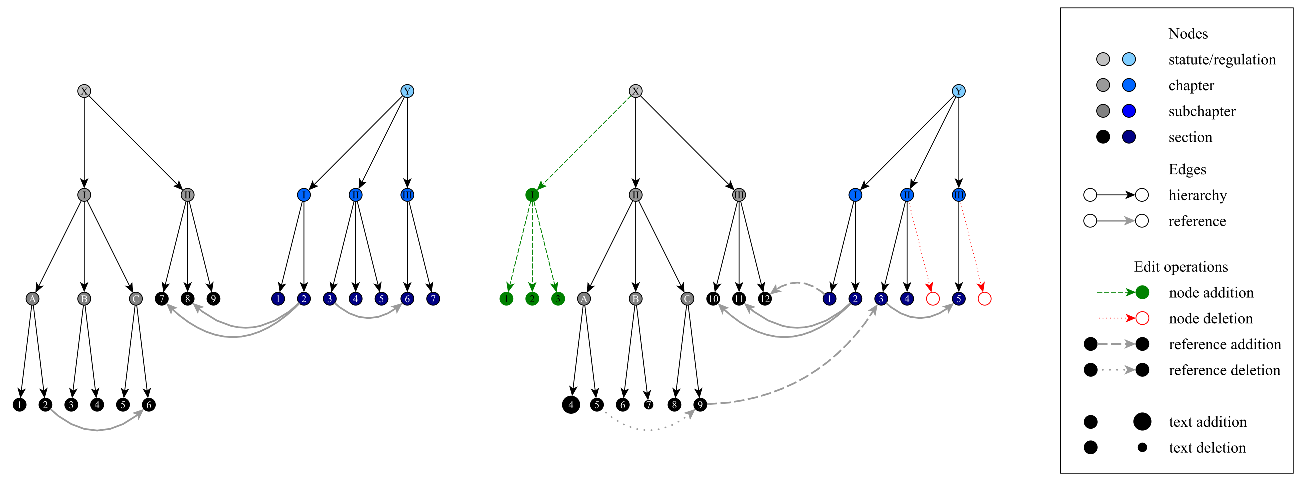


FIGURE 2 | Our legal network data model (adapted from [27]), illustrated for a dummy graph containing one statute and one regulation: initial configuration **(left)** and potential edit operations **(right)**.

special name) in Germany. These collections are actively maintained to reflect the latest consolidated state of the law (though, in the United States, the consolidation may lag several years behind the actual law). As such, they are a best-effort representation of all universally binding and directly enforceable rules at the federal level in their country at any point in time, commonly referred to as *codified law*.

For the United States, we obtain the annual versions (reflecting the state of the codified law at the *end* of the respective year) of the United States Code (USC) and the Code of Federal Regulations (CFR) from the United States Office of the Law Revision Council[1] and the United States Government Publishing Office[2], respectively. For Germany, we create a parallel set of annual snapshots for all federal statutes and regulations in effect at the *end* of the year in question based on documents from Germany's primary legal data provider, *juris GmbH*.[3] These data sources are the most complete presently available, but they may still be incomplete. They also reflect choices made by and events affecting the agents in charge of their maintenance, e.g., varying rates and orders of updates, purposeful or unintentional omissions or modifications, and changes in the agents' composition (e.g., as a consequence of elections).

We perform several preprocessing steps on the raw input data, detailed in the **Supplementary Material**, to extract the hierarchy, sequence, and reference structure contained in each collection. The results are directed multigraphs, one per country and year, akin to those illustrated in **Figure 2**. These graphs contain all structural elements of the USC and the CFR (in the United States) or the federal statutes and regulations (in Germany) as nodes and all direct inclusion relationships (*hierarchy*) and atomic references (*reference*) as edges, where the references are resolved to the section level (*sequence*). Each graph represents the *codified law* of a particular country in a particular year, containing documents of two *document types* (statutes and regulations) at the federal level.

When modeling codified law as just described, we take a couple of design decisions that limit the scope of the results presented in **Section 4**. First, we focus on *codified* law, i.e., *law in books*, excluding other legal materials listed in **Table 1**, especially those representing *law in action* (in the sense of [51]), or even other representations of legislative materials such as the United States Statutes at Large or the German Federal Law Gazette (Bundesgesetzblatt). These materials all merit investigation, and they need to be included in an all-encompassing assessment of the legal system. Our current work also serves as a preparatory step toward realizing this larger vision.

Second, we extract *atomic* and *explicit* references that follow a specified set of common patterns only, i.e., references including—in a typical format—a particular section (called "Paragraph" or "Artikel" in German law), a list of sections, or a range of sections. With this procedure, we exclude *container* references (e.g., references to an entire chapter of the USC), *pinpoint* references (e.g., references to a codified Act of the United States Congress by its popular name), *implicit* references (e.g., the use of a certain term implying its definition), and *explicit* references following *uncommon* patterns. As sketched in Section 3.3 of the **Supplementary Material**, there are plenty of such references, especially in the CFR, and including them would produce results different from those presented in **Section 4**. However, the graph representation of such references is inherently ambiguous, and their extraction is inherently more challenging than the extraction of atomic citations. Solving these problems falls outside the scope of this paper but presents an interesting opportunity for future work.

[1] https://uscode.house.gov/download/annualhistoricalarchives/downloadxhtml.shtml

[2] https://www.govinfo.gov/bulkdata/CFR

[3] This differs from the approach taken in [27], where the annual snapshots represented the law in effect at the *beginning* of the year in question.

Third, we resolve the atomic references we extract to the level of sections, rather than the smallest referenced unit (which might be a subsection or even an item in an enumeration), thereby effectively discarding potentially valuable information. Since for statutes and regulations, the section level corresponds to the documents' sequence level, this is consistent with our data model. It also reflects a focus on the perspective of the user, who tends to navigate the law on the section level because it is the only level at which the individual German laws or their United States counterparts, the chapters of the USC and the CFR, are uniquely sequentially labeled. Finally, it ensures a certain degree of comparability because sections are the only structural elements in which text is (with very few exceptions) guaranteed to be contained (albeit the amount of text varies widely across sections). Therefore, resolving references to the section level is reasonable for our purposes, but further research is needed on how the choice of the resolution level impacts the analysis of legal networks.

3 METHODS

Since the legal system produces a diverse set of outputs, many of which are rich in internal structure, a methodological framework for its dynamic network analysis must allow for many different foci and units of analysis. Our methods are designed to match this need, enabling us to describe the legal system and its evolution in its entirety (*macro level*), through selected sets of legal documents (*meso level*), or using individual documents and their substructures (*micro level*), all while integrating documents of potentially different types. More precisely, we provide tools for measuring

1. the *growth* of the legal system (3.1),
2. the macro-level, meso-level, and micro-level *connectivity* of the legal system (3.2), and
3. the *evolution* of individual units of law (e.g., single statutes, single regulations, or their sections or chapters) in the legal system (3.3).

While most network analytical tools we employ are conceptually simple, the challenge lies in selecting adequate units of analysis and metrics allowing for substantive interpretation. In the following, we refer to the object of study as *the legal system* for brevity, but one should keep in mind that this system is explored through the window provided by the document collection underlying the analysis (cf. **Figure 1**), and therefore can also refer to a *national legal system*.

3.1 Growth

As detailed in **Section 2.1**, despite their diversity, the typical outputs of the legal system have *hierarchy*, *sequence*, and *reference* as common structural features, and they also all contain *text*. Therefore, to assess the growth of the legal system, we track the number of tokens (roughly corresponding to words), the number of structural elements (i.e., hierarchical structures), and the number of references between documents of the same type at the documents' sequence level (e.g., the section level for statutes and regulations) across all documents in the collection, separately for each document type and across all temporal snapshots (e.g., all years). For the token counts, we concatenate the text of all materials for one snapshot and document type and split on whitespace characters. For a given document type, the structural element counts reflect the number of nodes of that type, and the reference counts reflect the number of reference edges between nodes of that type in our graphs. These measures give a first, high-level idea of how the legal system evolves, and the aggregation by document type allows us to uncover, e.g., differences in growth rates.

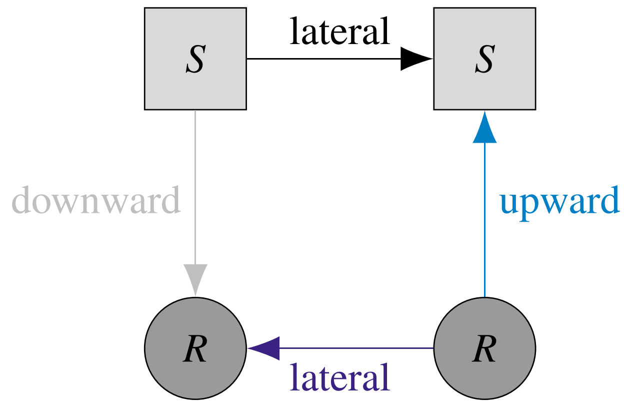


FIGURE 3 | Differentiation of reference types for a document collection containing two types of documents: statutes and regulations. Light grey squares marked *S* represent sections of statutes, dark grey circles marked *R* represent sections of regulations. Given that statutes stand above regulations in the hierarchy of legal rules, a reference from a statute to a regulation is *downward* (silver), a reference from a regulation to a statute is *upward* (light blue), and references between documents of the same type are *lateral* (black and blue).

To explore the relevance of references between documents of different types, if the document collection contains x types of documents, we distinguish between x^2 types of references. **Figure 3** illustrates the idea for $x = 2$ with statutes and regulations as document types. We count the number of references of each type for each temporal snapshot.

Raw reference counts do not show how the incoming and outgoing references are distributed across the individual sequence-level items. The user experience of a legal system, however, depends crucially on these distributions: If the typical item on the sequence level has very few outgoing references, the *expected* cost of navigating the law is much smaller than if the outgoing references are uniformly distributed, and if the distribution of incoming references is very skewed, when reading the law, we are much more likely to encounter a few prominent sequence-level items than a large number of less prominent ones (of course, the *actual* cost of the user also depends on the size of the items to be navigated, which can vary widely, e.g., among the sections in the USC). Therefore, we inspect the evolution of the in-degree distribution and the out-degree distribution of the subgraph induced by the reference edges. We compute these distributions separately for each combination of reference edge types (e.g., considering any combination of the reference types depicted in **Figure 3** for a document collection containing statutes and regulations) and across all snapshots. This allows us to evaluate whether growth in the number of references further amplifies differences in the prominence of different parts of the law, which would be reflected in a lengthening or thickening of the distributions' tails, and to assess how this affects the navigability of a legal system.

3.2 Connectivity

When exploring the connectivity of the legal system over time, we distinguish between macro-level connectivity (3.2.1), meso-level connectivity (3.2.2), and micro-level connectivity (3.2.3).

3.2.1 Macro-Level Connectivity

Investigating connectivity at the macro level helps us understand how information in the legal system is organized and processed. As the basis of all analyses, we consider the graph induced by the structural items on the documents' sequence level (referred to as *seqitems* in [27]) as nodes and the references between them as edges. For each snapshot, we count the number of non-trivial connected components (i.e., components with more than one node). Furthermore, we compute the fraction of nodes in the largest (weakly) connected component, the fraction of nodes in satellites, i.e., non-trivial components that are not the largest connected component, and the fraction of isolated nodes. We do this for the graph containing nodes of all document types as well as for the graphs containing only nodes of a single document type. These statistics provide a high-level overview of the system's information infrastructure and how it changes over time, and they enable a differentiated assessment of the role of documents of different types.

For a more detailed picture, we draw on concepts introduced in the study of the Web graph [52], which have also proven useful in the analysis of complex and self-organizing systems, e.g., in biology [53–55]. More precisely, we analyze the largest connected component of each of the sequence-level reference graphs, tracking the fraction of nodes contained in its strongly connected component, its in-only component (i.e., the nodes which can reach *to* but cannot be reached *from* the strongly connected component), its out-only component (i.e., the nodes which can be reached *from* but cannot reach *to* the strongly connected component), and its tendrils and tubes (whatever remains), again across all snapshots. We ask to what extent the legal system has a bowtie structure (i.e., a small strongly connected core joined by larger in-only and out-only components), which has been associated with "effective trade-offs among efficiency, robustness and evolvability" [54], inter alia, in complex biological systems, and whether any empirical deviations from that structure are characteristic of legal information processing.

3.2.2 Meso-Level Connectivity

One fundamental question concerning a legal system's connectivity at the meso level is how it self-organizes into areas of law. Existing taxonomies categorizing the law into distinct fields are largely based on tradition (e.g., the titles of the USC) or intuition (e.g., the thematic categories used by some legal database providers). Exploiting the connectivity provided by references between legal documents at the meso level, network analytical methods provide an alternative, data-driven approach to mapping the law. To implement such an approach, we follow a multi-step procedure:

1. We preprocess the graphs for each snapshot by taking the quotient graph at the granularity we are interested in (e.g., at the level of individual chapters for an analysis of the USC and the CFR). That is, we remove all nodes above and below that level and reroute all references outgoing from or incoming to a lower-level node to the node's unique ancestor that lies on the level of interest.
2. We cluster each of the *undirected* versions of the graphs from Step 1 separately using the *Infomap* algorithm [56, 57] with a parametrization that mirrors domain knowledge, and passes sensitivity and robustness checks. Leveraging the randomness inherent in this algorithm, we increase the robustness of the clustering for each graph by computing the *consensus clustering* [58] of 1000 *Infomap* runs with different seeds, where two nodes are put into the same cluster if they are in the same cluster in 95% of all runs. We choose the *Infomap* algorithm as our clustering algorithm because it is scalable, has a solid information-theoretic foundation, and mirrors the process in which users like lawyers navigate law (inter alia, by identifying a relevant section of a statute, reading that section, then potentially following a reference).
3. We compute pairwise alignments between the clusterings of all temporally adjacent snapshots based on the nodes of the *unpreprocessed* graphs that wrap text (for details on our alignment procedure, see Section 4.3.1 in the **Supplementary Material**). This is most relevant for collections containing documents that can change over time (e.g., statutes and regulations), and it allows us to assess, inter alia, what amount of text from a cluster A in year y is contained in a cluster B in year $y + 1$.
4. We use the clusterings from Step 2 and the alignments from Step 3 to define a *cluster family graph* as introduced in [27]. This graph contains all clusters from all snapshots as nodes, and two clusters A and B are connected by a (weighted) edge if A lies in snapshot y, B lies in snapshot $y + 1$, at least p% of the tokens from A are contained in B, and at least p% of the tokens from B are contained in A, where p is chosen based on the analytical resolution we are interested in.
5. We define a *cluster family* as a connected component in a cluster family graph from Step 4 and compute, for each cluster family in each year, the number of tokens it contains from each document type.

This process is a variant of the family graph construction developed for statutes in [27], with the modification that we now allow for input data containing documents of different types. It results in a dynamic, data-driven map of the legal system that accounts for the information provided by the references between its documents.

3.2.3 Micro-Level Connectivity

At the micro level, the connectivity created by references allows us to assess the roles of individual units of law. Regardless of the level at which we aggregate the references between documents, the shapes of the nodes' neighborhoods at that level contain valuable information about their function in the legal system. This information is partly accounted for in the meso-level connectivity assessment, which

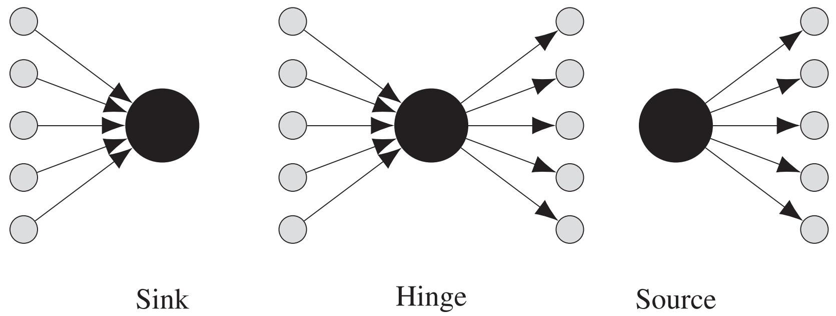


FIGURE 4 | Star types. If the hub's in-degree is at least ten times its out-degree, the star is a *sink*. If the hub's out-degree is at least ten times its in-degree, the star is a *source*. Otherwise, the star is a *hinge*.

leverages local *density*. While local density can help us find out which nodes interact strongly, local *sparsity* lets us identify nodes that play a particularly prominent role for the information flow in the network: If a node's neighbors are themselves only very sparsely connected (i.e., its neighbors almost form an independent set), the node provides an important *bridge* between them. We call the ego graph of such a node a *star*, with the node as its *hub* and the node's neighbors as the *spokes*.

In directed graphs, we can classify stars according to the ratio between the hub's in-degree and the hub's out-degree as depicted in **Figure 4**. More precisely, we define the type of a star s with hub v as follows:

$$\text{type}(s) := \begin{cases} \dfrac{\delta^-(v)}{\delta^+(v)} \geq 10 & \text{sink} \\ \dfrac{1}{10} < \dfrac{\delta^-(v)}{\delta^+(v)} < 10 & \text{hinge} \\ \dfrac{\delta^+(v)}{\delta^-(v)} \geq 10 & \text{source,} \end{cases}$$

where $\delta^+(v)$ is v's out-degree and $\delta^-(v)$ is v's in-degree. In the legal system, the type of a star captures the hub's role in mediating the information flow in its neighborhood.

To identify and classify stars in the legal system at the documents' sequence level, for each snapshot, we create the ego graph for each node v in the graph induced by the reference edges (where we exclude parallel edges). We then iteratively remove the node w that is connected to most of v's neighbors while w is connected to more than 5% of v's neighborhood (excluding w), and keep the ego graph if it has a certain minimum size determined by the size of the collection (e.g., 10 nodes for collections with several thousands of items on the sequence level). The stars produced in this way contain no spoke that is connected to more than 5% of the other spokes, and we classify them to identify those sequence-level items that are vital to the information flow in their neighborhoods and to describe the type of mediation they perform. To find stars at levels above the documents' sequence level, we can apply the methodology just described on graphs that aggregate references at those levels (e.g., on the quotient graphs described in **Section 3.2.2**).

3.3 Profiles

To assess the evolution of individual units of law (e.g., individual court decisions or chapters of a regulation) in the legal system, we create profiles of these units covering all temporal snapshots in the document collection under study. More specifically, based on the quotient graphs that are created on the level of our unit of interest and contain only reference edges (like the preprocessed graphs described in Step 1, **Section 3.2.2**), we track ten statistics in five groups (note that not all of these statistics can change over time for units of all legal document types):

1. the number of tokens and the number of unique tokens,
2. the number of items above, on, and below the sequence level (provided our unit of analysis lies above the sequence level),
3. the number of self-loops,
4. the weighted in- and out-degree (accounting for parallel edges), and
5. the binary in- and out-degree (excluding parallel edges).

These statistics capture how the unit of law in focus evolves in size (number of tokens), topical breadth (number of unique tokens), structure (number of items above, on, and below the sequence level), self-referentiality (number of self-loops), scope of interdependence within the legal system (weighted in-degree and weighted out-degree), and diversity of interdependence within the legal system (binary in-degree and binary out-degree). Finally, by constructing the ego graphs of the profiled unit for its out-neighborhood and its in-neighborhood and following the evolution of these ego graphs across snapshots, we assess to what extent a profiled unit references which other units (*reliance*) and to what extent it is referenced by which other units (*responsibility*).

4 RESULTS

In the following, we apply the framework presented in **Section 3** to the data introduced in **Section 2.2**, i.e., codified law comprising federal statutes and regulations in the United States and Germany over the 22 years from 1998 to 2019 (inclusive). We start by examining the

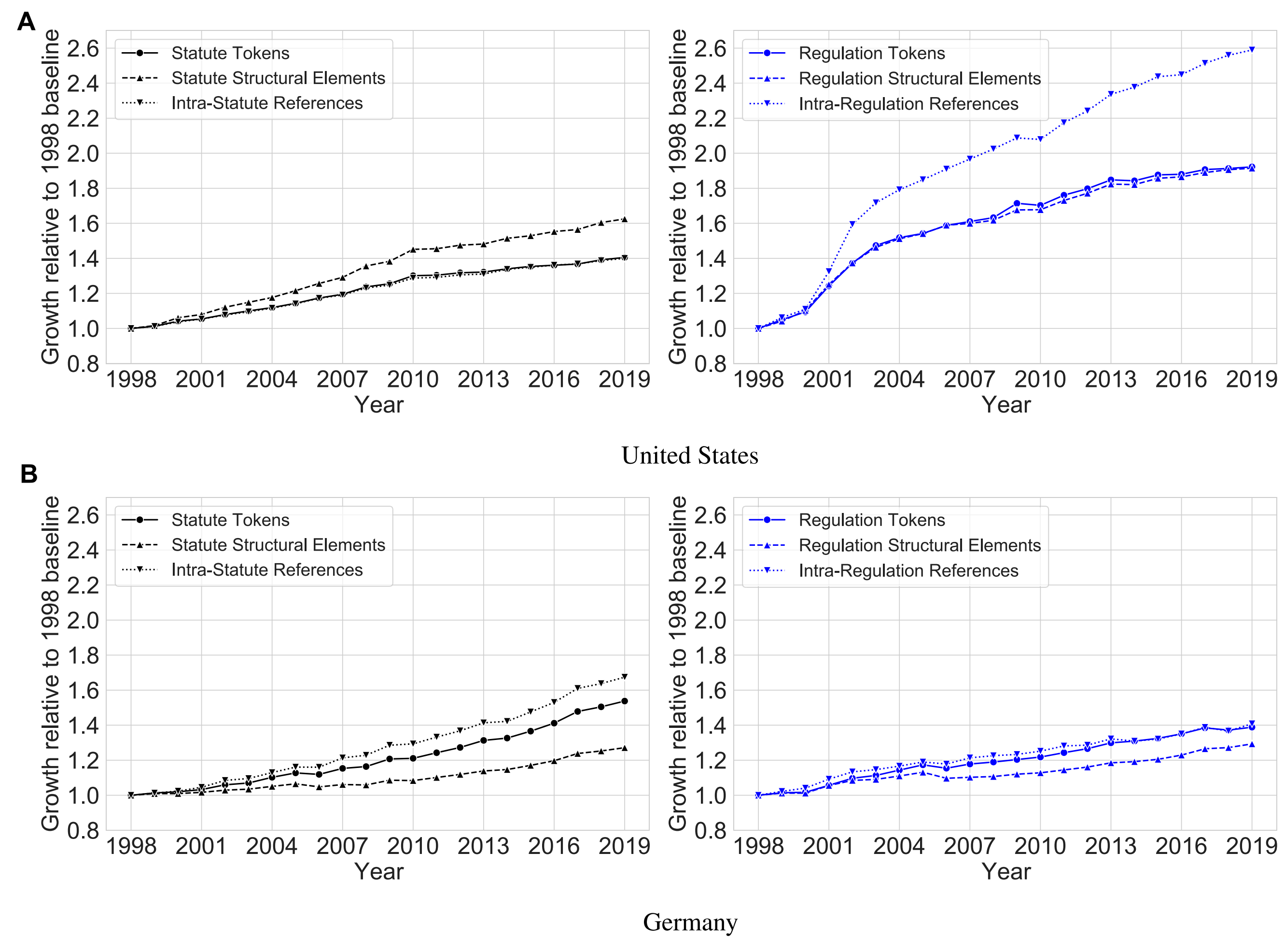


FIGURE 5 | Growth relative to the 1998 baseline for statutes **(left)** and regulations **(right)** in the United States **(top)** and Germany **(bottom)**.

growth of the United States and German national legal systems (henceforth: the national legal systems) as viewed through the lens of our data (**Section 4.1**). Next, we investigate the macro-level, meso-level, and micro-level connectivity of these national legal systems (**Section 4.2**). Finally, we explore the evolution of selected chapters of the USC and the CFR and selected German statutes and regulations within their national legal systems in a case study focusing on financial regulation (**Section 4.3**). The results we report are mostly descriptive, and as discussed in **Section 5**, identifying causal factors behind the dynamics we observe or interpreting our results using a qualitative approach is left to future research.

4.1 Growth

Figure 5 summarizes the growth of the United States legal system and the German legal system as measured by the tokens, structural elements, and lateral references contained in their codified law. Each row of the figure corresponds to a country, and each column corresponds to a document type. All counts are divided by their value in 1998, i.e., **Figure 5** depicts growth relative to the 1998 baseline. Supplementing the time series data,

TABLE 2 | (Rounded) size of the national legal systems of the United States (top) and Germany (bottom) as measured by the tokens, structural elements, and references in their codified law in 1998 and 2019, including the total percentage change between these years (Δ).

	Statutes			Regulations		
	1998	**2019**	**Δ**	**1998**	**2019**	**Δ**
			(a) United States			
Tokens	15.2 M	21.4 M	41	43.9 M	84.3 M	92
Structures	516.2 K	838.8 K	63	1.4 M	2.7 M	91
References	80.1 K	112.1 K	40	134.6 K	348.4 K	159
			(b) Germany			
Tokens	5.0 M	7.7 M	54	3.9 M	5.4 M	39
Structures	130.6 K	166.0 K	27	87.9 K	113.7 K	29
References	86.4 K	144.6 K	67	33.5 K	47.1 K	41

Table 2 provides the absolute counts for 1998 and 2019 and the total percentage change between these years (Δ).

Figure 5 and **Table 2** show that over the last two decades, the legal systems of both countries have grown substantially. In the

United States, the USC (containing codified statutes) has over 60 new structural elements (e.g., chapters, parts, or sections) in 2019 for every 100 such elements it had in 1998. Notably, as evident from the upper left panel of **Figure 5**, the growth rate of the USC appears to have experienced two distinct periods when measured by its structural elements: one period with a monotonic growth rate of approximately 4% per year (1998–2010), followed by another period with a decelerated monotonic growth rate of approximately 2% per year (2010–2019). At a slightly lower level, this trend also occurs for both the number of tokens and the number of intra-USC references. For example, there are approximately 40 new tokens or references in 2019 for every 100 tokens or references that existed in 1998. Shifting the focus for the United States to the CFR (containing codified regulations), as observable from the upper right panel of **Figure 5**, the quantity of regulations has increased by an even greater factor. For every 100 structural elements or tokens that were present in 1998, approximately 90 additional elements or tokens exist in 2019. This increase is even more extreme for intra-CFR references, where there are almost 160 *new* references in 2019 for every 100 that existed in 1998. Apart from brief intervals of stagnation or slight decrease (2009–2010, 2013–2014), these increases have been monotonic.

Corresponding trends for German statutes and regulations are presented in the bottom row of **Figure 5**. Growth in the German legal system has been qualitatively similar to that in the United States legal system but quantitatively less pronounced and of different functional shape. For both German statutes and German regulations, there are approximately 30 new structural elements in 2019 for every 100 that existed in 1998. Unlike in the United States, however, this growth has been non-monotonic: When measured through structural elements, both statutes and regulations experienced some periods of shallow decline between 2005 and 2010. These shrinking periods are generally not mirrored by the token and lateral reference counts, with one notable exception: In the period from 2005 to 2006, *all* German statistics decreased. This is likely due to statutes aiming to cleanse the law (*Rechtsbereinigungsgesetze*), eight of which were introduced in 2006 (recall that our 2006 snapshot represents the law at the *end* of 2006).[4] In total, there are approximately 55 new statute tokens in Germany in 2019 for every 100 such tokens that existed in 1998. Like *regulations* in the United States, German *statutes* experienced a greater increase in the quantity of lateral references than in other metrics: For every 100 references in 1998, there are approximately 70 new references in 2019. For German *regulations*, as for *statutes* in the United States, the rate of change has been more similar across metrics, with growth varying roughly between 30% and 40% (as noted in **Table 2**). At a high level, the growth of the German legal system thus seems to be driven by statutes, whereas the growth of the United States legal system appears to be driven by regulations.

Figure 5 and **Table 2** only account for *lateral* references, excluding references between documents of different types. Therefore, **Figure 6** shows growth relative to the 1998 baseline for lateral and *upward* (i.e., regulation-to-statute) references. We exclude *downward* references because they are very few in number (which means that even a small absolute increase results in a large relative increase) but note that, contrary to the legal theory intuition, they *do* occur.[5] As evident from **Figure 6**, upward references have grown at similar rates in both countries, with approximately 80 new upward references existing in 2019 for every 100 upward references that existed in 1998. This relative increase is larger than that of the lateral references in both countries, with the exception of lateral regulation references in the United States, whose growth rate dwarfs all others. Since the token and structural element growth rates of German regulations are lower than or similar to those of German statutes, this means that over the period under study, connectivity between statutes and regulations in Germany has grown faster than connectivity within statutes or within regulations.

To evaluate how the growth in the number of references affects the differences in the prominence of individual sections of codified law in the legal systems under study, in **Figure 7**, we examine the in-degree distribution and the out-degree distribution of the graphs induced by the reference edges in 1998 and 2019 (an analogous figure normalizing section degrees by section size in tokens can be found in Section 4.1 of the **Supplementary Material**). Since these distributions are highly skewed (as in many graphs arising from complex systems), we plot them on a log-log scale.

All distributions plotted in **Figure 7** demonstrate features common among graphs arising from bibliometric dynamics. For example, most sections of statutes and regulations in the United States and Germany are referenced very few times (if at all). The detailed characteristics of the distributions, however, differ between distribution types, document types, and countries: For the United States, the out-degree distributions exhibit less right skew than their in-degree counterparts, while in Germany, we observe the opposite: All out-degree distributions are either within an order of magnitude of or have a longer and thicker right tail than their in-degree counterparts. Similarly, the sections contained in United States regulations exhibit a higher degree

[4]These statutes are: (1) Erstes Gesetz über die Bereinigung von Bundesrecht im Zuständigkeitsbereich des Bundesministeriums des Innern vom 19. Februar 2006 (BGBl. I S. 334), (2) Gesetz zur Bereinigung des Bundesrechts im Zuständigkeitsbereich des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz vom 13. April 2006 (BGBl. I S. 855), (3) Erstes Gesetz über die Bereinigung von Bundesrecht im Zuständigkeitsbereich des Bundesministeriums der Justiz vom 19. April 2006 (BGBl. I S. 866), (4) Erstes Gesetz zur Bereinigung des Bundesrechts im Zuständigkeitsbereich des Bundesministeriums für Wirtschaft und Technologie und im Zuständigkeitsbereich des Bundesministeriums für Arbeit und Soziales vom 19. April 2006 (BGBl. I S. 894), (5) Gesetz zur Änderung und Bereinigung des Lastenausgleichsrechts vom 21. Juni 2006 (BGBl. I S. 1323), (6) Gesetz über die Bereinigung von Bundesrecht im Zuständigkeitsbereich des Bundesministeriums für Arbeit und Soziales und des Bundesministeriums für Gesundheit vom 14. August 2006 (BGBl. I S. 1869), (7) Erstes Gesetz über die Bereinigung von Bundesrecht im Zuständigkeitsbereich des Bundesministeriums für Verkehr, Bau und Stadtentwicklung vom 19. September 2006 (BGBl. I S. 2146), and (8) Zweites Gesetz über die Bereinigung von Bundesrecht im Zuständigkeitsbereich des Bundesministeriums des Innern vom 2. Dezember 2006 (BGBl. I S. 2674).

[5]The total number of downward references in the United States increases from 24 in 1998 to 90 in 2019. In Germany, it rises from 305 to 833.

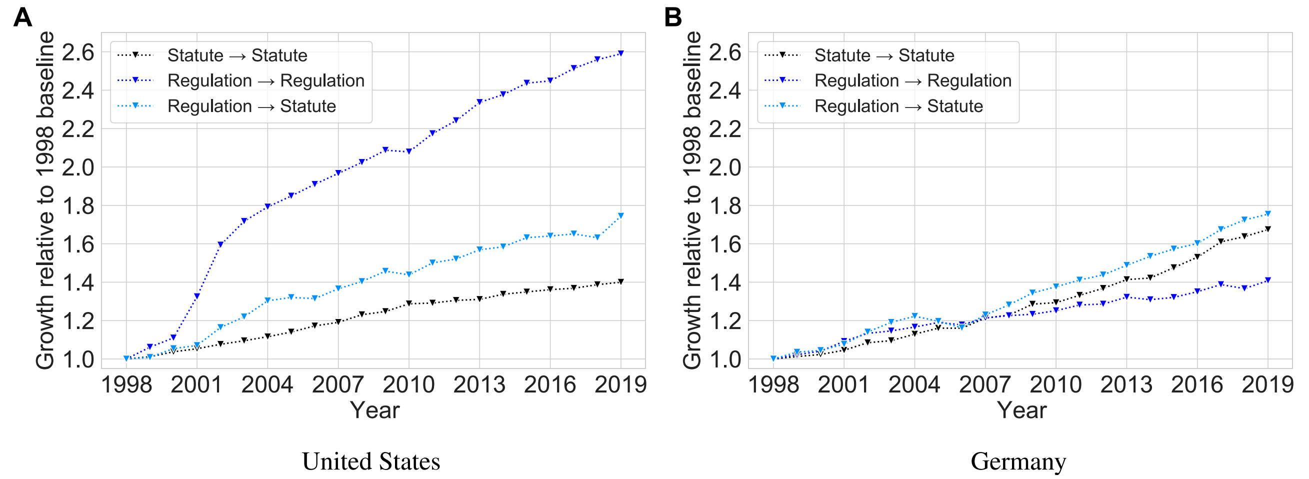


FIGURE 6 | Growth relative to the 1998 baseline for lateral and upward references in the United States **(left)** and Germany **(right)**.

of skew in their in-degree distributions than the sections contained in United States statutes (e.g., a higher fraction of these sections has more than 500 ingoing references), but in Germany, the opposite phenomenon occurs at a lower absolute level: There are many statute sections with more than 100 ingoing references but hardly any regulation sections clearing that threshold. These national divergences might be partly due to the differing ratio between statutes and regulations, but they could also point to peculiarities in United States and German drafting style.

Comparing *all* distributions across countries, we observe that the United States legal system exhibits more extreme statistics than the German legal system (which might, at least in part, be due to its larger size). Finally, we see that from 1998 to 2019, most distributions shift to the right, i.e., the tails become both longer and thicker, which is in line with bibliometric preferential attachment dynamics [59, 60]. This indicates that reference growth has disparate impact, amplifying the differences in relevance between the individual sections contained in United States and German statutes and regulations. As a consequence, the difficulty of navigating the law increases more slowly than the growth of the reference count may suggest.

4.2 Connectivity

When exploring the connectivity of the national legal systems of the United States and Germany over time, we distinguish between the macro level, the meso level, and the micro level as suggested in **Section 3.2**.

4.2.1 Macro-Level Connectivity

To understand how the United States legal system and its German counterpart organize and process information, we investigate the connectivity of the graphs containing code sections as nodes and references between them as edges. **Figure 8** displays the number of non-trivial (weakly) connected components (i.e., components with more than one node) in these graphs over time, in absolute terms and per 1000 tokens. It shows that the connectivity in the graphs containing only statute sections is generally higher than that in the graphs containing only regulation sections or sections of both document types. Furthermore, while the United States system seems more fragmented than the German system (**Figure 8A**) when considering absolute numbers, it turns out to be relatively less fragmented when normalizing for system size (**Figure 8B**).

For a more granular connectivity assessment over time, **Figure 9** shows, for each year from 1998 to 2019, what fraction of statute sections and regulation sections in each country is contained in the largest connected component, satellite components, or isolates, and how the largest connected component is composed internally. The underlying graphs do not distinguish between sections of different document types; analogous figures considering statute sections only and regulation sections only can be found in **Section 4.2** of the **Supplementary Material**. In both the United States and Germany, the largest connected component is growing as the fraction of sections contained in both satellites and singletons decreases, and the difference between the largest connected component fraction in 1998 and that in 2019 is around 10%. However, the relative size of these largest connected components varies substantially between the two countries: In the United States, the largest connected component has grown from about 40% to nearly 50%, while in Germany, its size has increased from circa 55% to roughly 65%. Furthermore, the fraction of isolates (sections that neither reference another section nor get referenced by another section) is larger in the United States (around 45% in 2019) than in Germany (below 30% in 2019).

When focusing on the largest connected component and taking edge directions into account, the differences between the two countries become even more pronounced. In the United States, the fraction of the largest connected component contained in the in-only component is almost equal to that contained in its tendrils and tubes in 1998, and both lie around 45%. Over time, these fractions diverge as the strongly connected component and the out-only component grow and the in-only component stagnates.

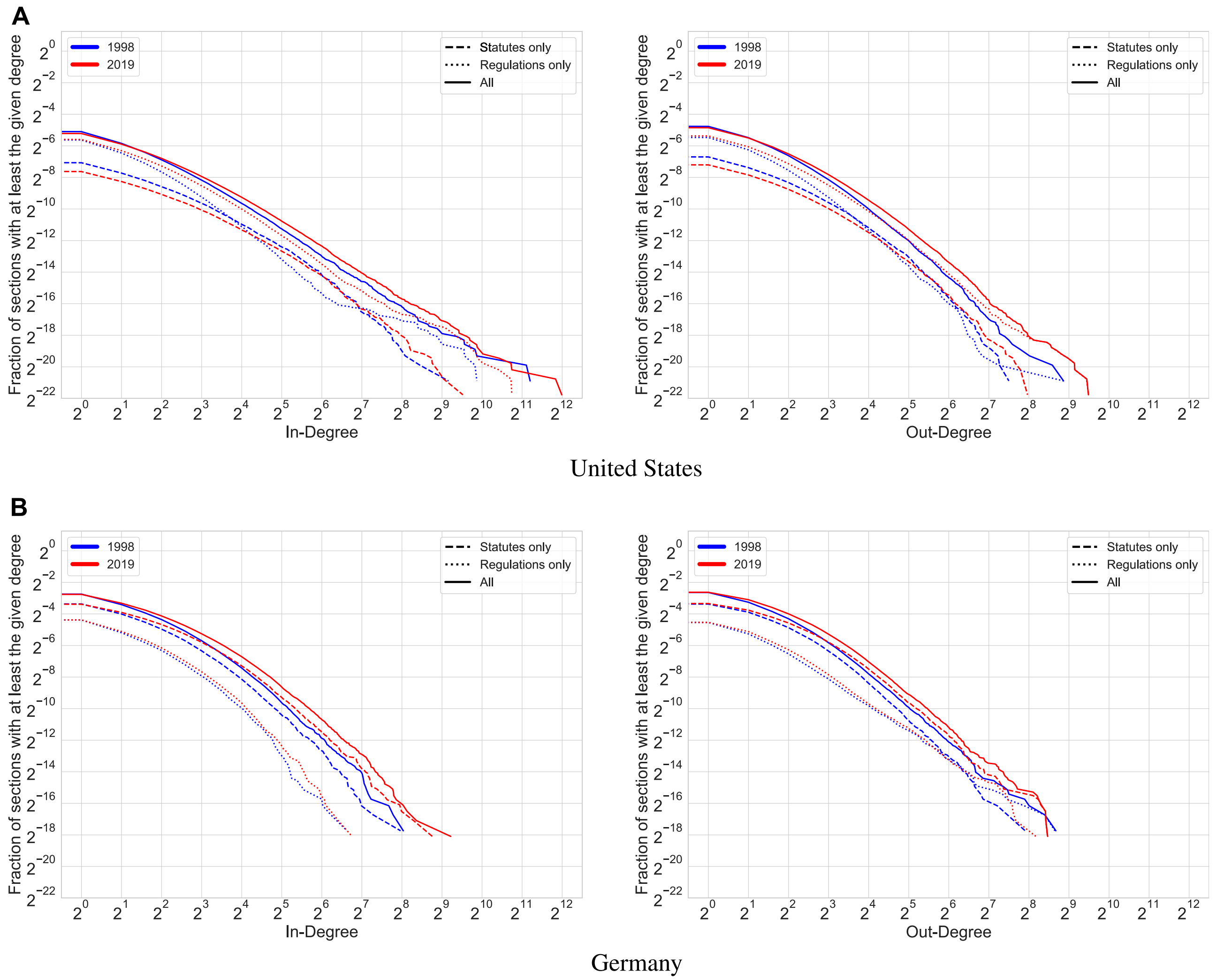
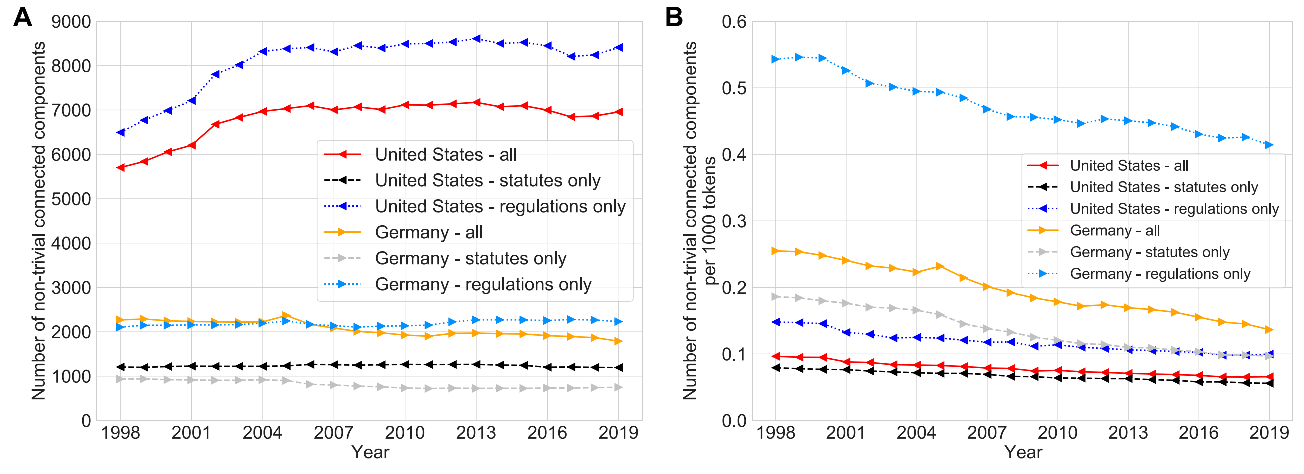


FIGURE 7 | In-degree **(left)** and out-degree **(right)** distributions for the United States **(top)** and Germany **(bottom)** in 1998 (blue) and 2019 (red) when considering statutes only (dashed line), regulations only (dotted line), or statutes and regulations (solid line).

In Germany, however, tendrils and tubes dominate in 1998, accounting for more than 65% of nodes, but by 2019, their fraction has declined to less than 45%, while the strongly connected component and the out-only component have grown at low levels and the in-only component has gained more than 50% in fractional size (growing from less than 20% to over 30%).

Notably, in both legal systems, the out-only component accounts for the smallest fraction of sections in 2019 (around 7% in the United States and around 12% in Germany), followed by the strongly connected component, with none of them containing more than 15% of all sections, while the in-component is twice as large in Germany and thrice as large in the United States. Hence, at least when considering code sections as nodes and references between them as edges, both national legal systems do not exhibit the bowtie structure observed in biological systems (small strongly connected component with larger in-only and out-only components [53]) or that found in early measurements of the World Wide Web (all components, including tendrils and tubes, of roughly the same size, with a slightly larger strongly connected component [52]). Rather, the legal systems we study are shaped more like rockets, with the in-only component as their base, tendrils and tubes as their fins, the strongly connected component as their body, and the out-component as their nose cone (see **Figure 10** for an illustration). The rocket structure mirrors both the hierarchical structure of legal systems (large in-only component, small out-only component) and the fact that some areas of law function relatively independently (many tendrils and tubes; also evident from the nontrivial fraction of nodes outside the largest connected component). This suggests that it might be characteristic of legal systems in general, but more research is needed to corroborate this hypothesis. Similarly, it would be interesting to investigate how our observations change if we include, e.g., non-atomic references (which, by definition, interconnect multiple sections).



Non-trivial connected components total

Non-trivial connected components per 1000 tokens

FIGURE 8 | Development of reference connectivity as measured by the absolute number of non-trivial (weakly) connected components **(left)** and the number of non-trivial (weakly) connected components per 1,000 tokens **(right)** in the graphs induced by reference edges between all sections (solid lines), statute sections only (dashed lines), and regulation sections only (dotted lines) in the United States (left-pointing triangles) and Germany (right-pointing triangles).

4.2.2 Meso-Level Connectivity

When analyzing the connectivity of the United States and German legal systems at the meso level, our goal is to create a dynamic, data-driven map of their codified law. To this end, for both the United States and Germany, we compute cluster families as described in **Section 3.2.2**, using quotient graphs on the chapter level in the United States and on the statute or regulation level (or the book level, if available) in Germany. Here, we choose 100 as the preferred number of *Infomap* clusters and 15% of tokens as the edge threshold for constructing the cluster family graph (for details on how we handle text that does not lie in a chapter as well as a sensitivity analysis of the parameter choice, see Sections 4.3.1 and 4.3.4 of the **Supplementary Material**). We calculate how many tokens from statutes and regulations these families contain in each year from 1998 to 2019. By construction, our cluster families unite sets of related rules that can be thought of as different areas of law, where—unlike in, e.g., the title structure of the USC or the German finding aids' subject classification (Fundstellennachweise, FNA)—the categorization is based solely on the empirically observed reference relationships between the legal documents in our data.

Figure 11 shows the evolution (1998–2019) of the ten cluster families with the largest number of tokens in 2019 (henceforth: top ten cluster families) for each country, which we label leveraging our subject matter expertise (details on the labeling procedure and complementary linguistic statistics can be found in Section 4.3.2 of the **Supplementary Material**). Most families either represent a traditional field of law (e.g., property law or financial regulation) or concern a real-life domain (e.g., energy or vocational training). A few families hold clusters from more diverse backgrounds and are therefore hard to interpret at first sight (e.g., a family containing military, public finance, and research regulation in the United States or a family containing court procedure, data security, and telecommunications in Germany). However, a more detailed examination of the individual clusters constituting these families uncovers nuanced underlying topics (e.g., grants and commercial activity by the federal government in the example from the United States, and data protection in public [including court] proceedings in the example from Germany). Hence, in summary, the method sketched in **Section 3.2.2** produces an informative map of the codified law for both countries we investigate (at the resolution level determined by our parametrization).

Inspecting the panels in **Figure 11** in more detail, we observe that the families' ratios of statute tokens to regulation tokens span the whole possible range: Some families are *statute-heavy* (i.e., contain mostly statute tokens), others are *regulation-heavy* (i.e., contain mostly regulation tokens), and yet others are *mixed* (i.e., lie between the aforementioned extremes). For a robust categorization of the ten largest families, the data suggests a threshold of an average 80% (i.e., an average ratio of 4:1) over the entire investigation period to classify a family as *x-heavy* for $x \in \{\text{statute}, \text{regulation}\}$. In the United States, this leads to four mixed families (Agriculture and Food; Financial Regulation; Energy; Housing) and six regulation-heavy families. In Germany, we find four statute-heavy families (Courts and Data Protection; Criminal Law and Justice; Corporate Taxes; Property), one regulation-heavy family (Vocational Training) and five mixed families. The overall situation remains similar even if we adopt a simple majority for the classification (eliminating the *mixed* category): With the exception of three singular years in two families (Energy in 1998, and Housing in 1999 and 2000), all top families in the United States contain a majority of regulation tokens. The German data then presents three regulation-heavy families (Vocational Training; Environmental and Workplace Safety; Environmental Protection) and seven statute-heavy families. A particularly striking example of a statute-heavy family is the Property family in Germany, in which there are nine times as many statute tokens as there are regulation tokens in all years except

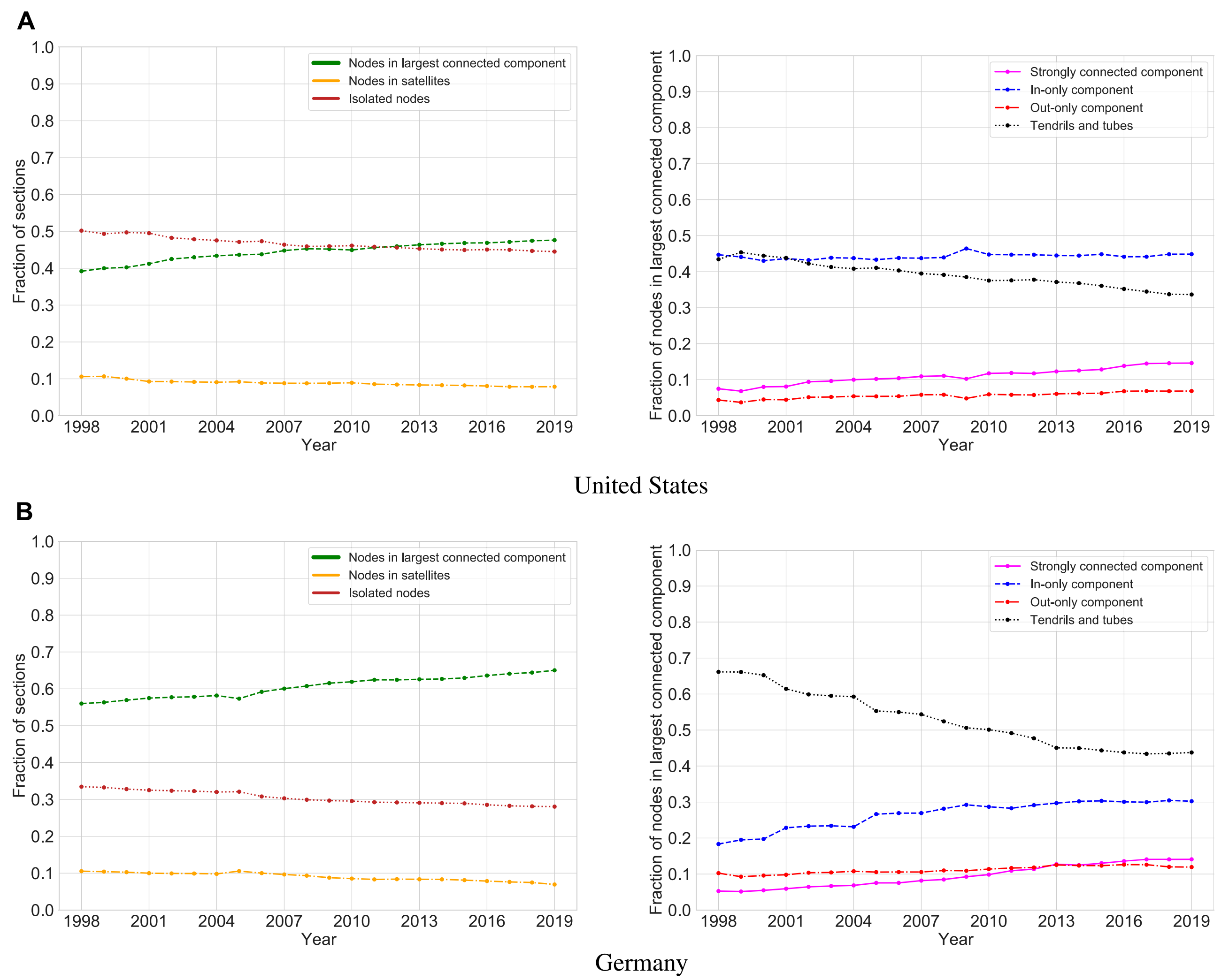


FIGURE 9 | Development of reference connectivity in the United States **(top)** and Germany **(bottom)** as measured by the fraction of sections contained in the largest connected component **(left)**, along with the internal structure of that component **(right)**.

between 2002 and 2006. The United States cluster family concerning Healthcare and Tax (two topics connected, inter alia, via the tax-based funding of Medicare and Medicaid) represents the opposite extreme, containing almost no statute tokens over the entire period under study. Interestingly, no family in either country is constantly balanced between statutes and regulations, with the family concerning Personal and Consumption Taxes in Germany coming closest in the period from 1998 to 2015.

As **Figure 11** traces the development of the top ten cluster families in each country over time, we can also observe changes in the families' composition. Extending the terminology adopted above, we can classify the families' growth based on the fraction of growth that is attributable to each of our document types. We say that a family is *x-driven* for $x \in$ {statute, regulation} if tokens from x account for at least 80% of the family's net growth when comparing 1998 and 2019, otherwise, we say that its growth is *mixed*. Using these categories, we can classify all of the United States top ten families as regulation-driven and half of the German top ten families as statute-driven (Social Security; Personal and Consumption Taxes; Criminal Law and Justice; Corporate Taxes; Property), while only one German family is classified as regulation-driven (Vocational Training). The full categorization of all top ten families for both countries, both in terms of their average composition and in terms of their growth, can be found in Section 4.3.3 of the **Supplementary Material**, where we further show that the general tendencies described above also hold for the entire population (although the trends are neither monotone nor universal and their extent differs from family to family).

Overall, the dynamics of the largest cluster families reflect the growth patterns documented in **Figure 5**. In absolute terms, regulations outgrow statutes by large margins in all of the top ten United States families, and statutes moderately outgrow regulations in most of the top ten German families. In relative terms, regulations still dominate in the United States, and statutes

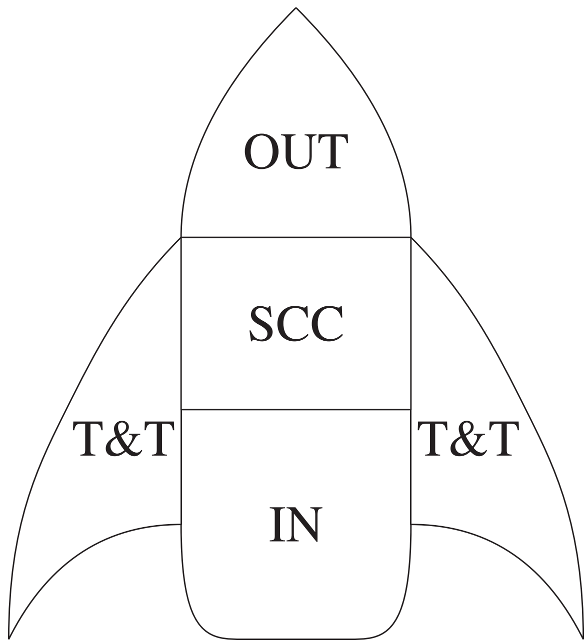


FIGURE 10 | Rocket structure of a legal system when viewed through the lens of macro-level connectivity, with the in-component (IN) as the rocket's base, the strongly connected component (SCC) as the body, the out-component (OUT) as the nose cone, and tendrils and tubes (T&T) as the fins.

still dominate in Germany (although they are less prominent than they appear when considering absolute numbers). In summary, based on the top ten families depicted in **Figure 11**, the United States seems to favor rule making via regulations, while Germany seems to favor rule making via statutes, and both countries' preferences appear to get stronger over time.

Finally, to evaluate how federal regulations impact our data-driven map of the United States and German legal systems, we compare the cluster families depicted in **Figure 11** to those derived in prior work that considers only federal statutes [27]. For the United States, the top ten cluster families based on statutes only have topics similar to those derived from statutes and regulations combined, including Environmental and Health Protection, Public Health and Social Welfare, Taxes, Agriculture and Food, Financial Regulation, Public Procurement, Telecommunications, and Federal Grants and Commercial Activity including Small Business Aid. The topic of Education makes the top ten in the statutes-only data but not in the data containing regulations, while Maritime Affairs and Transport as well as Energy only rise to prominence in the combined data. In Germany, topics such as Financial Regulation, Taxes, Social Security, Environmental Protection, Criminal Law and Justice, and Property represent sizeable cluster families based on both datasets. The topic of Public Servants, Judges, and Soldiers features prominently only in the results excluding regulations, while the families of Vocational Training and of Environmental and Workplace Safety make the top ten only in the combined data.

First and foremost, however, comparing our results to those from [27] demonstrates that adding federal regulations to the data results in a more accurate map of law. For example, the German data from [27] features a family on Market and Network Regulation that includes a leading cluster on (renewable) energy law, while no comparable family exists in the United States. Having added federal regulations, we now see such a family in the top ten also in the United States, whose prominent position is explained by its mixed composition (including more than 70% regulation tokens on average). At the same time, a cluster concerning (renewable) energy law is now part of the Environmental and Workplace Safety family in Germany because its regulations connect it more closely to rules concerning the protection of the environment than its statutes alone. This suggests that adding yet further document types, e.g., federal court decisions, to our data will continue to improve the legal maps produced using our methodology, making this a promising avenue for further research.

4.2.3 Micro-Level Connectivity

We analyze the connectivity of the United States and German legal systems on the micro level in order to identify those code sections that play a particularly important role in mediating the information flow between the sections which they reference and the sections by which they are referenced. More precisely, we apply the method sketched in **Section 3.2** to the graphs induced by the reference edges, where we keep a star if it has at least ten nodes in total. We classify these stars (and their hubs) into sinks, hinges, and sources depending on the ratio between their hubs' in-degree and their hubs' out-degree, and hypothesize that hubs of the same type have a similar function within the legal system. We explore the merits of this hypothesis by identifying and classifying the stars of each type in 1998 and 2019 and analyzing the content of the top five stars (i.e., those with the largest number of nodes) of each type in 2019 as shown in **Table 3**.

In the United States, we find that hubs of the same type indeed play similar roles in the legal system: *Sinks* contain delegation of authority and general procedures, e.g., for record keeping, that are relevant for and therefore referenced by many other sections. *Hinges* connect entire collections, enumerations, and definitions to one another. 49 CFR § 171.7 is an example, as its only function is the incorporation of material collections by external parties (such as the American National Standards Institute) into other sections of the CFR like 49 CFR § 173.306, which itself serves as a hub for other sections. *Sources* enumerate duties contained in other statutes (four of the top five stem from the CFR title on Agriculture, in which this drafting technique seems to be popular) or provide a document map for their respective chapter.

In Germany, the results paint a similar picture: *Sinks* contain provisions for delegation of legislative authority (as expected by legal theory), competencies, statements of goals, or definitions. *Hinges* contain transitional provisions, which are designed to bridge between old and new rules, as well as definitions. The definition classified as a hinge (§ 100a Strafprozeßordnung) establishes a well-known connection between the Criminal Code and its definition of crimes and investigative methods described both inside and outside the Code of Criminal Procedure. All *sources* (and one hinge) are collections of misdemeanors to sanction violations of rules contained in their respective statute or regulation, and they encompass activities as diverse as road traffic, securities trading,

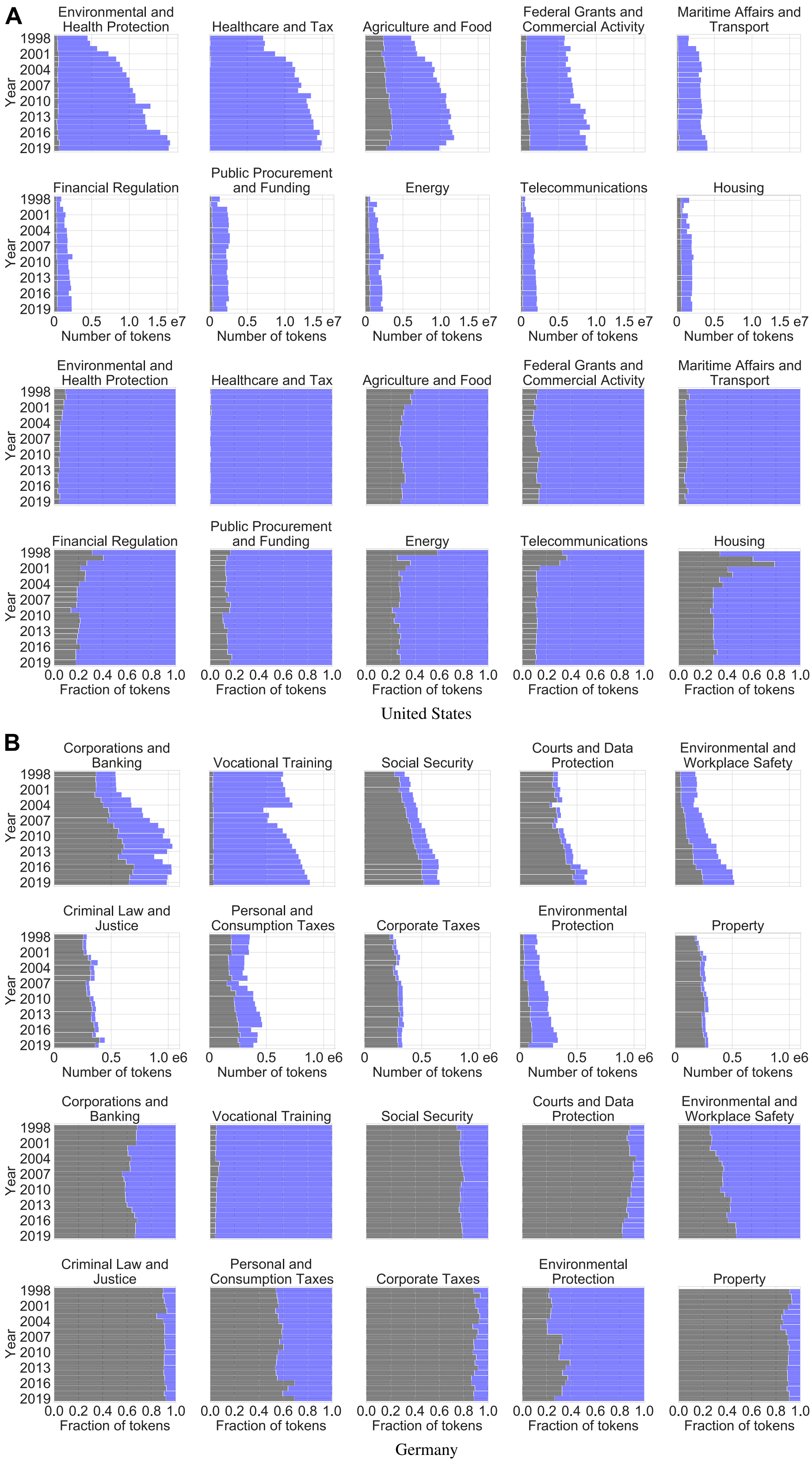


FIGURE 11 | Development of the top ten cluster families from 1998 to 2019 as measured by their absolute size in tokens (rows {1, 2, 5, 6}) and their document type composition (rows {3, 4, 7, 8}) in the United States **(top)** and Germany **(bottom)**. Black areas represent tokens from statutes and blue areas represent tokens from regulations.

TABLE 3 | Top five reference stars of each type in 2019 for the United States (top) and Germany (bottom), with stars whose hubs are contained in regulations marked grey. Edge and degree counts exclude multi-edges; ms is the number of edges between spokes, δ+ is the hub's out-degree, and δ− is the hub's in-degree.

n	m_s	δ^+	δ^-	Type	Hub	Description
					(a) United States	
2721	933	2	2719	Sink	5 USC 552 Public information; agency rules, opinions, orders, records, and proceedings	Authority to delegate agency rules, records, etc. to regulations
1702	26	1	1700	Sink	40 CFR 721.125 Recordkeeping requirements	Authority to require particular records to be kept
1684	28	3	1680	Sink	40 CFR 721.185 Limitation or revocation of certain notification requirements	Criteria and procedure for limitation or revocation of notifications by an agency
1173	298	3	1171	Sink	5 USC 552a Records maintained on individuals	General definitions and procedure for keeping records on individuals
1023	13	0	1022	Sink	40 CFR 721.80 Industrial, commercial, and consumer activities	Definition of a new use of a regulated substance
283	74	31	254	Hinge	8 USC 1101 Definitions	Definitions for subchapter on immigration and nationality
218	150	114	213	Hinge	49 CFR 171.7 Reference Material	Collection of materials to be incorporated by References in other subchapters
141	34	34	127	Hinge	49 CFR 172.101 Purpose and use of hazardous materials table	Collection of substances deemed hazardous materials
138	18	20	117	Hinge	10 USC 101 Definitions	Definitions including bundling of statutes
127	8	24	103	Hinge	15 USC 637 Additional Powers	Authority to carry out actions required throughout the chapter
215	23	213	1	Source	7 CFR 2.22 Under Secretary for Marketing and Regulatory Programs	Enumeration of stand-in duties contained in other statutes
177	13	174	2	Source	7 CFR 2.21 Under Secretary for Research, Education, and Economics	Enumeration of stand-in duties contained in other statutes
150	36	149	0	Source	19 CFR 178.2 Listing of OMB control numbers	Mapping of documents in other parts to control numbers from the office of management and budget
133	16	128	4	Source	7 CFR 2.16 Under Secretary for Farm Production and Conservation	Enumeration of stand-in duties contained in other statutes
129	8	128	0	Source	7 CFR 2.79 Administrator, Agricultural Marketing Service	Enumeration of stand-in duties contained in other statutes
					(b) Germany	
256	19	0	255	Sink	§ 36 Gesetz über Ordnungswidrigkeiten	Determination of the competent authority to prosecute misdemeanor
224	6	0	223	Sink	§ 4 Berufsbildungsgesetz	Authority to delegate vocational training regulations (professions)
194	0	1	192	Sink	§ 25 Gesetz zur Ordnung des Handwerks	Authority to delegate vocational training regulations (crafts)
191	3	0	190	Sink	§ 1 Berufsbildungsgesetz	Goal and definitions for vocational training
180	24	16	168	Sink	§ 1 Gesetz über das Kreditwesen	Definitions for financial and banking regulation
131	27	88	46	Hinge	§ 3 Einkommensteuergesetz	Enumeration of tax-free income types
88	0	74	13	Hinge	Art 229 Weitere überleitungsvorschriften Einführungsgesetz zum Bürgerlichen Gesetzbuche	Transitional provisions of the civil code
86	1	18	68	Hinge	§ 60 Lebensmittel-, Bedarfsgegenstände-und Futtermittelgesetzbuch	Misdemeanors in food and feed safety
84	0	73	10	Hinge	Art 97 übergangsvorschriften Einführungsgesetz zur Abgabenordnung	Transitional provisions of the fiscal code
82	0	61	20	Hinge	§ 100a Strafprozeßordnung	Definition of particularly serious crimes allowing for telecommunication surveillance
76	0	75	0	Source	§ 69a Straßenverkehrs-Zulassungs-Ordnung	Misdemeanors in traffic and road safety
73	1	71	2	Source	§ 340 Kapitalanlagegesetzbuch	Misdemeanors in the capital investment code
59	0	56	2	Source	§ 194 Gesetz zum Schutz vor der schädlichen Wirkung ionisierender Strahlung	Misdemeanors in the radiation protection statute
48	0	47	0	Source	§ 184 Verordnung zum Schutz vor der schädlichen Wirkung ionisierender Strahlung	Misdemeanors in the radiation protection regulation
48	1	45	3	Source	§ 120 Gesetz über den Wertpapierhandel	Misdemeanors and authority to delegate in the securities trading act

and handling radioactive materials. Hence, our classification correctly identifies examples of this popular drafting technique.

As suggested in **Section 3.2.3** and confirmed for the largest stars, the type of a star contains information about a section's function within the legal system. Examining the hundred largest stars, whose types are shown in **Table 4**, exposes different trends in both countries. In the United States, sinks dominate both across document types and over time, accounting for three out of four stars in 1998 and six out of seven stars in 2019, which points to a pronounced drafting preference. At the local level, sections of the United States codified law are mostly connected (only) by referencing the same section, which often contains a definition or the description of a procedure. In Germany, the composition is more balanced to begin with, but sinks still make up the largest

TABLE 4 | Types of the top hundred reference stars (i.e., those with the largest number of nodes) in 1998 and 2019 for the United States (left) and Germany (right). S-Hubs are hubs contained in statutes and R-Hubs are hubs contained in regulations.

	1998		2019	
	S-Hub	R-Hub	S-Hub	R-Hub
(a) United States				
Sink	60	16	56	30
Hinge	8	1	6	3
Source	0	15	0	5
(b) Germany				
Sink	42	0	33	0
Hinge	30	0	52	1
Source	17	11	8	6

share in 1998. Over time, though, the number of sinks and sources among the hundred largest stars decreases in favor of hinges. Hence, individual sections are no longer only connected by a reference to the same section, but the frequently referenced sections themselves increasingly reference other sections. As a consequence, the number of sections that are reachable from any individual section in two hops (i.e., following two references) increases. This makes the information flow via references more efficient, which could explain the reduced need for structural elements to guide information flow via hierarchy in Germany when compared to the United States. But it also increases the prevalence of reference chains, possibly making the German legal system progressively harder to navigate.

Mirroring the larger trends described in **Sections 4.2.1** and **4.2.2**, regulations play a more important role in the United States than in Germany at the micro level of connectivity as well. While the total share of regulation hubs in Germany is small, they make up almost half of the sources among the top hundred reference stars in 2019, which again follows the larger pattern of regulations referencing rather than being referenced. In the United States, regulation hubs account for just under 40% of the top hundred reference stars, but almost all of the largest stars are sinks, regardless of the document type of their hub. This suggests that the United States drafting dynamics resulting in sinks affect both the executive and the legislative branches of government.

4.3 Profiles

In a step toward developing a dashboard for measuring and monitoring the law, we demonstrate the utility of the profiling procedure described in **Section 3.3** by applying it to selected statutes and regulations from the United States and Germany in a case study focusing on financial regulation. We profile a total of four statutes (two from each country) that constitute landmark legislation in this domain and trace their statistics over time in **Figure 12**, along with those of two additional regulations from the same area.

12 USC Ch. 16, popularly known as the Gramm-Leach-Bliley Act or Financial Services Modernization Act of 1999 (GLBA), liberalized the United States financial market by allowing the combination of investment banks, commercial banks, and insurance companies in one institution. It has been in effect for nearly our entire investigation period (1999–2019) and, as indicated by nearly flat lines in all but the panels related to in-degree, has not materially changed. However, the interaction of the GLBA with other parts of the legal system has been anything but static, with its initial weighted in-degree of 1000 increasing by 60% between 1999 and 2019 due to incoming references from other statutes and regulations. Unlike the growth trend in the weighted in-degree, the growth trend in the binary in-degree is nearly monotonic. This indicates that most of the fluctuations in the GLBA's regulatory environment occur within individual chapters of the USC or the CFR. In summary, the GLBA can therefore be rightfully regarded as a landmark statute, which has required little engineering but has remained an important reference throughout the period under study.

The profile of 12 USC Ch. 53, popularly known as the Dodd–Frank Act or the Wall Street Reform and Consumer Protection Act (DFA), shows similarities with the GLBA in most statistics we track. Introduced in response to the Great Recession in 2010, it is approximately half as old as the GLBA, and like the GLBA, it has barely changed in size, breadth, or structure. But although the DFA is comparable to the GLBA in size, its interaction with the environment seems much more dynamic, with its weighted in-degree increasing by a factor of almost ten over its lifetime. In absolute terms, however, the references increase by less than 500, i.e., in the same order of magnitude as for the GLBA. This is in line with the fact that both statutes are part of the same legal domain, and it highlights how much the evolution of individual pieces of legislation is influenced by their initial conditions, e.g., whether they are strongly connected with their environment already at birth. For the DFA, the growth of the weighted in-degree again is not monotonic, but it is visibly steeper before 2017 than afterwards, leveling off in the last years of the period under study. The binary in-degree, whose gradient is almost monotonically decreasing from the start, anticipates this deceleration by several years. This suggests the existence of an onboarding period, in which the DFA is integrated into the regulatory environment before finding its place in the United States legal system (see the related discussion in [61]).

The profiles of the two German statutes we examine are starkly different from those of the United States statutes. Statutes under the names of both the Stock Exchange Act (SEA) and the Securities Trading Act (STA) have been in effect for the entire observation period. As indicated by their unique token count, they are both constantly narrow in thematic scope (with the STA an order of magnitude more narrow than the SEA to start with), and their token count increases over time. While the SEA and the STA start at comparable sizes in 1998, the STA grows by a factor of seven, while the SEA merely doubles. Possibly as a result, the SEA largely maintains its original number of structures, while the number of structural elements in the STA triples. This is accompanied by an expected, nearly parallel increase in the number of STA sections, and even a decrease of about 25% in the number of SEA sections. Beyond the general growth trends present in almost all STA statistics, the period from 2006 to 2007 stands out, as most of its statistics experience a relatively steep increase between these years. The doctrinal investigation prompted by this observation reveals that the source of the

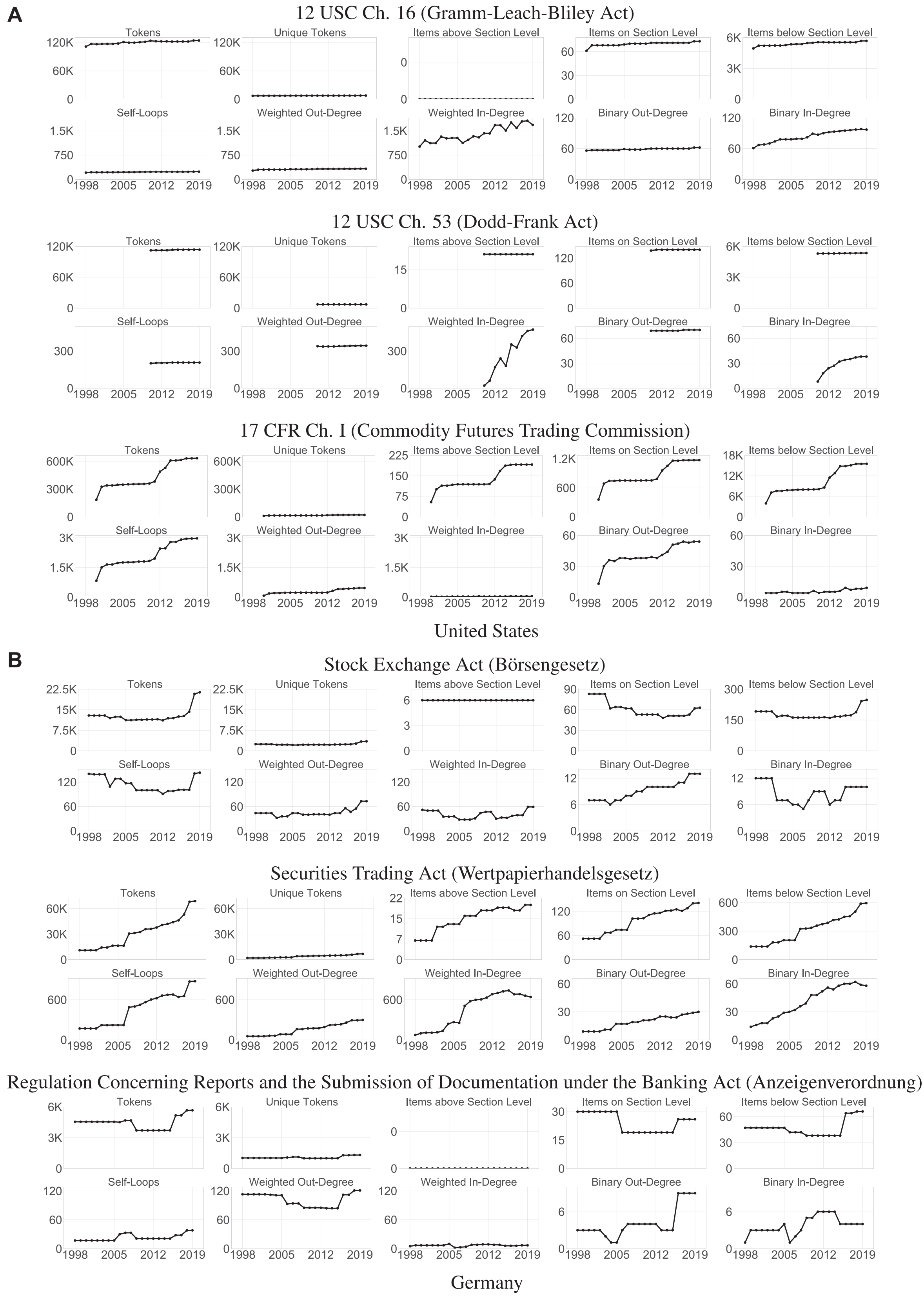


FIGURE 12 | Profiles tracking the evolution of selected laws related to financial regulation for the United States **(top)** and Germany **(bottom)** from 1998 to 2019.

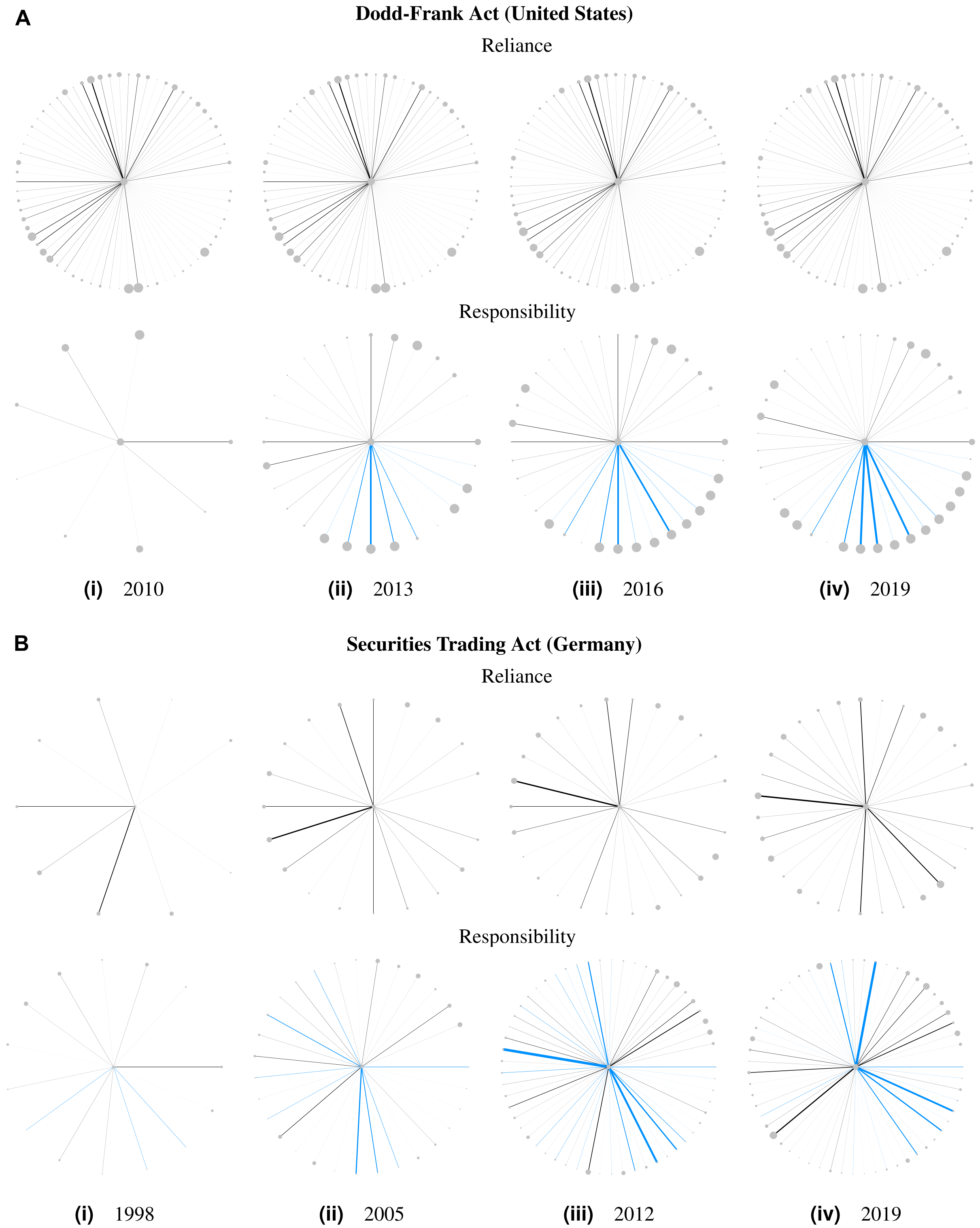


FIGURE 13 | Reliance and responsibility of the Dodd–Frank Act (DFA, top) and the Securities Trading Act (STA, Wertpapierhandelsgesetz, bottom) from 1998 to 2019. Leaf nodes are chapters of the USC or the CFR that are cited by (reliance) or cite (responsibility) the DFA **(top)**, and statutes or regulations (or their books, if applicable) that are cited by (reliance) or cite (responsibility) the STA **(bottom)**. Edge types indicate reference types as used in **Figure 3** (black for lateral statute references, light blue for upward references, and silver for downward references). Node size is proportional to the node's number of tokens; edge width is proportional to the number of references represented by the edge.

increases is a legislative project translating extensive transparency requirements mandated by the European Union into German law (Transparenzrichtlinie-Umsetzungsgesetz), which came into effect in January 2007. This finding also shows how our methods can complement doctrinal legal scholarship, as has been demonstrated, e.g., for the development of the STA over its entire lifetime [28].

Our statistics produce interesting insights not only for statutes but also for regulations. For example, there is a noticeable increase in the self-referentiality of the CFR chapter about the Commodity Futures Trading Commission (CFTCR), and the German Reports and Documents Regulation (RDR) shows structural changes between 2005 and 2006 as well as between 2015 and 2016. As our framework enables the joint modeling of data from different document types, its application can surface characteristic differences between these types, too. Examining our exemplary regulations and statutes in **Figure 12**, we find that the CFTCR experiences noticeable growth (about 200%), while the size of the featured statutes barely changes. At the same time, the regulation's weighted in-degree is several orders of magnitude smaller than that of the DFA or the GLBA, supporting the intuition that statutes should be referenced more frequently than regulations for this particular case. In Germany, the RDR is smaller than the featured German statutes, and it covers less diverse content (as would be expected for a regulation from a legal theory perspective). Its structural organization is minimal, as is its self-referentiality. This confirms that smaller units of law require less internal organization by both structure and references. The RDR references, and is referenced by, a small number of different documents, indicating homogeneity in its regulatory environment. Its weighted out-degree falls between the SEA and STA, i.e., it has a non-negligible number of references to a limited number of targets. In summary, the RDR has most characteristics expected from a German regulation, and its overall profile can clearly be distinguished from that of the featured German statutes.

Our framework enables comparisons not only across document types but also across countries. When examining the DFA and the STA, we see that the STA starts at a size of roughly a third of the DFA but grows to two thirds of the DFA over time. Both statutes have similar degrees of structure at the section level and above, but the DFA contains ten to fifteen times the amount of items below section level, indicating a vastly more granular hierarchical organization. Conversely, the STA contains three to four times more self-loops than the DFA, with its weighted in-degree about 1.5 times and its binary in-degree between 1.5 and 2 times higher than those of the DFA after the first couple of years. This mirrors the more general finding that rule-making agents in the United States and Germany favor different mechanisms to handle the token growth of their statutory corpora: Americans like adding hierarchical structure, while Germans prefer adding references [27].

Figure 13 combines profile statistics concerning size and interdependence to enable direct visual comparisons. Here, we compare the ego graphs of the DFA and the STA for every quarter of their existence during our investigation period. Note that the distance between the snapshots is different for both statutes, as the DFA was adopted only in the middle of our study period, but both series end in 2019. Complementing the statistics presented in **Figure 12**, the visualization attributes the references to their actual sources and targets, indicating their number by the weight of the edges. For the DFA, its *reliance* (i.e., how much and how diversely it references other statutes and regulations) barely changes, while its *responsibility* (i.e., how much and how diversely it is referenced by other statutes and regulations) discernibly increases, as the DFA becomes more and more integrated with its environment. In contrast, both the reliance and the responsibility of the STA increase over time, with its responsibility starting nearly twice as high and increasing at a much faster rate than its reliance. As shown by the edge colors, the diverse responsibility of the STA concerns both statutes and regulations, and the DFA is most intensively responsible for regulations. In line with the expectation derived from legal theory, both the DFA and the STA rely mostly on statutes.

5 DISCUSSION

We have introduced an analytical framework for the dynamic network analysis of legal documents and demonstrated its utility by applying it to a dataset comprising federal statutes and regulations in the United States and Germany over a period of more than 20 years. The limitations of this work concern two separate areas: the methods introduced in **Section 3** and the results presented in **Section 4**.

To gravitate toward its ideal formulation, our framework requires further refinement based on experiences from applications to diverse datasets. Our model is deliberately document type and country agnostic, such that it can be easily instantiated for new data. Similar studies using legal documents from a variety of jurisdictions would be of immense value for improving our framework, and they could provide further context for the results reported in **Section 4**. Furthermore, our network analytical framework could be complemented by a framework for natural legal language processing, as the combination of relational information and linguistic information will likely lead to insights that would not be possible using either of these sources alone.

When preparing this article, we found that combining documents of different types in one graph representation raises many conceptual questions. Some of these questions relate to the presentation of our results, e.g., whether to depict dynamics in absolute or relative terms (thereby either impairing comparisons across document types of different sizes or visually overstating dynamics for small baselines). Others concern design decisions when defining our methods, e.g., whether tokens from documents of different types should have the same weight when determining cluster families even if there is a striking imbalance between the total number of tokens in documents of these types (as is the case in our United States data). Here, one alternative would be to rescale the token counts before constructing the cluster family graph, such that the total influence of tokens from one specific document type is equal across all types. While this would change the results to a certain extent, it is difficult to assess whether the modified method would be superior because comparable investigations of multimodal legal document networks currently do not exist.

The results stated in **Section 4** are limited in geographic scope (United States and Germany), temporal scope (1998–2019), and institutional scope (legislative and executive branch on the federal level). Most importantly, our findings cover only codified law. As the United States and Germany are typically assumed to follow distinct legal traditions (common law and civil law), which are often thought to differ, inter alia, in their reliance on court precedent, including court decisions might have disparate impact on our results for both countries. However, it could also provide empirical evidence against the traditional classification. Irrespective of legal traditions, unlocking and integrating judicial data is an important direction for future work.

Regarding both growth and connectivity, the next steps consist in eliminating the uncertainties and limitations affecting our data. For example, as highlighted in **Section 2.2**, one important stride toward a more comprehensive picture of the connectivity between legal documents is the extraction and resolution of non-atomic references. At the macro level, connectivity could also be evaluated at other resolutions (e.g., the chapter level) or when including hierarchy edges, and our analysis could be expanded using further statistics, such as motif counts and their evolution over time. Furthermore, applying our methods to other document types or other countries would help us assess whether the rocket structure we found in our data is characteristic of legal systems in general. When assessing connectivity at the meso level, the dynamic map of law provided by our cluster families could be further refined, especially at other resolution levels. At the micro level of connectivity, a more fine-grained star taxonomy might be in order because in both countries, there exists some functional overlap between hinge stars and sink stars. For the profiles, a sensible step forward would be to apply the tracing methodology at other levels of resolution (e.g., at the level of individual sections), and the statistics we track could be complemented by similarity measures allowing us to compare between the different units of law we analyze.

Beyond the specific opportunities for further research outlined above, our work raises three larger questions to be explored in the future:

1. *When quantitatively analyzing legal documents, how should we choose the unit of analysis?*
   On the one hand, no clear consensus exists as to what constitutes a *unit of law* or a *legal rule*. But on the other hand, the choice has far-reaching consequences for all analyses. Furthermore, even analyzing all documents at the same structural level presents problems: Legal rules come in various sizes, and at times, a single paragraph might be longer than the average document due to drafting decisions by the agents in the legal system. This complicates comparisons and creates countless opportunities for erroneous interpretations. Detailing the full rationale behind all choices we made when presenting our results in **Section 4** is beyond the scope of this article. However, an extensive exposition of the possible choices and the tradeoffs surrounding them would benefit the research community at large and, therefore, constitutes a fecund field for future findings.
2. *How can we measure the regulatory energy of statutes?*
   The analysis of individual statutes such as the Gramm-Leach-Bliley Act and the Dodd–Frank Act suggests that legislative outputs impact their environments at potentially different rates (e.g., by prompting further rule making), i.e., that they have a certain *regulatory energy* that they emit over time. This hypothesis could be validated, inter alia, using external data on regulatory relevance, e.g., the filings concerning regulatory risk that are required for annual and transition reports pursuant to sections 13 or 15(d) of the Securities Exchange Act of 1934 under 17 CFR 249.310 – Form 10-K [62]. However, other approaches are equally possible and merit further investigation.
3. *How can we connect our empirical findings to established theories in law and political science?*
   Although beyond the scope of this work, some of our findings can be combined with analyses using established theories on the composition and evolution of codified law in both legal scholarship and political science. The most prominent example here is the question of delegation: How does it happen and what are its limits, in theory and in practice? This touches the heart of democratic legitimacy, and it presents a promising opportunity for empirical legal studies to contribute to mainstream legal and political science discourse that we are planning to seize in the future.

DATA AVAILABILITY STATEMENT

For the United States, the raw input data used in this study is publicly available from the Annual Historical Archives published by the Office of the Law Revision Counsel of the U.S. House of Representatives and the United States Government Publishing Office, and is also available from the authors upon reasonable request. For Germany, the raw input data used in this study was obtained from juris GmbH but restrictions apply to the availability of this data, which was used under license for the current study, and so is not publicly available. For details, see Section 1 of the **Supplementary Material**. The preprocessed data used in this study (for both the United States and Germany) is archived under the following DOI: https://doi.org/10.5281/zenodo.4660133. The code used in this study is available on GitHub in the following repositories:

- Paper: https://github.com/QuantLaw/Measuring-Law-Over-Time
  DOI of publication release: https://doi.org/10.5281/zenodo.4660191

- Data preprocessing: https://github.com/QuantLaw/legal-data-preprocessing DOI of publication release: https://doi.org/10.5281/zenodo.4660168
- Clustering: https://github.com/QuantLaw/legal-data-clustering DOI of publication release: https://doi.org/10.5281/zenodo.4660184

AUTHOR CONTRIBUTIONS

CC and JB have contributed equally to this work and share first authorship. All authors conceived of the research project. CC, JB, and MB performed the computational analysis in consultation with DM, and DH. CC, DH, and MB drafted the manuscript, which was revised and reviewed by all authors. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

FUNDING

The Center for Legal Technology and Data Science, Bucerius Law School is covering the open access publication fees.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2021.658463/full#supplementary-material

REFERENCES

1. Mitchell M. *Complexity: A Guided Tour*. Oxford: Oxford University Press (2009).
2. Miller JH, and Page SE. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*. Princeton: Princeton University Press (2009). doi:10.1515/9781400835522
3. Ruhl JB, Katz DM, and Bommarito MJ. Harnessing Legal Complexity. *Science* (2017) 355:1377–8. doi:10.1126/science.aag3013
4. Murray J, Webb T, and Wheatley S. *Complexity Theory And Law: Mapping An Emergent Jurisprudence*. Oxfordshire: Routledge (2018).
5. Ruhl JB. Complexity Theory as a Paradigm for the Dynamical Law-And-Society System: A Wake-Up Call for Legal Reductionism and the Modern Administrative State. *Duke LJ* (1995) 45:849–928.
6. Scott RE. Chaos Theory and the Justice Paradox. *William Mary L Rev* (1993) 35:329–51.
7. Ruhl J, and Katz DM. Measuring, Monitoring, and Managing Legal Complexity. *Iowa L Rev* (2015) 101:191–244.
8. Amaral LAN, and Ottino JM. Complex Networks. *Eur Phys J B Condens Matter* (2004) 38:147–62. doi:10.1140/epjb/e2004-00110-5
9. Albert R, and Barabási A-L. Statistical Mechanics of Complex Networks. *Rev Mod Phys* (2002) 74:47–97. doi:10.1103/revmodphys.74.47
10. Watts DJ, and Strogatz SH. Collective Dynamics of 'Small-World' Networks. *Nature* (1998) 393:440–2. doi:10.1038/30918
11. Coupette C. *Juristische Netzwerkforschung: Modellierung, Quantifizierung und Visualisierung relationaler Daten im Recht [Legal Network Science: Modeling, Measuring, and Mapping Relational Data in Law]*. Tübingen: Mohr Siebeck (2019). doi:10.1628/978-3-16-157012-4
12. Winkels R (2019). "Exploiting the Web of Law," in *Law, Public Policies And Complex Systems: Networks in Action*. Editors R. Boulet, C. Lajaunie, and P. Mazzega (Berlin: Springer), p. 205–19. doi:10.1007/978-3-030-11506-7_10
13. Black RC, and Spriggs JF. The Citation and Depreciation of U.S. Supreme Court Precedent. *J Empirical Leg Stud* (2013) 10:325–58. doi:10.1111/jels.12012
14. Lupu Y, and Fowler JH. Strategic Citations to Precedent on the U.S. Supreme Court. *J Leg Stud* (2013) 42:151–86. doi:10.1086/669125
15. Bommarito MJ, Katz DM, and Isaacs-See J. An Empirical Survey of the Population of U.S. Tax Court Written Decisions. *Va Tax Rev* (2011) 30: 523–57. doi:10.1016/0148-2963(87)90018-x
16. Cross FB, Spriggs JF, Johnson TR, and Wahlbeck PJ. *Citations in the U.S. Supreme Court: An Empirical Study of Their Use and Significance. Univ. Ill. L. Rev.* (2010) 489–575.
17. Fowler JH, Johnson TR, Spriggs JF, Jeon S, and Wahlbeck PJ. Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Polit Anal* (2007) 15:324–46. doi:10.1093/pan/mpm011
18. Olsen HP, and Esmark M. In: R Whalen, editor. *Needles in a Haystack: Using Network Analysis to Identify Cases that Are Cited for General Principles of Law by the European Court of Human Rights. Computational Legal Studies*. Cheltenham: Edward Elgar Publishing (2020). p. 293–311. doi:10.4337/9781788977456.00018
19. Alschner W, and Charlotin D. The Growing Complexity of the International Court of Justice's Self-Citation Network. *Eur J Int L* (2018) 29:83–112. doi:10.1093/ejil/chy002
20. Larsson O, Naurin D, Derlén M, and Lindholm J. Speaking Law to Power: the Strategic Use of Precedent of the Court of Justice of the European Union. *Comp Polit Stud* (2017) 50:879–907. doi:10.1177/0010414016639709
21. Tarissan F, Panagis Y, and Šadl U. In: VS Subrahmanian and J Rokne, editors. *Selecting the Cases that Defined Europe: Complementary Metrics for a Network Analysis. IEEE/ACM ASONAM 2016*. (New York City: IEEE Press) (2016). p. 661–8.
22. Panagis Y, and Šadl U. The Force of EU Case Law: A Multi-Dimensional Study of Case Citations. In: Rotolo, editor. *Legal Knowledge and Information Systems*. Amsterdam: IOS Press (2015). p. 71–80.
23. Pelc KJ. The Politics of Precedent in International Law: A Social Network Application. *Am Polit Sci Rev* (2014) 108:547–64. doi:10.1017/s0003055414000276
24. Lupu Y, and Voeten E. Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights. *Br J. Polit. Sci.* (2012) 42:413–39. doi:10.1017/s0007123411000433
25. Charlotin D. The Place of Investment Awards and WTO Decisions in International Law: A Citation Analysis. *J Int Econ L* (2017) 20:279–99. doi:10.1093/jiel/jgx010
26. Langford M, Behn D, and Lie RH. The Revolving Door in International Investment Arbitration. *J Int Econ L* (2017) 20:301–32. doi:10.1093/jiel/jgx018
27. Katz DM, Coupette C, Beckedorf J, and Hartung D. Complex Societies and the Growth of the Law. *Sci Rep* (2020) 10:18737. doi:10.1038/s41598-020-73623-x
28. Coupette C, and Fleckner AM. *Das Wertpapierhandelsgesetz (1994–2019)—Eine quantitative juristische Studie [The Securities Tradining Act (1994–2019)—A Quantitative Legal Study]. Festschrift 25 Jahre WpHG*. (Berlin: De Gruyter) (2019). p. 53–86. doi:10.1515/9783110632323-005
29. Boulet R, Mazzega P, and Bourcier D. Network Approach to the French System of Legal Codes Part II: the Role of the Weights in a Network. *Artif Intell L* (2018) 26:23–47. doi:10.1007/s10506-017-9204-y
30. Koniaris M, Anagnostopoulos I, and Vassiliou Y. Network Analysis in the Legal Domain: A Complex Model for European Union Legal Sources. *J Complex Networks* (2018) 6:243–68. doi:10.1093/comnet/cnx029
31. Li W, Azar P, Larochelle D, Hill P, and Lo AW. Law Is Code: A Software Engineering Approach to Analyzing the United States Code. *J Bus Technol L* (2015) 10:297–374. doi:10.2139/ssrn.2511947
32. Katz DM, and Bommarito MJ. Measuring the Complexity of the Law: the United States Code. *Artif Intell L* (2014) 22:337–74. doi:10.1007/s10506-014-9160-8
33. Bommarito MJ, II, and Katz DM. A Mathematical Approach to the Study of the United States Code. *Physica A: Stat Mech its Appl* (2010) 389:4195–200. doi:10.1016/j.physa.2010.05.057

34. Lee B, Lee KM, and Yang JS. Network Structure Reveals Patterns of Legal Complexity in Human Society: The Case of the Constitutional Legal Network. *PLoS ONE* (2019) 14:e0209844. doi:10.1371/journal.pone.0209844
35. Rutherford A, Lupu Y, Cebrian M, Rahwan I, LeVeck BL, and Garcia-Herranz M. Inferring Mechanisms for Global Constitutional Progress. *Nat Hum Behav* (2018) 2:592–9. doi:10.1038/s41562-018-0382-8
36. Rockmore DN, Fang C, Foti NJ, Ginsburg T, and Krakauer DC. The Cultural Evolution of National Constitutions. *J Assoc Inf Sci Technol* (2017) 69:483–94. doi:10.1002/asi.23971
37. Boulet R, Barros-Platiau AF, and Mazzega P. In: R Boulet, C Lajaunie, and P Mazzega, editors. *Environmental and Trade Regimes: Comparison of Hypergraphs Modeling the Ratifications of Un Multilateral Treaties*. *Law, Public Policies and Complex Systems: Networks in Action*. Cham: Springer (2019). p. 221–42. doi:10.1007/978-3-030-11506-7_11
38. Alschner W, and Skougarevskiy D. Mapping the Universe of International Investment Agreements. *J Int Econ L* (2016) 19:561–88. doi:10.1093/jiel/jgw056
39. Kim RE. The Emergent Network Structure of the Multilateral Environmental Agreement System. *Glob Environ Change* (2013) 23:980–91. doi:10.1016/j.gloenvcha.2013.07.006
40. Kinne BJ. Network Dynamics and the Evolution of International Cooperation. *Am Polit Sci Rev* (2013) 107:766–85. doi:10.1017/s0003055413000440
41. Saban D, Bonomo F, and Stier-Moses NE. Analysis and Models of Bilateral Investment Treaties Using a Social Networks Approach. *Physica A: Stat Mech Appl* (2010) 389:3661–73. doi:10.1016/j.physa.2010.04.001
42. Katz DM, and Stafford DK. Hustle and Flow: a Social Network Analysis of the American Federal Judiciary. *Ohio St LJ* (2010) 71:457–510. doi:10.5749/minnesota/9780816692286.001.0001
43. Easley D, and Kleinberg J *Networks, Crowds, and Markets*. Cambridge: Cambridge University Press (2010). doi:10.1017/cbo9780511761942
44. Barabási AL. *Network Science*. Cambridge: Cambridge University Press (2016).
45. Newman M. *Networks*. Oxford: Oxford University Press (2018). doi:10.1093/oso/9780198805090.003.0016
46. Katz DM, Gubler JR, Zelner J, and Bommarito MJ. Reproduction of Hierarchy-A Social Network Analysis of the American Law Professoriate. *J Leg Education* (2011) 61:76–103. doi:10.1007/978-1-4614-6170-8_100596
47. Newton BE. Law Review Scholarship in the Eyes of the Twenty-First-Century Supreme Court Justices: An Empirical Analysis. *Drexel L Rev* (2011) 4: 399–416. doi:10.1017/cbo9780511977480.010
48. Schwartz DL, and Petherbridge L. The Use of Legal Scholarship by the Federal Courts of Appeals: An Empirical Study. *Cornell L Rev* (2011) 96:1345–72. doi:10.1111/jels.12081
49. George TE. An Empirical Study of Empirical Legal Scholarship: the Top Law Schools. *Indiana L J* (2006) 81:141–61. doi:10.1111/jels.12007
50. Ellickson RC. Trends in Legal Scholarship: A Statistical Study. *J Leg Stud* (2000) 29:517–43. doi:10.1086/468084
51. Pound R. Law in Books and Law in Action. *Am L Rev* (1910) 44:12–36.
52. Broder A, Kumar R, Maghoul F, Raghavan P, Rajagopalan S, Stata R, et al. Graph Structure in the Web. *Computer Networks* (2000) 33:309–20. doi:10.1016/s1389-1286(00)00083-9
53. Friedlander T, Mayo AE, Tlusty T, and Alon U. Evolution of Bow-Tie Architectures in Biology. *PLOS Comput Biol* (2015) 11:e1004055. doi:10.1371/journal.pcbi.1004055
54. Csete M, and Doyle J. Bow Ties, Metabolism and Disease. *Trends Biotechnol* (2004) 22:446–50. doi:10.1016/j.tibtech.2004.07.007
55. Ma H-W, and Zeng A-P. The Connectivity Structure, Giant Strong Component and Centrality of Metabolic Networks. *Bioinformatics* (2003) 19:1423–30. doi:10.1093/bioinformatics/btg177
56. Rosvall M, and Bergstrom CT. Maps of Random Walks on Complex Networks Reveal Community Structure. *Proc Natl Acad Sci* (2008) 105:1118–23. doi:10.1073/pnas.0706851105
57. Rosvall M, Axelsson D, and Bergstrom CT. The Map Equation. *Eur Phys J Spec Top* (2009) 178:13–23. doi:10.1140/epjst/e2010-01179-1
58. Lancichinetti A, and Fortunato S. Consensus Clustering in Complex Networks. *Sci Rep* (2012) 2:336. doi:10.1038/srep00336
59. Merton RK. The Matthew Effect in Science: The Reward and Communication Systems of Science Are Considered. *Science* (1968) 159:56–63. doi:10.1126/science.159.3810.56
60. Price DDS. A General Theory of Bibliometric and Other Cumulative Advantage Processes. *J Am Soc Inf Sci* (1976) 27:292–306. doi:10.1002/asi.4630270505
61. McLaughlin PA, Sherouse O, Febrizio M, and King MS. Is Dodd–Frank the Biggest Law Ever? *J Financ Regul* (2021). doi:10.1093/jfr/fjab001
62. Bommarito MJ, and Katz DM. Measuring and Modeling the US Regulatory Ecosystem. *J Stat Phys* (2017) 168:1125–35. doi:10.1007/s10955-017-1846-3

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 04 June 2021
doi: 10.3389/fphy.2021.671882

Complexity and Entropy in Legal Language

Roland Friedrich *

ETH Zurich, D-GESS, Zurich, Switzerland

We study the language of legal codes from different countries and legal traditions, using concepts from physics, algorithmic complexity theory and information theory. We show that vocabulary entropy, which measures the diversity of the author's choice of words, in combination with the compression factor, which is derived from a lossless compression algorithm and measures the redundancy present in a text, is well suited for separating different writing styles in different languages, in particular also legal language. We show that different types of (legal) text, e.g. acts, regulations or literature, are located in distinct regions of the complexity-entropy plane, spanned by the information and complexity measure. This two-dimensional approach already gives new insights into the drafting style and structure of statutory texts and complements other methods.

Keywords: information theory, complex systems, linguistics, legal theory, algorithmic complexity theory, lossless compression algorithms, Shannon entropy

OPEN ACCESS

Edited by:
Pierpaolo Vivo, King's College London, United Kingdom

Reviewed by:
Eric DeGiuli, Ryerson University, Canada
Alessandro Vezzani, National Research Council (CNR), Italy

***Correspondence:**
Roland Friedrich
roland.friedrich@gess.ethz.ch

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *24 February 2021*
Accepted: *03 May 2021*
Published: *04 June 2021*

Citation:
Friedrich R (2021) Complexity and Entropy in Legal Language. Front. Phys. 9:671882. doi: 10.3389/fphy.2021.671882

1 INTRODUCTION

The complexity of the law has been the topic of both scholarly writing and scientific investigation, with the main challenge being the proper definition of "complexity". Historically, articles in law journals took a conceptual and non-technical approach toward the "complexity of the law", motivated by practical reasons, such as the ever increasing amount of legislation produced every year and the resulting cost of knowledge acquisition, e.g. [1, 2]. Although this approach is important, it remains technically vague and not accessible to quantitative analysis and measurement. Over the past decade, with the increasing availability of digitized (legal) data and the steady growth of computational power, a new type of literature has emerged within legal theory, the authors of which use various mathematical notions that come from areas as diverse as physics and information theory or graph theory, to analyze the complexity of the law, cf. e.g. [3–5]. The complexity considered results mainly from the exogenous structure of the positive law, i.e. the tree-like hierarchical organization of the legal texts in a forest consisting of codes (root nodes), chapters, sections, etc., but also from the associated reference network.

According to the dichotomy introduced by [6]; one can distinguish between structure-based measures and content-based measures of complexity, with the former pertaining to the field of knowledge representation (knowledge engineering) and the latter relating to the complexity of the norms, which includes, e.g. the (certainty of) legal commands, their efficiency and socio-economic impact.

In this article, we advance the measurement of legal complexity by focusing on the language using a method originating in the physics literature, cf. [7]. So, we map legal documents from several major legal systems into a two-dimensional complexity-entropy plane, spanned by the (normalized) vocabulary entropy and the compression factor, cf. **Section 2.1**. Using an abstract and rigorous measurement of the complexity of the law, should have significant practical benefits for policy, as discussed previously by, e.g. [1, 2]. For example, it could potentially identify parts of the law that need to be rewritten in order to remain manageable, thereby reducing the costs for citizens and firms who are supposed to comply. Most notably, the French Constitutional Court has ruled that articles of unjustified "excessive-complexity" are unconstitutional[1]. However, in order to render the notion of "excessive complexity" functional, quantitative methods are needed such as those used by [5, 8]; and which our version of the complexity-entropy plane ideally complements.

2 COMPLEXITY AND ENTROPY

A non-trivial question that arises in several disciplines is how the complexity of a hierarchical structure, i.e. of a multi-scale object, can be measured. Different areas of human knowledge are coded as written texts that are organized hierarchically, e.g. each book's Table of Contents reflects its inherent hierarchical organization as a tree, and all books together form a forest. Furthermore, a tree-like structure appears again at the sentence level in the form of the syntax tree and its semantics as an additional degree of freedom. Although various measures of complexity have been introduced that are specially adapted to a particular class of problems, there is still no unified theory. The first concept we consider is Shannon entropy, [9]; which is a measure of information. It is an observable on the space of probability distributions with values in the non-negative real numbers. For a discrete probability distribution $P := \{p_1, \ldots, p_N\}$, with $p_i > 0$, for all i, and $\sum_{i=1}^{N} p_i = 1$, the Shannon entropy $H(P)$, is defined as:

$$H(P) := -\sum_{i=1}^{N} p_i \log_2(p_i), \tag{1}$$

with $\log_2$, the logarithm with base 2. The normalized Shannon entropy $H_n(P)$, is given by

$$H_n(P) := \frac{H(P)}{\log_2(N)}, \tag{2}$$

i.e. by dividing $H(P)$ by the entropy $H(P_N)$ of the discrete uniform distribution $P_N := \{1/N, \ldots, 1/N\}$, for N different outcomes. We shall use the normalized entropy in order to measure the information content of the vocabulary of individual legal texts, for details cf. **Section 6.3**. Word entropies have previously been used by various authors. In the legal domain [5], calculated the word entropy, after removing stop words, for the individual Titles of the U.S. Code. [10] used word entropies to gauge Shakespeare's and Jin Yong's writing capacity, based on the 100 most frequent words in each text.

The second concept we consider is related to Kolmogorov complexity (cf. [11, 12] and references therein), which is the prime example of algorithmic (computational) complexity. Heuristically, the complexity of an object is defined as the length of the shortest of all possible descriptions. Further fundamental examples of algorithmic complexity include Lempel-Ziv complexity C_{76}, [13]; or Wolfram's complexity measure of a regular language, [14]. The latter is defined as (logarithm of) the minimal number of nodes of a deterministic finite automaton (DFA) that recognizes the language (Meyhill-Nerode theorem). In order to facilitate the discussion, let us propose a set of axioms for a complexity measure. This measure is basically a general form of an outer measure.

Let X be (at least) a monoid $(X, \circ, \varepsilon)$, with binary composition $\circ : X \times X \to X$, and identity element ε, and additionally, let $\geq$ be a partial order on X.

A complexity measure C on X, is a functional $C : X \to \mathbb{R}_+$, such that for all $a, b \in X$, we have

pointed:

$$C(\varepsilon) = 0, \tag{3}$$

monotone:

$$\text{if } a \leq b \text{ then } C(a) \leq C(b), \tag{4}$$

sub-additive:

$$C(a \circ b) \leq C(a) + C(b). \tag{5}$$

Examples satisfying the above axioms include tree structures, with the (simple) complexity measure given by the number of levels, i.e. the depth from the baseline. Then the empty tree has zero complexity, the partial order being given by being a sub-tree and composition being given by grafting trees. Further, the Lempel-Ziv complexity C_{76}, and Wolfram's complexity measure for regular languages, if slightly differently defined via recognizable series, satisfy the axioms. However, plain Kolmogorov complexity does not satisfy, e.g. sub-additivity, cf. the discussion by [12].

2.1 Compression Factor

A derived complexity measure is the compression factor, which we consider next, and which is obtained from a lossless compression algorithm, such as, [15, 16].

A lossless compression algorithm, i.e. a compressor γ, reversibly transforms an input string s into a sequence $\gamma(s)$ which is shorter than the original one, i.e. $|\gamma(s)| \leq |s|$, but contains exactly the same information as s, cf. e.g. [17, 18].

For a string s, the compression factor $r = r(s)$, is defined as

$$r(s) := \frac{|s|}{|\gamma(s)|}. \tag{6}$$

The inverse r^{-1}, is called the compression ratio. These derived complexity measures quantify the relative amount of redundancy or structure present in a string, or more generally data.

The compression factor, as the entropy rate, is a relative quantity which permits to directly compare individual data items, independently of their size.

Let us illustrate this for the Lempel-Ziv complexity measure C_{76}, cf. [13]; and the following strings of length 20:

$$s_1 := 00000000000000000000,$$
$$s_2 := 01010101010101010101,$$
$$s_3 := 01001010100110101101.$$

[1]Conseil Constitutionnel, Décision n 2005–530 DC du 29 décembre 2005 (Loi de Finances pour 2006) 77–89, available at https://www.conseil-constitutionnel.fr/decision/2005/2005530DC.htm.

Then we have $C_{76}(s_1) = 2$, $C_{76}(s_2) = 3$ and $C_{76}(s_3) = 7$, from which one immediately obtains the respective compression factors. [19]; showed that a generic string of length n has complexity close to n, i.e. it is "random", however the meaningful strings for humans, i.e. representing text, images etc., are not random and have a structure between the completely uniform and the random string, cf. [18, 19]. [20] introduced a quantity related to the compression factor, called the "computable information density", which is a measure of order and correlation in (physical) systems in and out of equilibrium. Compression factors (ratios) were previously used by [21]; who measured the complexity of mulitple languages by compressing texts and their shuffled versions to measure the inherent linguistic order. [22]; additionally to a neural language model, utilized compression ratios to measure the complexity of the language used by the Supreme Courts of the U.S. (USSC) and Germany (BGH). [23]; using the Lempel-Ziv complexity measure C_{76}, took into account not only the order inherent in a grammatically correct sentence, but also the larger organization of a text document, e.g. sections, by selectively shuffling the data belonging to each level of the hierarchy.

3 SOME REMARKS ON COMPLEXITY, ENTROPY AND LANGUAGE

[24] (pp. 10–11) intuitively describe the broad difference between classical information theory and algorithmic complexity, which we summarize next. Whereas information theory (entropy), as conceived by Shannon, determines the minimal number of bits needed to transmit a set of messages, it does not provide the number of bits necessary to transmit a particular message from the set. Kolmogorov complexity on the other hand, focuses on the information content of an individual finite object, e.g. a play by Shakespeare, accounting for the (empirical) fact that strings which are meaningful to humans, are compressible, cf. [19]. In order to relate entropy, Kolmogorov complexity or Ziv-Lempel compression to one another, various mathematical assumptions such as stationarity, ergodicity or infinity are required, cf. [11, 17, 25]. Also, the convergence of various quantities found in natural languages, e.g. entropy estimates, [26]; are based on some of these assumptions. Despite the fact that the different approximations and assumptions proved valuable for language models, natural language is not necessarily generated by a stationary ergodic process, cf. [11]; as e.g., cf. [25]; the probability of upcoming words can depend on words which are far away. But, as argued by [27]; it is precisely due to the non-ergodic nature of natural language that one can empirically distinguish different topics, e.g. by determining the uneven distribution of keywords in texts, cf. also [28]. [29] considered a model of a random languages and showed how structure emerges as a result of the competition between energy and entropy.

Finally, let us comment on the relation between relative frequencies and probabilities in the context of entropy. Given a standard n-simplex, Δ_n, i.e. $(x_0, \ldots, x_n) \in \mathbb{R}^{n+1}$, $\sum_{i=0}^{n} x_i = 1$, and $x_i \geq 0$, for $i = 0, \ldots, n$, its points can either be interpreted as discrete probability distributions on $(n+1)$ elements or as the set of relative frequencies of $(n+1)$ elements. The distinction between the two concepts is relevant as the Shannon entropy H, provides in both cases a functional (observable) $H : \Delta_n \to \mathbb{R}_+$, which, in our context, has two possible interpretations. Namely, as a component of a coordinate system on (law) texts, which is the interpretation in the present study, but also as an estimate of the Shannon entropy of the language used if considered as a sample from the space of all (law) texts of a certain type. In the latter case, it is known that the "naive" estimation of the Shannon entropy **Eq. 1** from finite samples is biased. Therefore, several estimators have been developed to solve this problem. We utilize the entropy estimator introduced by [30]; in order to reexamine some of our results in the light of a probabilistic interpretation, and find that it has no qualitative effect on the outcome, cf. **Supplementary Material**.

4 THE COMPLEXITY-ENTROPY PLANE

Complex systems, e.g. biological, physical or social ones, are high-dimensional multi-scale objects. [31]; and [32] realized that in order to describe them, entropy is not enough, and an independent complexity measure is needed. Guided by the insight that the intuitive notion of complexity for patterns, when ordered by the degree of disorder, is at odds with its algorithmic description, the notion of the physical complexity of a system emerged, cf. [7, 31, 33]. The corresponding physical complexity measure, pioneered by [33]; should not be a monotone function of the disorder or the entropy, but should attain its maximum between complete order (perfect crystal) and total disorder (isolated ideal gas). [7]; introduced the excess Shannon entropy as a statistical complexity measure for physical systems, and later [34] introduced another physical complexity measure, the product of a system's entropy with its disequilibrium measure. [35]; introduced a novel approach to handle the complexity of patterns on multiple scales using a multi-level renormalization technique to quantify the complexity of a (two- or three-dimensional) pattern by a scalar quantity that should ultimately better fit the intuitive notion of complexity.

[7]; paired both the entropy and the physical complexity measure into what has become a complexity-entropy diagram, in order to describe non-linear dynamical systems; for a review cf. [36]. Remarkably, these low-dimensional coordinates are often sufficient to characterize such systems (in analogy to principal component analysis), since they capture the inherent randomness, but also the degree of organization. Several variants of entropy-complexity diagrams are now widely used, even outside the original context. [37]; by combining the normalized word entropy, cf. **Eq. 7**, with a version of a statistical complexity measure, quantitatively study Shakespeare and other English Renaissance authors. [23]; used for the complexity-entropy plane the entropy rate and the entropy density and studied the organization of literary texts (Shakespeare, Abbott and Doyle) at different levels of the hierarchy. In order to calculate the entropy rate and density, which are asymptotic quantities, they used the Lempel-Ziv

complexity C_{76}. Strictly speaking this approach would require the source to be stationary and ergodic, cf. [11].

We introduce a new variant Γ of the complexity-entropy plane, spanned by the normalized word entropy and the compression factor, in order to study text data. So, every text t, can be represented by a point in Γ, via the map $t \mapsto (H_n(t), r(t))$, with coordinates H_n, the normalized Shannon entropy of the underlying vocabulary, and r, the compression factor. Let us note, that Γ is naturally a metric space, e.g. with the Euclidean metric, but other metrics may be more appropriate, depending on the particular question at hand.

5 THE NORM HIERARCHY AND BOUNDARIES OF NATURAL LANGUAGE

Let us now motivate some of our research questions from the perspective of Legal Theory.

[38] and his school introduced and formalized the notion of the "Stufenbau der Rechtsordnung",[2] which led to the concept of the hierarchy of norms. The hierarchy starts with the Constitution (often originating in a revolutionary charter written by the "pouvoir constituant"), which governs the creation of statutes or acts, which themselves govern the creation (by delegation) of regulations, administrative actions, and also the judiciary. At the national level these (abstract) concepts are taken into account, e.g. Guide de légistique [39]; when drafting positive law. This is valid for, e.g. Austria, France, Germany, Italy, Switzerland and the European Union, although strictly speaking, it does not have a formal Constitution. Every new piece of legislation has to fit the preexisting order, so at each level, the content outlined at an upper level, has to be made more precise, which leads to the supposed linguistic gradient of abstraction. A new phenomenon can be observed for regulations, namely that the legislature, or more precisely its drafting agencies, is being forced to abandon the realm of natural language and take an approach that is common to all scientific writing, namely the inclusion of images, figures and formulae. The purpose of figures, tables and formulae is not only the ability to succinctly visualize or summarize large amounts of abstract information, but most often it is the only mean to convey complex scientific information at all. As regulations increasingly leave the domain of jurisprudence, novel methods should be adopted. For example [2], advocated the inclusion of mathematical formulae in a statute if this statue contains a computation that is based on this formula. Ultimately, a natural scientific approach (including the writing style) to law would be beneficial, however, this might be at odds with the idea of law being intelligible to a wide audience.

Our hypothesis is that these functional differences between the levels of the hierarchy of legal norms should manifest themselves as differences in vocabulary entropy or in the compression factor.

[2]This could be translated with "hierarchy of the legal order" or "hierarchy of norms".

6 MATERIALS AND METHODS

6.1 Data

Our analysis is based on the valid (in effect) and online available national codes from Canada, Germany, France, Switzerland, the United States, Great Britain and Shakespeare's collected works, for a summary statistics, cf. **Table 1**. We also included the online available constitutions of Canada, Germany, and Switzerland in the analysis, cf. **Table 2**. In addition, we use the online available German EuroParl corpus from [40] and its aligned English and French translations (proceedings of the European Parliament from 1996 to 2006) to measure language-specific effects for German, English and French.

In detail, we use all Consolidated Canadian Acts and Regulations in English and French (2020); all Federal German acts (Gesetze) and Federal regulations (Verordnungen) in German (2020); all French Codes (en vigueur) (2020); all Swiss Internal Laws (Acts and Ordinances) which have been translated into English, containing the following areas: 1 State - People - Authorities; 2 Private law - Administration of civil justice - Enforcement; Criminal law - Administration of criminal justice - Execution of sentences; 4 Education - Science - Culture; 5 National defense; 6 Finance; 7 Public works - Energy - Transport; 8 Health - Employment - Social security; 9 Economy - Technical cooperation (2020); the United Kingdom Public General Acts (partial dataset 1801–1987 and complete dataset 1988–2020); U.S. Code Titles 1–54 (Title 53 is reserved, including the appendices) (2020); U.S. Code of Federal Regulations for (2000) and (2019).

The collected works of Shakespeare are obtained from "The Folger Shakespeare - Complete Set, June 2, 2020", https://shakespeare.folger.edu/download/

6.2 Pre-Processing

For our analysis we use Python 3.7. If available, we downloaded the bulk data as XML-files, from which we extracted the legal content (without any metadata), and saved it as a TXT-file, after removing multiple white spaces or line breaks. If no XML-files were available, we extracted the texts from the PDF versions, removed multiple white spaces or line breaks, and saved it as TXT-files.

6.3 Measuring Vocabulary Entropy

For an individual text t, let $V := V(t) := \{v_1, \ldots, v_{|V|}\}$, be the underlying vocabulary, and $|V|$ the size of V. Let f_i be the frequency (total number of occurrences) of a unique word $v_i \in t$, and let $|t|$ be the total number of words in t (with repetitions), i.e. $|t| = \sum_{i=1}^{|V|} f_i$. The relative frequency is given by $\widehat{p}_i := f_i/|t|$, which can also be interpreted as the empirical probability distribution $\widehat{p}_i$. The word entropy $H(t)$ of a text t (but cf. **Section 3**), is then given by

$$H(t) := -\sum_{i=1}^{|V|} \widehat{p}_i \log_2(\widehat{p}_i), \tag{7}$$

and correspondingly, the normalized word entropy $H_n(t)$, cf. **Eq. 2**. Let us remark, that the word entropy is invariant under permutation of the words in a sentence.

TABLE 1 | Summary statistics on acts, regulations and English literature showing the language used and size (in MB) of the respective corpora, the number of items, the mean size (in KB) and the standard deviation.

Corpus (language)	Size [MB]	# Texts	Mean (size) [KB]	Std (size)
CA acts (EN)	52.4	823	63.6	254.7
CA reg. (EN)	55.6	3,725	14.9	59.7
CA acts (FR)	56.9	833	64.6	264.5
CA reg. (FR)	62.4	3,718	15.9	64.5
F codes (FR)	127.6	74	1664.0	2275.8
D acts (DE)	53.6	1,306	40.3	108.3
D reg. (DE)	69.6	3,316	20.6	61.5
United Kingdom PGA (EN)	269.5	3,512	76.3	192.7
USC 1–54 (2020) (EN)	139.6	57	2442.6	3835.6
U.S. CFR (2000) (EN)	940.2	200	4701.9	8156.2
U.S. CFR (2019) (EN)	572.9	242	2360.9	1079.7
CH acts (EN)	7.0	103	343.2	286.6
CH reg. (EN)	6.3	118	53.4	58.3
Shakespeare (EN)	5.2	42	124.9	32.0

TABLE 2 | Summary statistics for the Constitutions of Canada (EN), Germany (DE), Switzerland (DE,EN,FR), showing the language used, the original size (in KB), the compression factor and the normalized vocabulary entropy (after cutoff at 150 K).

Corpus (language)	Size [KB]	Comp. Factor	n-voc. Entropy
CH constitution (DE)	156	3.74	0.79
CH constitution (EN)	157	3.88	0.77
CH constitution (FR)	172	3.80	0.77
Ca constitution (EN)	215	3.67	0.75
D Grundgesetz (DE)	180	3.57	0.79

First we read the individual TXT-files, then filter the punctuation or special characters out and then split the remaining text into a list of items. In order to account for prefixes in French, the splitting separates expressions which are written with an apostrophe into separate entities. However, we do not lowercase letters, lemmatize or stem the remaining text, nor do we consider any bi- or trigrams. Keeping the original case-sensitivity, allows us to capture some syntactic or semantic information. Then we determine the relative frequencies (empirical probability values) of all unique items, from which we calculate the normalized entropy values according to **Eq. 2**. We truncate each text file at 150,000 characters, and discard files which are smaller than the cutoff value. For the EuroParl corpus we sampled 400 strings, consisting of 150 K characters each (with a gap of 300 K characters between consecutive strings) from the English, German and French texts, in order to calculate the corresponding normalized vocabulary entropy.

6.4 Measuring Compression Factors Using Gzip

In order to compute the compression factor as our derived complexity measure, we use as lossless compressor gzip.[3] After reading the individual TXT-files as strings, we compress them using Python's gzip compression module, with the compression level set to its maximum value (= 9). The individual compression factors are calculated according to **Eq. 6**. After analyzing all of our data, we choose 150,000 characters as the cutoff in order to minimize the effects of the overhead generated by the compression algorithm for very small text sizes. For the EuroParl corpora (English, French, German), we calculated the compression factors based on 400 samples each, as described above. Note that in the future it might make sense to also consider other (e.g. language specific) lossless compression algorithms in order to deal with short strings.

[3]Note that we do not consider quantities in the limit or issues like the convergence of entropy estimates.

7 RESULTS

Our first analysis, cf. **Table 1**, is a summary of the sizes of the different corpora, the languages used, the number of individual items, the mean text sizes and standard deviations. The analysis shows different approaches to the organization of national law, namely either by thousands of small texts of around 50 KB (Canada, Germany, United Kingdom) or less than a hundred large codes, several MB in size (France, United States), with the regulations significantly exceeding the number of acts. Note that the French codes contain both the law and the corresponding regulation in the same text. The size of a corpus within the same category, i.e. act or regulation, differs from country to country by an order of magnitude or even two, which is noteworthy as broadly similar or even identical areas are regulated within the law, e.g. banking, criminal, finance or tax law. This begs the question of what an efficient codification should ideally look like. The Swiss Federal codification is remarkably compact, despite the fact that the English version does not contain all acts or regulations available in German, French or Italian (which are the official languages); nevertheless all important and recent ones are included, cf. **Section 6.1**.

7.1 Normalized Entropy and Compression Factor

The normalized vocabulary entropies per corpus, cf. **Table 3**, have a standard deviation of approximately 0.01, and average entropy values that are distributed as follows: English in $[0.73, 0.80]$, German in

TABLE 3 | Summary statistics on acts, regulations (reg.) and English literature.#

Corpus	# Texts	Mean (cfc.)	Std. (cfc.)	Mean (nve.)	Std. (nve.)
CA acts (EN)	75	5.00	0.94	0.73	0.01
CA reg. (EN)	54	5.23	1.18	0.73	0.02
CA acts (FR)	74	4.64	0.93	0.75	0.01
CA reg. (FR)	60	4.98	1.24	0.74	0.02
F codes	58	4.10	0.28	0.76	0.01
D acts	78	4.12	0.42	0.78	0.01
D reg.	69	4.28	1.06	0.79	0.01
United Kingdom PGA	431	4.68	0.44	0.74	0.01
U.S. Codes (2020)	49	4.11	0.29	0.74	0.01
U.S. CFR (2000)	200	4.04	0.72	0.77	0.02
U.S. CFR (2019)	241	4.16	1.06	0.78	0.02
CH fed. acts (EN)	4	3.75	0.14	0.76	0.00
CH fed. reg. (EN)	5	4.00	0.23	0.77	0.01
EuroParl (DE)	—	2.95	0.05	0.81	0.00
EuroParl (EN)	—	3.02	0.05	0.77	0.00
EuroParl (FR)	—	3.06	0.06	0.77	0.00
Shakespeare	10	2.52	0.03	0.80	0.00

notes: cfac = compression factor; nve. = normalized vocabulary entropy; # texts = number of texts considered at 150 K.

$[0.78, 0.81]$ and French in $[0.74, 0.77]$. The analysis of the mean compression factors, based on the individual texts truncated at 150 K, reveals three regions where the values accumulate, cf. **Table 3**. So, Shakespeare's works have a mean compression factor of 2.52 (std = 0.03), the EuroParl corpora in English, French and German of around 3.01 (std = 0.06 approximately), whereas the national codifications are located in the interval $[3.75, 5.23]$, with the standard deviations being in the interval $[0.14, 1.24]$. On average, all national acts have a lower compression factor and a lower standard deviation than the corresponding national regulations. The (French), German, Swiss and United States acts are in the sub-interval $[3.75, 4.12]$, and the respective regulations in $[4.00, 4.28]$, but with a large standard deviation (1.06), for Germany and the United States. Based on the mean compression factor, the variance, the number of acts and the total size of the corpus, the French and the US codes are most similar. The acts of Canada (English and French) and of the United Kingdom are located at the upper end of the interval, namely in $[4.68, 5.0]$, as are the Canadian regulations with 4.98 and 5.23, for French and English, respectively. The values for the constitutions can be found in the interval $[3.57, 3.88]$ (compression factors), and $[0.75, 0.79]$ (normalized vocabulary entropy). The value of the compression factor of the Canadian and German Constitution is smaller than the corresponding mean value of the acts or regulations, but larger than that of EuroParl (DE, EN, FR) or Shakespeare. In the case of the Swiss Federal Constitution and its aligned translations into English, French and German, the compression factor is significantly higher than the corresponding EuroParl average values, but between the mean of the acts (EN) and the mean of the regulations (EN), cf. **Tables 2, 3**.

7.2 Complexity-Entropy Plane

The general picture of all texts analyzed in this study, cf. **Figure 1**, reveals, that the literary works of Shakespeare occupy a region to the left and are well separated from all the other data points. The three points corresponding to the English, French and German EuroParl samples are also well separated from the vast majority of legal texts and Shakespeare's collected works. This indicates that legal texts are much more redundant than classic literary texts or parliamentary speeches. The picture for the constitutions is heterogeneous for the data considered.

The German (DE) and Canadian (EN) Constitution are located on the left border of the region, which contains the respective national acts and ordinances, while the Swiss Federal Constitution lies between the averages of the acts and ordinances, but is much closer to the mean of the acts.

The plot for U.S. Code (USC), Titles 1–54 for the year 2020, and U.S. Code of Federal Regulations (CFR) for the years 2000 and 2019, cf. **Figure 2**, shows that the Federal acts occupy a distinguishable region which is located below the domain populated by the Federal regulations. This is in line with the values from **Table 3**, as the mean vocabulary entropy for USC is 0.74, as compared to 0.77, for CFR 2000, and 0.78, for CFR 2019. On the other hand, the distribution pattern of the regulations in 2000 and 2019 is similar (small changes in the region around the means), but several points are more spread out in the 2019 data, which is in line with the larger standard deviation of 1.06 in 2019 vs. 0.72 in 2000. However, the overall size of CFR 2000 is 940 MB, vs. 572,9 MB, for CFR 2019, which is a quite substantial difference.

We have already noted the similarity of the U.S. Titles and the French Codes. As **Figure 3** shows, the French Codes (in French), German Federal acts (in German) and the U.S. Titles (in English) are situated in the complexity-entropy plane, almost as vertical, non-overlapping, translations of each other, with the German acts being highest up. The order of the average normalized vocabulary entropies appears to be language specific, although in this case we are not considering (aligned) translations, cf. **Section 7.3**.

The picture for the aligned translations of the Canadian acts and regulations into English and French, cf. **Figure 4**, reveals that the acts are located, depending on the language, in separated regions which are bounded by ellipses of the same size around the respective means. For both English and French, the regulations are more dispersed than the acts (in particular the French) and the regulations in French are more widespread than those in

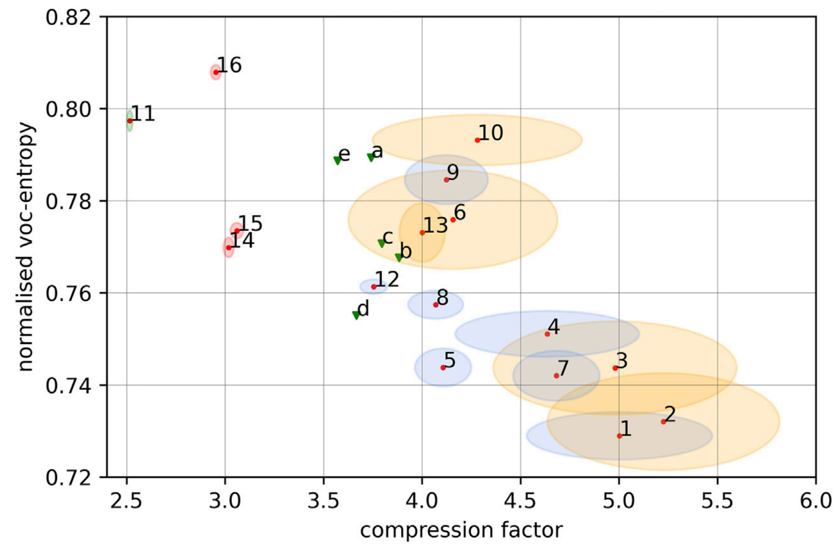


FIGURE 1 | Figure showing the mean compression factor and mean normalized vocabulary entropy for: 1 = Canadian acts (EN), 2 = Canadian regulations (EN), 3 = Canadian regulations (FR), 4 = Canadian acts (FR), 5 = U.S. Code Titles 1–54, 6 = U.S. CFR 2019, 7 = United Kingdom acts, 8 = French acts (FR), 9 = German Federal acts (DE), 10 = German Federal regulations (DE), 11 = Shakespeare's collected works, 12 = Swiss Federal acts (EN), 13 = Swiss Federal regulations (EN) 14 = EuroParl speeches (EN), 15 = EuroParl speeches (FR), 16 = EuroParl speeches (DE); and the compression factor and normalized vocabulary entropy (green marker) for: a = Swiss Federal Constitution (DE), b = Swiss Federal Constitution (EN), c = Swiss Federal Constitution (FR), d = Canadian Constitution (EN), e = German Constitution (Grundgesetz) (DE). The ellipses are centered around the mean values and have half-axes corresponding to $\sigma/2$ of the standard deviation of the compression factor and the normalized vocabulary entropy, respectively. Colors of ellipses correspond to: red = speeches (EuroParl), green = literature (Shakespeare), light blue = acts, orange = regulations; all texts truncated at 150 K.

English. The mean normalized entropy of the regulations in French is below the mean of the acts in French, but above the mean of the acts and regulations in English. The slightly odd position of the regulations in French could be due to the fact that after being truncated at 150 K, 60 (FR) vs. 54 (EN) regulations remain, while for the acts the number of texts remaining is the same. As we are dealing with aligned translations, the observed language specific pattern is quite meaningful, cf. **Section 7.3**. On

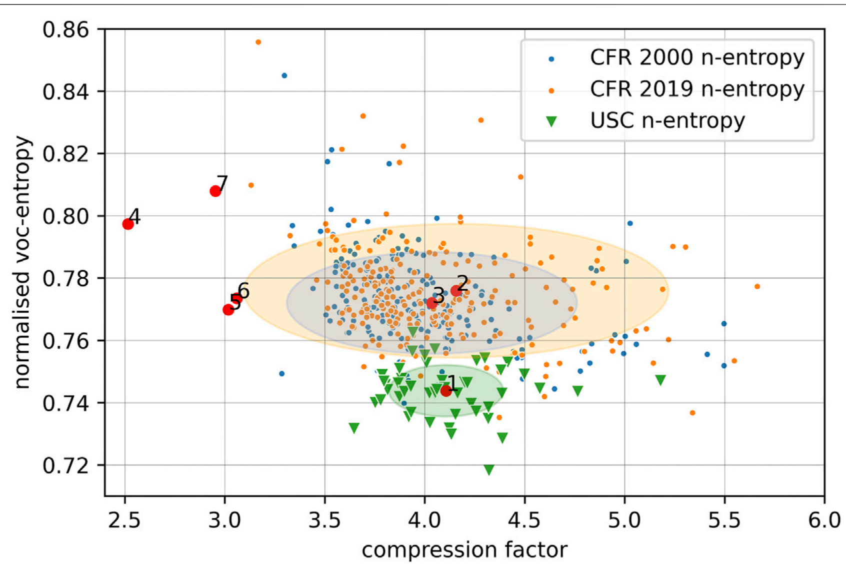


FIGURE 2 | Figure showing the mean compression factor and mean normalized vocabulary entropy for: 1 = U.S. Code Titles 1–54, 2 = U.S. CFR 2019, 3 = U.S. CFR 2000, 4 = Shakespeare's collected works, 5 = EuroParl speeches (EN), 6 = EuroParl speeches (FR), 7 = EuroParl speeches (DE). The ellipses are centered around the mean values and have axes corresponding to 1σ of the standard deviation of the compression factor and the normalized vocabulary entropy, respectively. Colors of ellipses correspond to: green = U.S. Federal acts (2020), orange = U.S. Federal regulations (2019), light blue = U.S. Federal regulations (2000); all texts truncated at 150 K.

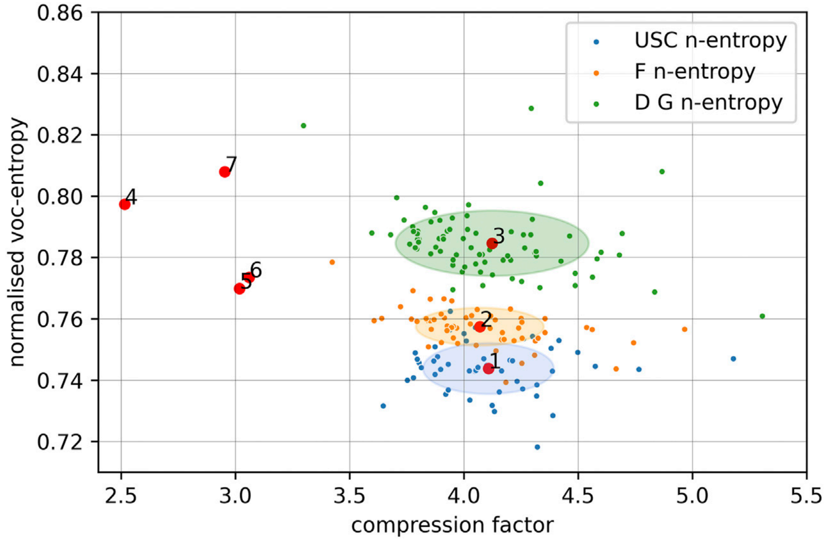


FIGURE 3 | Figure showing the mean compression factor and mean normalized vocabulary entropy for: 1 = U.S. Code Titles 1–54, 2 = French Codes (FR), 3 = German Federal acts (DE), 4 = Shakespeare's collected works, 5 = EuroParl speeches (EN), 6 = EuroParl speeches (FR), 7 = EuroParl speeches (DE). The ellipses are centered around the mean values and have axes corresponding to 1σ of the standard deviation of the compression factor and the normalized vocabulary entropy, respectively. Colors of ellipses correspond to: light blue = U.S. Code (2020), orange = French Codes, green = German Federal acts; all texts truncated at 150 K.

the other hand, Canadian acts and regulations in the same language are not easily separable, i.e. they show a distribution pattern that differs from the U.S. Titles and U.S. Federal regulations, cf. **Figure 2**.

The German Federal acts and regulations accumulate in nearby and overlapping areas of the plane, and cannot be clearly separated from each other, with the laws being more compactly grouped around the mean. The acts of Canada (EN), the United States and the United Kingdom are close to each other, but far below the German acts and regulations, cf. **Figure 5**. Indeed, this seems to reflect language-specific characteristics common to all genres.

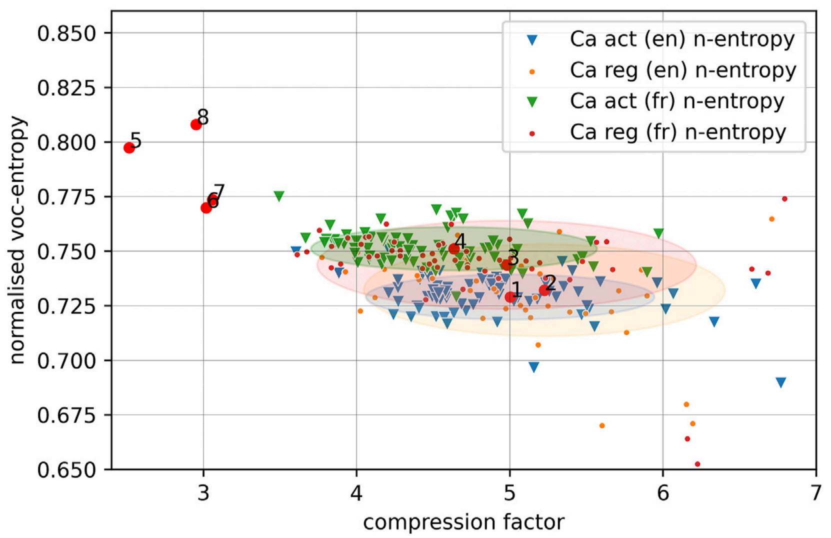


FIGURE 4 | Figure showing the mean compression factor and mean normalized vocabulary entropy for: 1 = Canadian acts (EN), 2 = Canadian regulations (EN), 3 = Canadian regulations (FR), 4 = Canadian acts (FR), 5 = Shakespeare's collected works, 6 = EuroParl speeches (EN), 7 = EuroParl speeches (FR), 8 = EuroParl speeches (DE). The ellipses are centered around the mean values and have axes corresponding to 1σ of the standard deviation of the compression factor and the normalized vocabulary entropy, respectively. Colors of ellipses correspond to: light blue = Canadian acts (EN), orange = Canadian regulations (EN), green = Canadian acts (FR), red = Canadian regulations (FR); all texts truncated at 150 K.

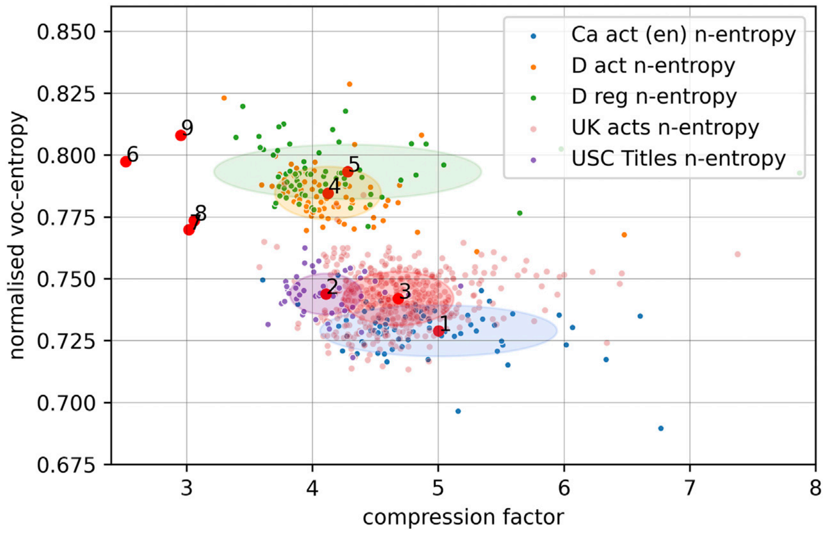


FIGURE 5 | Figure showing the mean compression factor and mean normalized vocabulary entropy for: 1 = Canadian acts (EN), 2 = U.S. Code, Titles 1–54 (USC), 3 = United Kingdom General Public Acts (PGA), 4 = German Federal acts (DE), 5 = German Federal regulations (DE), 6 = Shakespeare's collected works, 7 = EuroParl speeches (EN), 8 = EuroParl speeches (FR), 9 = EuroParl speeches (DE). The ellipses are centered around the mean values and have axes corresponding to 1σ of the standard deviation of the compression factor and the normalized vocabulary entropy, respectively. Colors of ellipses correspond to: light blue = Canadian acts (EN), orange = German Federal acts (DE), green = German Federal regulations (DE), red = United Kingdom PGA, purple = USC; all texts truncated at 150 K.

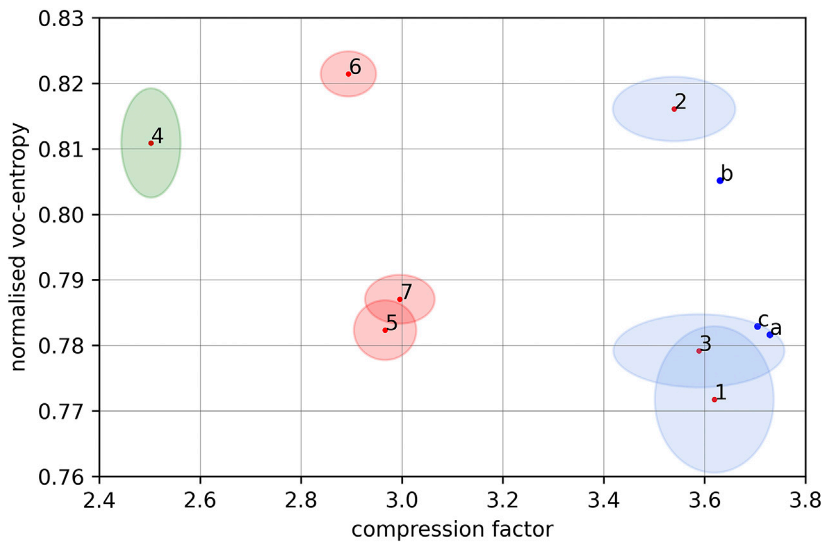


FIGURE 6 | Figure showing the mean compression factor and mean normalized vocabulary entropy for: 1 = Swiss Federal acts (EN), 2 = Swiss Federal acts (DE), 3 = Swiss Federal acts (FR), 4 = Shakespeare's collected works (EN), 5 = EuroParl speeches (EN), 6 = EuroParl speeches (DE), 7 = EuroParl speeches (FR), and the compression factor and normalized vocabulary entropy for: a = Swiss Federal Constitution (EN), b = Swiss Federal Constitution (DE), c = Swiss Federal Constitution (FR). The ellipses are centered around the mean values, and have axes corresponding to 1σ of the standard deviation of the compression factor and the normalized vocabulary entropy, respectively. Color code: red = EuroParl speeches, green = literature, light blue = acts; all texts truncated at 100 K.

The fact that the United States Code, unlike for Canada, Germany and Switzerland, is fairly well separated in the plane from its associated regulations could reflect differences in the way laws and regulations are drafted in the United States as compared to the countries mentioned above.

7.3 Distinguishing Different Languages

From the above discussion it can be seen that different languages can be distinguished by the normalized vocabulary entropy if the genre is kept constant. In order to further investigate the language effect on the position of the corpora in the complexity-entropy plane, we specifically considered aligned translations. So, additionally to the Swiss Federal Constitution (English, French and German), the German EuroParl corpus and its translation into English and French, we processed the nine largest Swiss Federal acts in English, French and German. However, in order to have enough Swiss Federal acts, we had to lower the cutoff to 100K, and correspondingly had to recalculate the EuroParl values. Additionally we added the collected works of Shakespeare (in English), with a cutoff of 100 K. Further, we have the Canadian acts and regulations, and their aligned translations into English and French. The results imply that (aligned) translations of the same collection of texts into different languages are primarily not distinguished by the compression factor but rather by the (normalized) vocabulary entropy, cf. **Figure 6** and **Figure 1**.

8 CONCLUSION

We introduced a tool that is new to the legal field but has already served other areas of scientific research well. Its main strength is the ability to simultaneously capture and visualize independent and fundamental information, namely entropy and complexity, of large collections of data, and to track changes over time. By devising a novel variant of the complexity-entropy plane, we were not only able to show that legal texts of different types and languages are located in distinguishable regions, but also to identify different drafting approaches with regard to laws and regulations. In addition, we have taken the first steps to follow the spatial evolution of the legislation over time. Although we observe that constitutions tend to have lower compression factors than acts and regulations, and regulations on average have higher compression factors than acts, which corresponds to the hierarchy of norms, we could not fully capture the assumed abstraction gradient. This suggests that other language-specific methods should also be used to investigate (possible) differences. On the other hand, the high(er) redundancy of the regulations reflects the increasing need to leave the realm of natural language and to borrow tools from the natural sciences. The analysis we perform can be modified in a number of ways to provide even more specific information. So, one might include *n*-grams, or perform additional pre-processing steps, or choose different compression algorithms. Also, one might add a third coordinate for even more visual information. In combination with other quantitative methods such as citation networks or the consideration of additional (internal) degrees of freedom such as local entropy, new types of quantitative research questions could be formulated, which may lead to more efficient and manageable legislation. In summary, we expect a broad range of further applications of complexity-entropy diagrams within the legal domain.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. The data can be found at: https://uscode.house.gov, https://www.govinfo.gov/bulkdata/CFR, https://www.legislation.gov.uk/ukpga, https://open.canada.ca/data/en/dataset/eb0dee21-9123-4d0d-b11d-0763fa1fb-403, https://www.fedlex.admin.ch/en/cc/internal-law/, https://www.gesetze-im-internet.de/aktuell.html, https://www.legifrance.gouv.fr/liste/code?etatTexte=VIGUEUR&page=1#code, https://www.statmt.org/europarl/.

AUTHOR CONTRIBUTIONS

RF contributed to the methods, analyzed the data and wrote the article.

FUNDING

The author received funding from the Max Planck Institute for the Physics of Complex Systems (MPIPKS).

ACKNOWLEDGMENTS

RF thanks Elliott Ash (ETH Zurich) and Holger Spamann (Harvard University) for numerous stimulating discussions. He thanks the condensed matter physics group at the Max Planck Institute for the Physics of Complex Systems in Dresden for its hospitality during his stay in 2020, and its support. Finally, he thanks the anonymous referees for their constructive comments and suggestions which helped to improve this article.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2021.671882/full#supplementary-material

REFERENCES

1. Schuck PH. Legal Complexity: Some Causes, Consequences, and Cures. *Duke L J* (1992) 42:1–52. doi:10.2307/1372753
2. Rook LW. Laying Down the Law: Canons for Drafting Complex Legislation. *Or L Rev* (1993) 72:663.
3. Mazzega P, Bourcier D, and Boulet R. *Proceedings of the 12th International Conference on Artificial Intelligence and Law*. New York, NY: Association for Computing Machinery) (2009). p. 236–7. The Network of French Legal Codes.
4. Bommarito MJ, and Katz DM. A Mathematical Approach to the Study of the united states Code. *Physica A: Stat Mech its Appl* (2010) 389:4195–200. doi:10.1016/j.physa.2010.05.057
5. Katz DM, and Bommarito MJ. Measuring the Complexity of the Law: the united states Code. *Artif Intell L* (2014) 22:337–74. doi:10.1007/s10506-014-9160-8
6. Bourcier D, and Mazzega P. Toward Measures of Complexity in Legal Systems. *Proceedings of the 11th International Conference on Artificial Intelligence and Law* (2007). p. 211–5.
7. Crutchfield JP, and Young K. Inferring Statistical Complexity. *Phys Rev Lett* (1989) 63:105–8. doi:10.1103/physrevlett.63.105

8. Ruhl J, and Katz DM. Measuring, Monitoring, and Managing Legal Complexity. *Iowa L Rev* (2015) 101:191–244.
9. Shannon CE. A Mathematical Theory of Communication. *Bell Syst Tech J* (1948) 27:379–423. doi:10.1002/j.1538-7305.1948.tb01338.x
10. Chang M-C, Yang AC-C, Eugene Stanley H, and Peng C-K. Measuring Information-Based Energy and Temperature of Literary Texts. *Physica A: Stat Mech its Appl* (2017) 468:783–9. doi:10.1016/j.physa.2016.11.106
11. Cover TM, and Thomas JA. *Elements of Information Theory (Wiley Series in Telecommunications and Signal Processing)*. USA: Wiley-Interscience (2006).
12. Li M, and Vitányi PM. *An Introduction to Kolmogorov Complexity and its Applications*. 4 edn. Incorporated: Springer Publishing Company (2019).
13. Lempel A, and Ziv J. On the Complexity of Finite Sequences. *IEEE Trans Inform Theor* (1976) 22:75–81. doi:10.1109/TIT.1976.1055501
14. Wolfram S. Computation Theory of Cellular Automata. *Commun.Math Phys* (1984) 96:15–57. doi:10.1007/bf01217347
15. Ziv J, and Lempel A. A Universal Algorithm for Sequential Data Compression. *IEEE Trans Inform Theor* (1977) 23:337–43. doi:10.1109/tit.1977.1055714
16. Ziv J, and Lempel A. Compression of Individual Sequences via Variable-Rate Coding. *IEEE Trans Inform Theor* (1978) 24:530–6. doi:10.1109/tit.1978.1055934
17. Hansel G, Perrin D, and Simon I. Compression and Entropy. *Annual Symposium on Theoretical Aspects of Computer Science*. Berlin, Heidelberg: Springer (1992).
18. Salomon D. *Data Compression: The Complete Reference*. 4 edn. London: Springer-Verlag (2007). doi:10.1007/978-1-84628-959-0
19. Chaitin GJ. Algorithmic Information Theory. *IBM J Res Dev* (1977) 21:350–9. doi:10.1147/rd.214.0350
20. Martiniani S, Lemberg Y, Chaikin PM, and Levine D. Correlation Lengths in the Language of Computable Information. *Phys Rev Lett* (2020) 125:170601. doi:10.1103/physrevlett.125.170601
21. Montemurro MA, and Zanette DH. Universal Entropy of Word Ordering across Linguistic Families. *PLoS ONE* (2011) 6:e19875. doi:10.1371/journal.pone.0019875
22. Friedrich R, Luzzatto M, and Ash E. *Entropy in Legal Language*. 2nd Workshop on Natural Legal Language Processing (NLLP, Collocated with KDD 2020) CEUR Workshop Proceedings. Vol. 2645, (virtual) (2020). Available at: http://ceur-ws.org/Vol-2645/.
23. Estevez-Rams E, Mesa-Rodriguez A, and Estevez-Moya D. Complexity-entropy Analysis at Different Levels of Organisation in Written Language. *PloS one* (2019) 14:e0214863. doi:10.1371/journal.pone.0214863
24. Grunwald P, and Vitányi P. Shannon Information and Kolmogorov Complexity. arXiv preprint cs/0410002. (2004).
25. Jurafsky D, and Martin JH. *Speech and Language Processing*. 2nd ed. USA: Prentice-Hall (2009). doi:10.1109/asru.2009.5373494
26. Shannon CE. Prediction and Entropy of Printed English. *Bell Syst Tech J* (1951) 30:50–64. doi:10.1002/j.1538-7305.1951.tb01366.x
27. Debowski Ł. Is Natural Language a Perigraphic Process? the Theorem about Facts and Words Revisited. *Entropy* (2018) 20:85.
28. Schürmann T, and Grassberger P. Entropy Estimation of Symbol Sequences. *Chaos* (1996) 6:414–27. doi:10.1063/1.166191
29. DeGiuli E. Random Language Model. *Phys Rev Lett* (2019) 122:128301. doi:10.1103/physrevlett.122.128301
30. Grassberger P. *Entropy Estimates from Insufficient Samplings* (2003) (arXiv preprint physics/*0307138*.
31. Grassberger P. Toward a Quantitative Theory of Self-Generated Complexity. *Int J Theor Phys* (1986) 25:907–38. doi:10.1007/bf00668821
32. Crutchfield JP. Between Order and Chaos. *Nat Phys* (2012) 8:17–24. doi:10.1038/nphys2190
33. Huberman BA, and Hogg T. Complexity and Adaptation. *Physica D: Nonlinear Phenomena* (1986) 22:376–84. doi:10.1016/0167-2789(86)90308-1
34. López-Ruiz R, Mancini HL, and Calbet X. A Statistical Measure of Complexity. *Phys Lett A* (1995) 209:321–6. doi:10.1016/0375-9601(95)00867-5
35. Bagrov AA, Iakovlev IA, Iliasov AA, Katsnelson MI, and Mazurenko VV. Multiscale Structural Complexity of Natural Patterns. *Proc Natl Acad Sci USA* (2020) 117:30241–51. doi:10.1073/pnas.2004976117
36. Feldman DP, McTague CS, and Crutchfield JP. The Organization of Intrinsic Computation: Complexity-Entropy Diagrams and the Diversity of Natural Information Processing. *Chaos: Interdiscip J Nonlinear Sci* (2008) 18:043106. doi:10.1063/1.2991106
37. Rosso OA, Craig H, and Moscato P. Shakespeare and Other English Renaissance Authors as Characterized by Information Theory Complexity Quantifiers. *Physica A: Stat Mech its Appl* (2009) 388:916–26. doi:10.1016/j.physa.2008.11.018
38. Kelsen H. *Reine Rechtslehre: Mit einem Anhang: Das Problem der Gerechtigkeit*. Tübingen: Mohr Siebeck (2017). doi:10.33196/9783704683991
39. General S. *Guide de légistique (3 édition mise à jour 2017)*. 3 edn. France: La documentation Française (2017).
40. Koehn P. *Conference Proceedings: The Tenth Machine Translation Summit. AAMT*. Phuket, Thailand: AAMT (2005). p. 79–86. Europarl: A Parallel Corpus for Statistical Machine Translation.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 09 June 2021
doi: 10.3389/fphy.2021.667471

Conspiracy of Corporate Networks in Corruption Scandals

J. R. Nicolás-Carlock and I. Luna-Pla*

Institute of Legal Research, National Autonomous University of Mexico, Mexico City, Mexico

Corruption in public procurement transforms state institutions into private entities where public resources get diverted for the benefit of a few. On this matter, much of the discussion centers on the legal fulfillment of the procurement process, while there are fewer formal analyses related to the corporate features which are most likely to signal organized crime and corruption. The lack of systematic evidence on this subject has the potential to bias our understanding of corruption, making it overly focused on the public sector. Nevertheless, corruption scandals worldwide tell of the importance of taking a better look at the misuse and abuse of corporations for corrupt purposes. In this context, the research presented here seeks to contribute to the understanding of the criminal conspiracy of companies involved in public procurement corruption scandals under a network and complexity science perspective. To that end, we make use of a unique dataset of the corporate ownership and management information of four important and recently documented cases of corruption in Mexico, where hundreds of companies were used to embezzle billions of dollars. Under a bipartite network approach, we explore the relations between companies and their personnel (shareholders, legal representatives, administrators, and commissioners) in order to characterize their static and dynamic networked structure. In terms of organized crime and using different network properties, we describe how these companies connect with each other due to the existence of shared personnel with role multiplicity, leading to very different conspiracy networks. To best quantify this behavior, we introduce a heuristic network-based conspiracy indicator that together with other network metrics describes the differences and similarities among the networks associated with each corruption case. Finally, we discuss some public policy elements that might be needed to be considered in anti-corruption efforts related to corporate organized crime.

Keywords: corruption, conspiracy, corruption networks, social network analysis (SNA), complex networks, complex systems, social physics, legal studies

Edited by:
Pierpaolo Vivo,
King's College London,
United Kingdom

Reviewed by:
Haroldo V. Ribeiro,
State University of Maringá, Brazil
Mihaly Fazekas,
Central European University, Hungary

****Correspondence:***
J. R. Nicolás-Carlock
jnicolas@unam.mx

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *13 February 2021*
Accepted: *14 May 2021*
Published: *09 June 2021*

Citation:
Nicolás-Carlock JR and Luna-Pla I (2021) Conspiracy of Corporate Networks in Corruption Scandals. Front. Phys. 9:667471. doi: 10.3389/fphy.2021.667471

1 INTRODUCTION

Corruption is a complex adaptive problem that threatens the integrity of modern societies. According to the United Nations, corruption is a transnational phenomenon that affects all societies in deep and multiple ways, at their political, economic, ecological, and social fronts [1, 2]. This is especially harmful in developing countries where it links to other forms of crime, such as organized crime [3, 4], economic crime, and where, on top of that, delegation of authority takes place [5].

For many years, academic researchers have provided insight into the corruption phenomenon through multidisciplinary approaches [6, 7]. Under these approaches, corruption is broken down into various types of individual or collective behaviors within several processes that affect private and public sectors likewise. Notably, the procurement of goods, services, and works is one of the governments' activities most vulnerable to corruption, with complex processes and high level of interactions between public officials and private businesses [8]. Here, corruption transforms public state institutions into private entities where public resources get diverted for the benefit of a few. Many legal corruption types are common at all stages of the procurement process, such as embezzlement, bribery, conflict of interest, fraud, conspiracy, and the use of shell companies to alter competition, among others [3, 9]. For those reasons, governments lose a significant amount of funds in corrupt practices, and thus, purchasing is considered a high-risk activity [4, 9].

The vast academic literature on government procurement and auctions is continuously developing sophisticated analyses to tackle fraud schemes in all stages of the process, while there are fewer formal analyses related to linking corporate criminal activities to the ownership and management features and collective behaviors that are most likely to signal corruption regardless of the procurement method [10–12]. The lack of systematic evidence based on real activities on this subject has the potential to bias our understanding of corruption, making it overly focused on the public sector, ignoring the structural dynamics that take place in the ecosystem of private entities (bidders) [13]. This issue is mainly due to the poor access to widely reliable data on the private sector, specifically the management and ownership of companies. On the one hand, the abuse of private (shell) companies is a well-known phenomenon that is well documented in few cases, but most of the time, it requires extensive research on government records, long waiting periods after freedom of information act requests, and handcrafted processing of information afterward. On the other hand, companies involved or misused for corrupt purposes tend to hide or misinform about their management and ownership structure either by being registered off-shore or by using straw men. In the case of the latter, the multiplicity of personnel roles (shareholders, managers, or legal representatives) within or among companies creates unusual corporate profiles that are still far unexplored and that might provide further insight into potential criminal activity. Therefore, the analysis of corruption related to private companies, especially in developing countries' context, is a relevant aspect at the forefront of international anti-corruption efforts that requires formal and extensive research [13].

Recently, data-driven studies based on complex systems and network theory approaches have delved into the description of the network characteristics of real corporations at the level of board members [14] as well as those that have participated in big corruption scandals [15]. Under a complex systems perspective, corruption can be seen as an adaptive phenomenon that is best understood in terms of the collective behavior of interrelated agents acting as a whole; that is, it could be hypothesized that this phenomenon is dominated by network effects and thus best characterized by unique collective properties that could provide more insight into corruption, organized crime, and cheating cases than by simply looking at isolated actors or events [12, 15–19]. Under this systemic perspective, the players are considered as connected actors on a network, regardless of their underlying interests or motivations, and the goal is to analyze the emerging structure, dynamics, organization, *modus operandi*, and the role that the actors play in practice as a collective. Linking actors by contracts, labor, or social relations may depict an organized network working for a common goal (win elections), multiple goals (private gain), or yielding criminal conspiracy (agreement to commit illegal acts) between companies bidding or having direct contracts while they intend to remain undetected and camouflage their activity. In public procurement, it is known that networked corruption schemes limit competition and affect the quality and efficacy of services, goods, and activities financed by governments [8, 10, 17]. However, to this day, there is no comprehensive understanding and convincing systematic evidence about the main structural and dynamical features of networked companies that might be able to define signs of corruption in organized crime activity despite the purchasing method.

In this context, the research presented here seeks to contribute to the understanding of the criminal conspiracy of companies involved in public procurement scandals under a network and complexity science perspective. In particular, due to the quality and scope of the data used, we focused on the collective features of the companies at the ownership and management levels in order to describe and identify structural and dynamical patterns that might be able to provide further insight into the corrupt schemes and nature of the cases. As stated above, the multiplicity of personnel roles within or among companies creates unusual corporate profiles that are still far unexplored and that might provide further insight into potential criminal activity. To that end, we make use of a unique microlevel corporate ownership and management dataset of four important and recently documented cases of corruption in Mexico, where hundreds of companies were used to embezzle billions of dollars. Under a bipartite network approach, we explore the relations between companies and their personnel (shareholders, legal representatives, administrators, and commissioners) in order to depict their network structure. Using diverse network metrics, we describe how these companies are not independent but connected with each other due to the existence of shared personnel with role multiplicity. To quantify this behavior, we introduce a heuristic network-based conspiracy indicator that together with other network metrics describes the differences and similarities among the networks associated with each corruption case. Finally, we discuss the advantages and shortcomings of data and network-based approaches, open research, and public policy elements that might need to be considered in anti-corruption efforts related to corporate organized crime.

1.1 Data and Methods

Important remark: The judicial and criminal investigation processes of some of the cases presented in this study are pending; therefore, the present analysis is only valid for academic purposes. For privacy protection, we have held in anonymity all the information regarding the identification of the companies and associated people. The data analyzed during the current study are available online (see Data Availability).

1.1.1 Data

The data used in this study were first gathered by investigative journalists and members of a local citizen participation committee, and it is related to four corruption scandals in public procurement involving three local administrations and the national government in Mexico: Veracruz (Case 1), Puebla (Case 2), Guanajuato (Case 3), and the Federation (Case 4), respectively. A short description of each case is presented below.

In order to have quality datasets that would allow for consistent comparisons among cases, we manually curated and double-verified the original data by using fiscal and official procurement records (when available). In some cases, the information regarding the tendering process and awarded contracts was missing or incomplete. Consequently, having consistent and verified data at the level of contracting bodies, suppliers, or bidders for all scandals was not possible. For this reason, the commonly employed analysis of the procurement environment as a network that considers the connections between the contracting institutions and suppliers, either under a bipartite institution–supplier network or a firm–firm co-bidding network [8, 17, 20], was not possible. However, the information regarding the connections of public servants and business people to their corresponding institutions or companies, respectively (beyond hidden political ties), is an aspect that can be further explored. In particular, the multiplicity of personnel roles within or among companies creates unusual corporate profiles and networks at the level of the supplier's ecosystem that are still far unexplored and that might provide further insight into potential criminal activity. Therefore, we focused on the collective features of the companies at the ownership and management levels in order to describe and identify structural and dynamical patterns that might be able to provide further insight into the corrupt schemes and nature of the cases.

After curating and cross-validating the records, we came up with datasets that contain information regarding the four roles that comprise the ownership and management of the firms as they appear in the companies' charters and contracts: shareholders (SHs), administrators (ADMs), legal representatives (LRs), and commissioners (COMs). The date of creation of the companies as well as identifier of whether the company has been classified as shell by fiscal authorities are also provided (more details in the data repository; see Data Availability).

The cases presented in this research fall into a certain type of procurement practices within an environment that deems common in Mexico. Since 2013, the tendency in federal public spending has increasingly turned over the direct contracting and restricted invitations, in substitute to the bidding procedures, to the point that by 2020, over 80 percent of the federal government contracts were directly awarded to companies [21]. This tendency replicates within the states, where heterogeneous legislation often allows budget thresholds to be exceeded, requires fewer requisites for tenders, and not all state governments make calcification and on-site visits [22]. As a consequence of this high-risk corruption context, market concentration and the creation of short-term companies that can offer multiple products and services prevail in the county. When federal or state public officials and private actors bypass legislation to obtain private gain (a corruption or criminal goal), they use companies (existing and new) to divert resources. Many of those companies become listed as shell given tax offenses within the course of three years [23]. Specifically, within the four cases selected for this study, nonstandard practices are often in place, such as bid-rigging, communication between companies during tender process, increased input costs and contract prices, and high degree of discretion of public servants in the application of the public procurement regulatory framework.

These four cases are selected for the following reasons: data availability, the governments assigned contracts to a set of companies that later were prosecuted by fiscal authorities and finally enlisted as shells, and these cases portrayed a great number of companies and personnel involved. Corruption in the cases presumably happened when companies failed to comply with the contracts, diverted public goods, product substitution, and/or simulated operations under Mexican fiscal legislation in order to commit embezzlement, fraud, and money laundering. The data used in the study imply certain limitations. Since most of the information is public or was made public after freedom of information requests, it is possible that governments answered in a restrained fashion, discarding documents and contracts that also relate to the network's or cartels' activities. Also, the contracts analyzed for the study contain straw men, criminal impersonation, and inauthentic companies' address, due to the criminal activities undertaken. And finally, the contract samples are not the result of statistical methods, but a selection of corruption media scandals set by a journalistic narrative; therefore, criminal files may or may not correspond to the grouping of participants. Although the data of the cases do not represent a large-scale sample of the procurement environment, other studies have used media corruption scandals based on government information to describe network structure and dynamics over the time [16, 24]. Hence, the results are not applicable to all procurement environments, but only to procurement and direct contracts activities that have been flagged by the fiscal and prosecution authorities, and where information is publicly available after corruption activity or the use of shell companies has been detected.

A shell company is commonly defined as an entity legally incorporated that, structurally speaking, lacks substantial assets, operations, or even employees, and when used for illicit purposes, the entity is typically oriented to the concealment of beneficial ownership. Likewise, it is characterized by the appointment of informal nominees, such as children, spouses, relatives, or associates who do not appear to be involved in the running of the corporate enterprise [25]. According to Mexican fiscal law, shell companies are categorized as such based on the government registry of tax evasion that is publicly available, when there is missing or in existent address, simulated operations, inauthentic documentation, lack of assets, or issue fiscal receipts to feign operations. Mexican fiscal authorities list them after an investigation that could take months or years after they were established.

Case 1-Veracruz: The data of this case were originally gathered from governmental sources (open to public access under Mexico's General Law of Transparency and Access to Information and local transparency legislation) by the nongovernmental organization known as Mexicanos Contra la Corrupción e Impunidad (Mexicans Against Corruption and Impunity) and the investigative journalism group known as Animal Político in 2016 [23]. Journalists also undertook field trips to verify shell companies' addresses. These companies participated in bidding procedures and direct award. At least four state government agencies contracted the group of companies and have been prosecuted for diverting public funds, amounting over 30 million of U.S. dollars between the years 2012 and 2014 [23]. Unlike the other three cases selected for this study, the procurement environment and legal framework in Veracruz favor an array of concentration practices. Although the end recipients of the funds are unclear, the journalist group has reasons to believe that resources were diverted to the political campaign for the national presidency held in 2012. This would explain the participation of a larger number of companies and collaborating under a meticulous design of fraud scheme for over 3 years. The corruption schemes under investigation by the fiscal authorities consisted in the use of straw men, other fiscal offenses, and tax fraud to divert public funds through tenders. In this scandal, there are 354 companies (96 of which are shell) and 306 personnel (or people) associated with those companies.

Case 2-Puebla: The contracts sampled for this case date from 2015 to 2018 where the state and municipal government of Puebla contracted services and goods from companies that were mostly listed as shell by the national fiscal authority. In total, the companies developed contracts over 17 million dollars. The research was made by journalism group *Datamos*, in coordination with the International Center for Journalists and Connectas [26]. The collection of data methods included filing of freedom of information requests under state legislation and from official websites available by legal transparency mandate (*Plataforma Nacional de Transparencia* and *Compranet*). In addition, journalists undertook field trips to verify shell companies' addresses and also interview several companies' stakeholders and public officials to validate data. The contracted companies were founded during the years 2012–2015, and three years later, 68 companies were listed as shell companies by the tax federal authority *Servicio de Administración Tributaria*. At least eleven state and municipal agencies contracted the companies from the sectors of education, infrastructure, health, security, and others. The corruption scheme under investigation by the fiscal authorities consisted in the use of straw men, other fiscal offenses, and tax fraud to divert public funds through tenders. In this scandal, there are 90 companies (87 of which are shell) and 230 personnel associated with those companies.

Case 3-Guanajuato: The data sample comprises the state and municipal governments of Guanajuato. The contracts were assigned to companies that have been listed as shell by the federal fiscal authority during the years 2014–2019. The source of this information is a special report of the Citizenship Committee of the State's Anti-corruption System, an official agency created by the anti-corruption state law with oversight powers. The committee filed numerous freedom of information requests and appeals to gather the data of the contracts with the state government, 4 municipalities, the state university, and other agencies. The report aimed to evaluate the state's situation regarding simulated operations and to track public funding diverted to shell companies for over 9 million U.S. dollars [27, 28]. In this case, the corruption scheme under investigation by the fiscal authorities also consisted in the use of straw men, other fiscal offenses, and tax fraud to divert public funds through tenders. There are 110 companies (101 of which are shell) and 511 personnel associated with those companies.

Case 4-Federation: The case was known in the media as the "Master Fraud" (*Estafa Maestra*) where federal and six state governments channeled public resources to regular and shell companies through state universities. The data were originally gathered from official sources and websites under Mexico's General Law of Transparency and Access to Information by the nongovernmental organization *Mexicanos Contra la Corrupción e Impunidad* (Mexicans Against Corruption and Impunity) and the investigative journalism group known as *Animal Político* [29]. The case sample was made after the general audit authority reviewed the contracts and identified possible fraud, embezzlement, money laundry, and misuse of public resources between the years 2013–2014 for over 380 million dollars. In this case, the government funds were diverted through contracts with public universities that hired shell companies to develop projects and activities, in which the general audit authority identified illegal contracts, fiscal offenses, and tax fraud in tenders. There are 120 companies (51 of which are shell) and 446 personnel associated with those companies.

2 METHODS

In this study, *criminal conspiracy* is seen as an organized crime act or as "an agreement between two or more people to commit an illegal act, along with an intent to achieve the agreement's goal" (e.g., see Whitfield v. United States, 453 U.S. 209 (2005)). Despite not having real evidence that the companies communicated previously or during the government purchases, we hypothesize that the emergence of networked structures due to the existence of shared personnel among companies represents a relevant proxy to the level of conspiracy of the actors involved; that is, if conspiracy is present, then it is encoded in the structure of the corporate networks.

We consider this a valid hypothesis in the context of the procurement corruption scandals studied here, where the corporate ownership and management information is not readily available or even not existent, and true ownership can be easily concealed. In this scenario, the abuse and misuse of small-size corporations is easily done by (although not restricted to), for example, the use of multiple straw men. Therefore, if conspiracy is encoded in networked structures of companies with shared personnel, network analysis will allow us to identify and quantify the static and dynamic network properties that might be able to signal organized crime and corruption [15, 16].

Under the previous considerations, our analysis is divided into three parts: I) an elemental statistical description of some quantities related to the companies and personnel per case, II) an analysis of the relationships among companies and individuals under a bipartite network analysis perspective, and III) an analysis based on a heuristic conspiracy indicator that together with other network metrics is able to classify and quantify different levels of conspiracy. Below, we present the elements considered for each part of the analysis.

Part I-Elemental corporate features: For each case, the exploratory analysis considers the following basic variables:

- Total number of companies, N_{CO}.
- Total number of personnel, N_{PE}.
- Number of companies created per year, $n_{CO,t}$.
- Number of new personnel added per year, $n_{PE,t}$.
- Number of personnel per role (SHs, ADMs, COMs, and LRs).
- Number of role pairs (these are the number of individuals that fall within the intersection of any role pairs).
- Number of personnel per number of roles performed.

Part II-Bipartite networks: Given the nature of the data, we considered an undirected weighted bipartite network approach in which companies and individuals represent two different sets of nodes that are connected by four types of edges (SHs, ADMs, COMs, and LRs), and where the weight of the edges is given by the number of roles an individual performs within a company. The network properties we considered are as follows [30]:

- Density, $\delta = L/L_{max}$, where L is the number of observed edges and L_{max} is the maximum number of edges for a network of N nodes. For a bipartite network, $L_{max} = N_1 N_2$; thus,

$$\delta = \frac{L}{N_1 N_2}, \tag{1}$$

where $N_1 = N_{CO}$ and $N_1 = N_{PE}$.

- Mean degree, $\langle k \rangle$:

$$\langle k \rangle = \frac{1}{N}\sum_{i=1}^{N} k_i, \tag{2}$$

where k_i is the degree of the ith node and N is the number of nodes of the corresponding set.

- Mean weighted degree or *strength*, $\langle s \rangle$:

$$\langle s \rangle = \frac{1}{N}\sum_{i=1}^{N} s_i;\; s_i = \sum_j w_{ij}, \tag{3}$$

where w_j is the weight of the edges between the ith and jth nodes, and N is the number of nodes of the corresponding set.

- Clustering coefficient of Robins–Alexander, C_{RA}. For a bipartite network, Robins and Alexander [31] defined the bipartite clustering coefficient as four times the number of four cycles, C_4, divided by the number of three paths, L_3, that is:

$$C_{RA} = \frac{4C_4}{L_3}. \tag{4}$$

- Number of connected components, N_{CC}. A connected component is a subnetwork containing one or more nodes such that there is a path connecting any pair of these nodes, but there is no path connecting them to other components.

Part III-Conspiracy indicator: According to our hypothesis—if conspiracy is present, then it is encoded in the network structure of companies with shared personnel—one straightforward measure of the conspiracy levels of a given set of companies is the ratio of connected components N_{CC} relative to the total number of companies N_{CO}. Under this consideration, the mathematical details of the indicator will be shown in the corresponding Results section.

All analyses were done using custom Python code. The network metrics were computed using the Python's NetworkX package [32]. The network visualizations were created using Cytoscape [33].

3 RESULTS

3.1 Part I-Elemental Corporate Features

The main results for the first part of our analysis are presented in **Figures 1**, **2**.

In **Figure 1A**, we show the total number of companies, N_{CO}, and personnel, N_{PE}, per case. In **Figure 1B**, we show the total number and cumulative number of companies created per year for all cases, $n_{CO,t}$. This provides an overall perspective of the temporal scale of action that spans several federal administrations (as indicated with light and shaded regions) and the dynamics of the companies' creation rate that spikes in recent years, with the cumulative number of companies indicating that around 2012, we got approximately 50% of all the companies. This is further explored in **Figures 2A–D**, where we present: (first row) the number and cumulative percentage of companies created per year, respectively, (second row) number and cumulative percentage of new personnel per year, respectively, and (third row) cumulative number of companies and personnel per year. Light and shaded regions indicate the changes in the local or

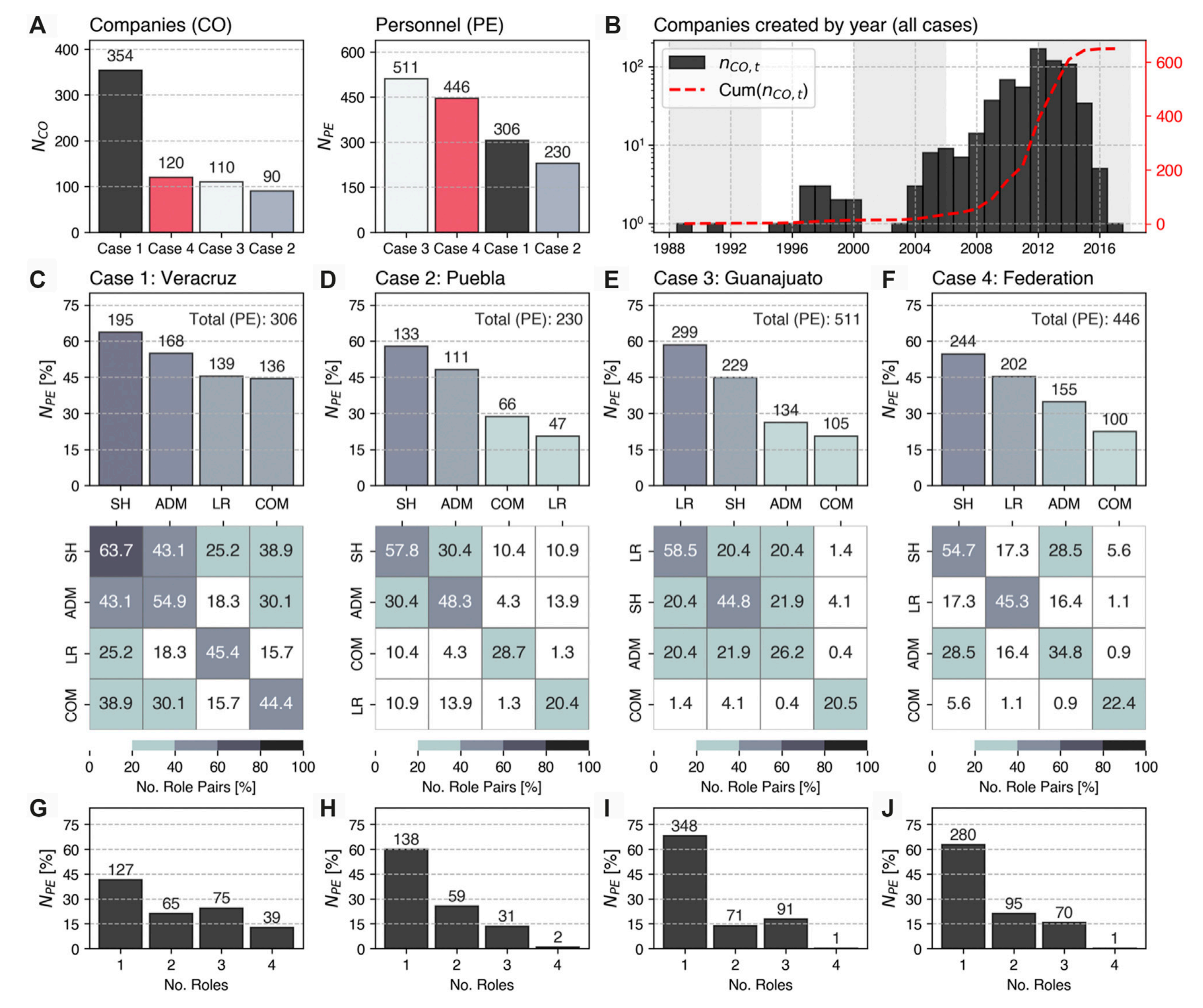


FIGURE 1 | Corporate features per case. **(A)** Number of companies and personnel, N_{CO} and N_{PE}, respectively. **(B)** Number and cumulative number of companies created yearly for all cases, $n_{CO,t}$ and Cum ($n_{CO,t}$), respectively. Light and shaded bands correspond to different federal administrations. **(C–F)** Ownership and management statistics: (top) percentage of personnel according to its role within a company, with exact numbers also indicated; (bottom) role pair co-occurrence percentage matrix. **(G–J)** Percentage of personnel according to the number of possible role types, with exact numbers also indicated.

federal government administrations accordingly. In particular, notice how the rate of creation of companies is always below the rate of addition of new personnel, except for Case 1.

In **Figures 1C–F**, we present some ownership and management statistics. First (top of the panel), we show the percentage of personnel according to its role within a company (SHs, ADMs, COMs, and LRs), with exact numbers indicated. Here, it is observed that the most prominent role is the one of shareholder (SH), followed by administrator (ADM) and legal representative (LR), and last, commissioner (COM). Second (bottom of the panel), we show the role-pairs in the form of a co-occurrence percentage matrix. The diagonal reproduces the distribution on top, while off-diagonal elements are associated with the percentage of individuals that fall within the intersection of any role-pairs. The most prominent role-pair is SH–ADM, which in all cases shows up above 20%, followed by SH–LR and ADM–LR that fall closely below 20%. Commissaries seem to be the least important actors as seen for Cases 2–4; however, they recover an important role for Case 1, where the different types of roles are highly connected, indicating a high degree of role multiplicity. This is even clearer in **Figures 1G–J**, where, in correspondence with the previous panel, we show the personnel profiles according to the number of roles performed. For Cases 2–4, most people (around 60%) perform just one role, while four roles is the least likely profile to be observed. However, Case 1 displays a maximum number of these very unlikely four-role profiles and the least percentage of one-role individuals.

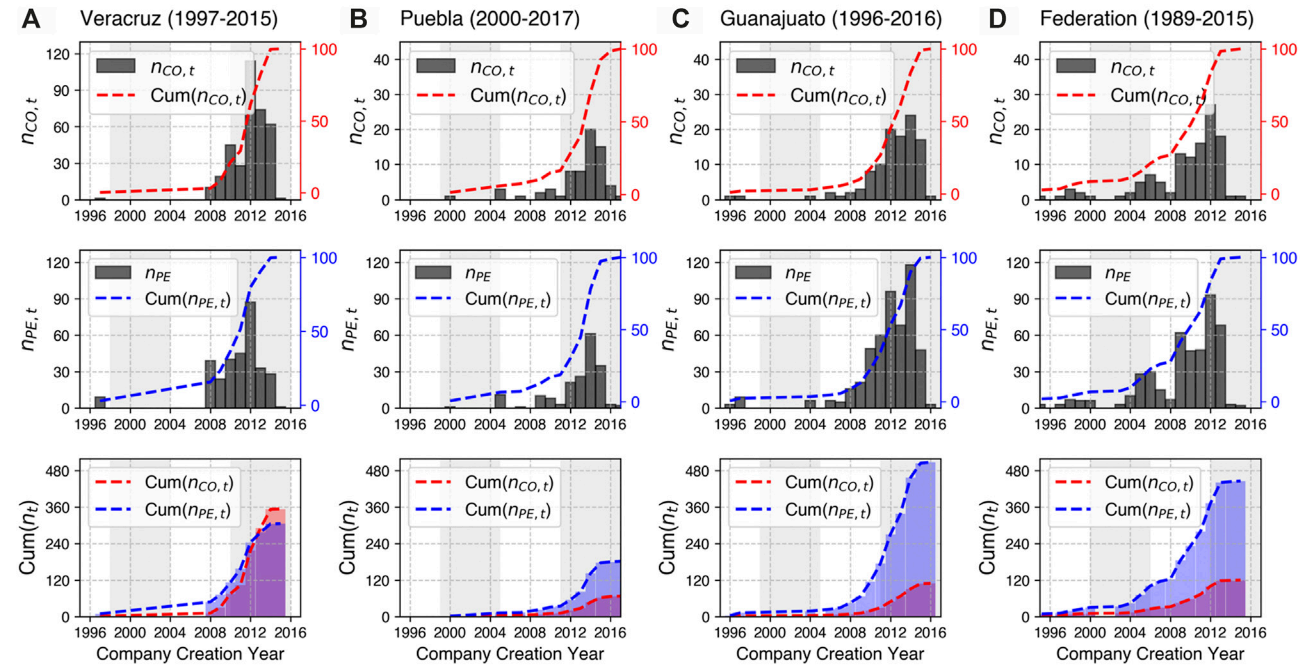


FIGURE 2 | Corporate temporal features per case. **(A–D)** (first row) Number and cumulative percentage of companies created per year, $n_{CO,t}$ and Cum ($n_{CO,t}$), respectively; (second row) number and cumulative percentage of new personnel per year, $n_{PE,t}$ and Cum ($n_{PE,t}$), respectively; (third row) cumulative number of companies and personnel per year, Cum ($n_{CO,t}$) and Cum ($n_{PE,t}$), respectively. Light and shaded regions indicate the changes in the local or federal government administrations according to case. On top, the year period according to the first and the last company created is indicated.

As observed in this part of the analysis, although Cases 2–4 vary in the number of companies and personnel, they share common properties regarding the distribution of role-pairs, role profiles, and temporal dynamics. Case 1 clearly presents an anomalous behavior that deviates from the rest.

3.1.1 Part II-Bipartite Networks

The main results for the second part of our analysis are presented in **Figures 3**, **4**.

In **Figure 3**, we show the network visualizations and some network metrics: Case 1 (Veracruz), **Figure 3A**; Case 2 (Puebla), **Figure 3B**; Case 3 (Guanajuato), **Figure 3C**; and Case 4 (Federation), **Figure 3D**. For each case, we indicate the number of company nodes, N_{CO}; the number of personnel nodes, N_{PE}; the mean degree of company and personnel nodes, $\langle k\rangle_{CO}$ and $\langle k\rangle_{PE}$, respectively; the mean strength of company and personnel nodes, $\langle s\rangle_{CO}$ and $\langle s\rangle_{PE}$, respectively; the number of edges, L; the number of connected components, N_{CC}; the number of shell companies, N_{SH}; the density, δ; and the clustering coefficient of Robins–Alexander, C_{RA}.

We found a remarkable difference in the network connectivity among cases, with Case 1 being the only one forming one giant connected component, while Cases 2–4 show very similar sparsity, although the density (given by **Eq. 1**) is quite similar for all cases, with $\delta \approx 0.01$. Another relevant difference was found in the values of the clustering coefficient of Robins–Alexander, C_{RA} (given by **Eq. 4**), for whose, the maximum value is associated with Case 4, $C_{RA} = 0.761$, while the minimum value is associated with Case 1, $C_{RA} = 0.05$. This counterintuitive result can be understood in terms of the definition of C_{RA}, in which only four cycles, C_4, are counted. Thus, the C_4 property is found in a great number of connected components in Cases 2–4 than in Case 1, for which this interpretation leads to the conclusion that companies and people tend to form more cycles of greater size or even chains than closely connected clusters.

In the case of the mean degree, $\langle k\rangle$, we found that there are important differences according to the node type. For the companies, $\langle k\rangle_{CO}$ represents their average number of personnel. We found that $3.09 \leq \langle k\rangle_{CO} \leq 5.12$, indicating the small size of the companies on average, with Case 3 being the one with the greatest corporate mean degree. For the personnel, $\langle k\rangle_{PE}$ represents their average number of associated companies. We found that $1.10 \leq \langle k\rangle_{PE} \leq 3.84$. In particular, the personnel mean degree of Cases 2–4 is quite similar, $\langle k\rangle_{PE} \approx 1$, while the one for Case 1 is $\langle k\rangle_{PE} \approx 4$, both in direct correspondence with the observed connectivity in **Figure 3**.

The mean strength, $\langle s\rangle$, is best understood in conjunction with the mean degree, $\langle k\rangle$. Specifically, the difference between the mean strength and mean degree, $\langle s\rangle - \langle k\rangle$, is a measure of the average role multiplicity in a corporate network. For example, given a corporate network for which $\langle k\rangle = 3$ and $\langle s\rangle = 5$ (either for the companies or the personnel), $\langle s\rangle - \langle k\rangle = 2$ implies that, on average, two out of three edges have a strength of two, indicating role multiplicity. Therefore, the greater the difference between the mean strength and the mean degree, the greater the multiplicity of roles in a corporate network.

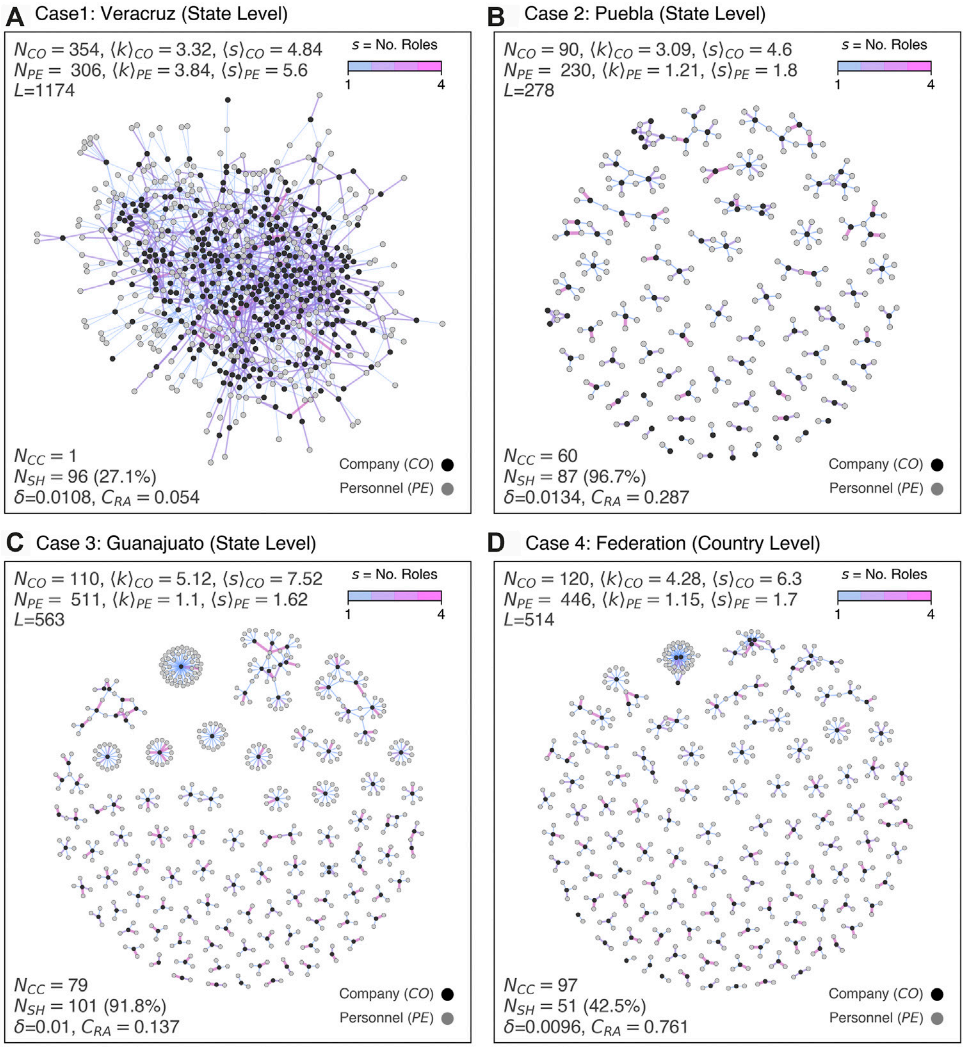


FIGURE 3 | Network visualizations and metrics. For each case **(A–D)**, we indicate the number of company nodes, N_{CO}; the number of personnel nodes, N_{PE}; the mean degree of company and personnel nodes, $\langle k \rangle_{CO}$ and $\langle k \rangle_{PE}$, respectively; the mean strength of company and personnel nodes, $\langle s \rangle_{CO}$ and $\langle s \rangle_{PE}$, respectively; the number of edges, L; the number of connected components, N_{CC}; the number of shell companies, N_{SH}; the density, δ; and the clustering coefficient of Robins–Alexander, C_{RA}. The colorbar is related to the strength of the edges.

Notably, the differences and similarities previously described are best appreciated by looking at the evolution of these quantities in time. In **Figure 4**, we present a comparative corporate network dynamics for all cases with respect to different network metrics. As the time variable, we consider the normalized cumulative number of companies per year, $\text{Cum}(n_{CO,t})^* \in [0, 1]$.

In **Figure 4A**, we show the mean degree and mean strength difference for company-type nodes (or the average role multiplicity in a company) as time progresses. Remarkably, Case 1 mimics Case 2 and Case 4 at initial stages; however, the latter has higher values and a different behavior (steeper slope) than the former one. The higher values of Case 3 indicate that, on average, companies tend to have personnel with higher role multiplicity. In **Figure 4B**, we show the mean degree and mean strength difference for personnel-type nodes (or the average role multiplicity of an individual). Remarkably, Cases 2–4 display identical (almost constant steady) dynamics, while Case 1 clearly differentiates from the rest, indicating a highly irregular increase in the role multiplicity of individuals as time progresses.

The previous observations regarding Case 1 can be complemented and best understood by comparing them with the cumulative number of companies, $\text{Cum}(n_{CO,t})$, and the cumulative number of new personnel, $\text{Cum}(n_{PE,t})$, per year (see **Figure 2**). For Cases 2–4, we have that the inequality $\text{Cum}(n_{CO,t}) < \text{Cum}(n_{PE,t})$ is always satisfied, indicating that companies and new personnel with same role multiplicity are added in the same proportion per year, while for Case 1, we have that the inequality is not always satisfied, indicating that companies and personnel are added at a disproportionate rate ($\text{Cum}(n_{CO,t}) > \text{Cum}(n_{PE,t})$) in which the same personnel is used in multiple roles for multiple companies.

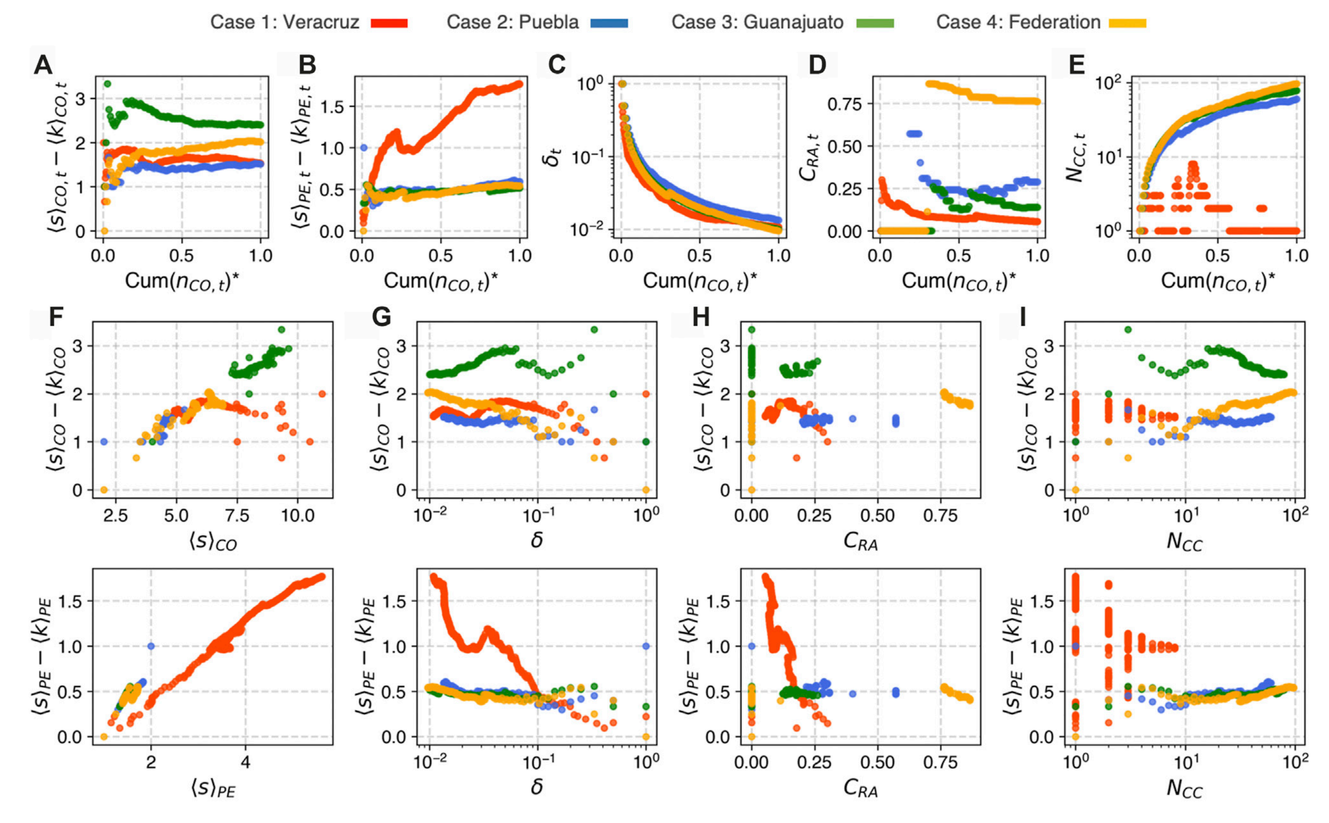


FIGURE 4 | Corporate network dynamics. **(A)** Mean degree and mean strength difference for company nodes, **(B)** mean degree and mean strength difference for personnel nodes, **(C)** density, **(D)** clustering coefficient of Robins–Alexander, and **(E)** number of connected components. In **(A–E)**, the temporal variable, Cum $(n_{CO,t})^*$, corresponds to the normalized cumulative number of companies per year. In **(F–I)**, time is an implicit variable of the data points for the mean degree and mean strength difference for company nodes (top), and the mean degree and mean strength difference for personnel (bottom), as a function of **(F)** companies mean strength, **(G)** density, **(H)** clustering coefficient, and **(I)** number of connected components.

In **Figure 4C**, we show the dynamics of the density, which is almost identical for all cases; in **Figure 4D**, we show the clustering coefficient of Robins–Alexander, which presents different dynamics for each case, with Case 1 and Case 4 being clearly differentiated at all times, while Cases 2 and 3 only for their later years; in **Figures 4D,E**, we show the number of connected components. Notably, Cases 2–4 display the same dynamics in their formation of connected components, while Case 1 is completely anomalous.

In **Figures 4F–I**, we present a comparative analysis for both role multiplicity indicators, $\langle s\rangle_{CO} - \langle k\rangle_{CO}$ and $\langle s\rangle_{PE} - \langle k\rangle_{PE}$, in which we display them as a function of the company's mean strength, the density, the clustering coefficient, and the number of connected components, respectively. Here, time is an implicit variable of the data points.

The results presented in this part of our analysis show that network science is a powerful analytical framework that is able to provide a deeper insight into the static and dynamical features of the cases studied. In particular, they clearly highlight Case 1 as a very irregular event compared even with other corruption cases. However, when it comes to having a better measure of the conspiracy levels for each case, it is hard to identify the set of network metrics that might be able to provide that description.

3.1.2 Part III-Conspiracy Indicator

According to our hypothesis—if conspiracy is present, then it is encoded in the network structure of companies with shared personnel—one straightforward measure of the conspiracy levels of a given set of companies is the ratio of connected components, N_{CC}, relative to the total number of companies, N_{CO}, that is,

$$\phi_C = 1 - \frac{N_{CC}}{N_{CO}}, \tag{5}$$

where $N_{CC} \in [1, N_{CO}]$. The minimum value, ϕ_C^{min}, is given when $N_{CC} = N_{CO}$, at which $\phi_C^{min} = 0$, indicating that all companies are disconnected (no conspiracy). Its maximum value, ϕ_C^{max}, is given when $N_{CC} = 1$, indicating that all companies are connected into one giant component (maximum conspiracy levels). In this case,

$$\phi_C^{max} = 1 - \frac{1}{N_{CO}}. \tag{6}$$

Notice that both ϕ_C and ϕ_C^{max} are dependent on N_{CO} and well defined for finite or small-size networks. Therefore, the distance to maximum conspiracy would simply be

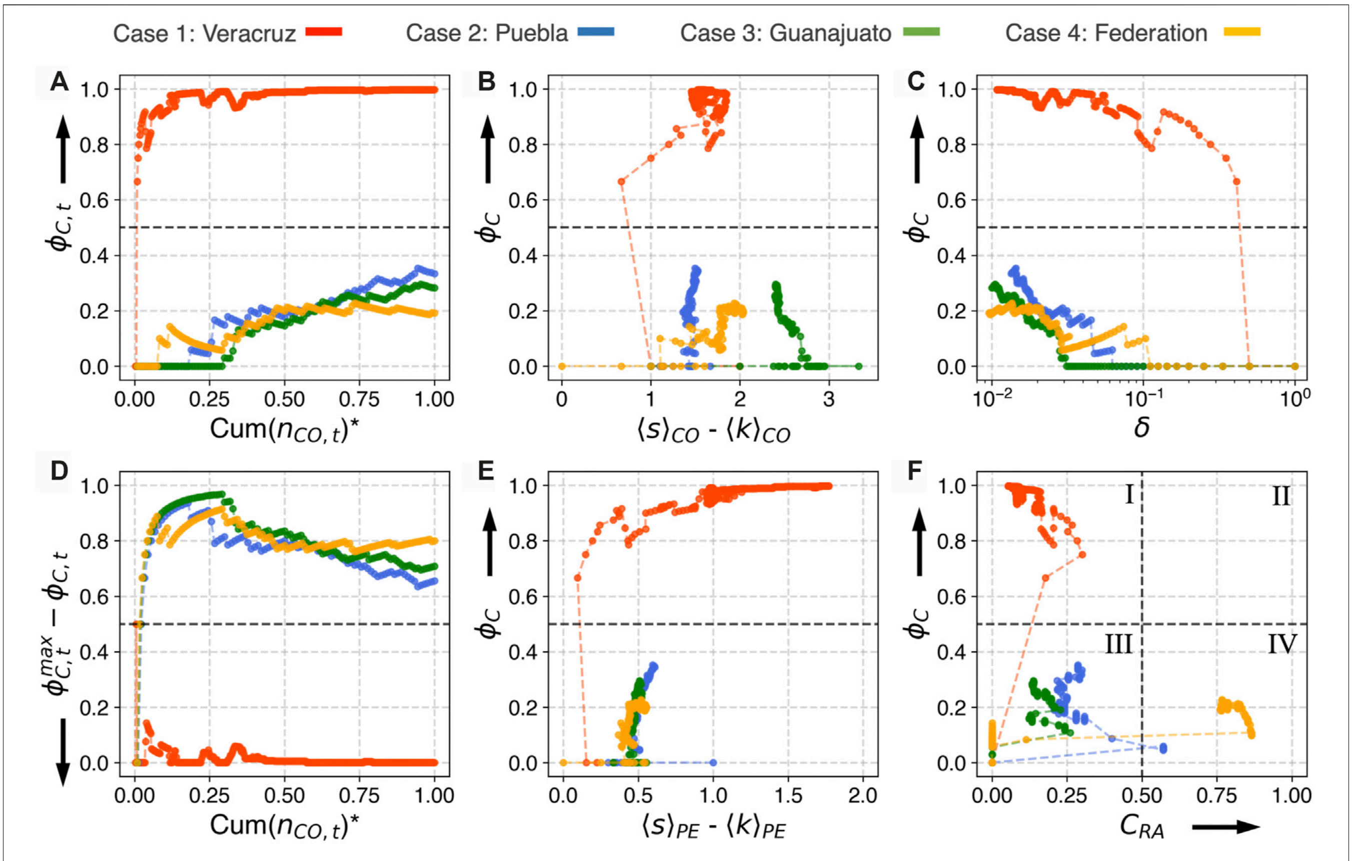


FIGURE 5 | Network-based conspiracy description. **(A)** Conspiracy indicator, $\phi_{C,t}$, and **(B)** distance to maximum conspiracy, $\phi_{C,t}^{max} - \phi_{C,t}$, as a function of the normalized cumulative number of companies per year, $\mathrm{Cum}\,(n_{CO,t})^*$; in **(C–F)**, time is an implicit variable of the data points for the conspiracy indicator as function of the **(C)** companies and **(D)** personnel mean strength and mean degree difference, **(E)** density, and **(F)** clustering coefficient. In all subplots, the arrows indicate the direction of increasing conspiracy.

$$\Delta\phi_C = \phi_C^{max} - \phi_C = \frac{N_{CC} - 1}{N_{CO}}. \quad (7)$$

Here, when $N_{CC} = 1$, it implies maximum conspiracy, while for $N_{CC} = N_{CO}$, it recovers ϕ_C^{max}.

In **Figure 5**, we present our description based on the conspiracy indicator ϕ_C given by **Eq. 5**. In **Figure 5A**, we show its temporal evolution as a function of the normalized cumulative number of companies per year, $\mathrm{Cum}\,(n_{CO,t})^* \in [0, 1]$. In **Figure 5B**, we show its dynamical evolution relative to the maximum conspiracy indicator given by **Eqs. 6**, **7**. In both figures, a clear distinction can be observed among Case 1 and Cases 2–4. In **Figures 5C,D**, we show its behavior with respect to the role multiplicity indicators $\langle s\rangle_{CO} - \langle k\rangle_{CO}$ and $\langle s\rangle_{PE} - \langle k\rangle_{PE}$, respectively. In **Figure 5E**, we compare it to the density. In **Figures 5C–E**, a clear distinction can again be observed among Case 1 and Cases 2–4 with respect to ϕ_C.

Finally, in **Figure 5F**, we show its behavior with respect to the clustering coefficient of Robins–Alexander. Considering that both quantities are bounded, $\phi_C \in [0, 1]$ and $C_{RA} \in [0, 1]$, the plane could be divided into the indicated quadrants or regions I–IV, with each one providing a qualitative description for each case in terms of the number of connected components information, given by ϕ_C, and the clustering, C_{RA}. In this way, Case 1 evolves in time toward the low clustering and high conspiracy region I, Cases 2 and 3 remain within the low clustering and low conspiracy region III, and Case 4 falls into the high clustering and low conspiracy region IV. Remarkably, the way each case is differentiated in terms of conspiracy-clustering regions is in an interesting correspondence with the general features of each corruption scandal, as it is discussed in the following section.

4 DISCUSSION

4.1 Organized Crime, Conspiracy, and Corruption in Procurement

The organized crime studies in law and social sciences are long familiar with the organization's analytical model of a network [34] and, more recently, with network analysis applied to criminal cases [24, 35]. According to Steffensmeier, the normative concept of organized crime groups is typically defined after three main characteristics: a formalized structure whose primary objective is to obtain money through illegal activities, the structure's activity has continuity over time, and it maintains its position through use of violence or threat of violence, corruption of public officials, and extortion [34]. Networks as criminal organizations are difficult to prosecute because they camouflage information and actions, and they grant impunity by adding security actors, such as public functionaries and politicians [34, 36], and also by acting in criminal conspiracy.

Conspiracy is a legal crime convenient when there is no substantive proof that an offense has been committed and evidence has been covered, but the agreement for fraud exists, and the centrality of the group activity is apparent as a "single invisible empire" [37]. Therefore, as a complexity problem, the recognition of factors that trigger criminal conspiracy for corruption in public procurement represents a step forward in predicting certain behaviors and discussing control strategies. In this research, criminal conspiracy is considered as an organized crime act or as "an agreement between two or more people to commit an illegal act, along with an intent to achieve the agreement's goal." The analysis of the procurement environment as a network commonly considers the connections between the contracting institutions and suppliers, either under a bipartite institution–supplier network or a firm–firm co-bidding network [8, 17, 20]. However, the information regarding the connections of public servants and business people to their corresponding institutions or companies, respectively (beyond hidden political ties), is an aspect that can be further explored. In particular, our results tell of the importance of the analysis of the supplier environment of a given country under a network and complex system approach, and of the relevance of creating network-based corruption indicators at the level of ownership and management of the companies involved in procurement.

The risk in procurement networks has also been qualified by a centrality proxy that opposes competition legislation [17]. The present research argues that while centrality is relevant as a normative risk indicator, it might not be useful in a political context where the rule of law is deteriorating and large-scale public and private actors acting unlawfully is common (see Data and Methods regarding the Mexican scenario).

4.1.1 Complexity Science Approaches to Conspiracy Networks

Corruption is a phenomenon that occurs within the structure and dynamics of complex social, economic, political, and technological systems. Although there is no general consensus on a comprehensive definition for complex system, in this study, we considered the following: complex systems are networks made of a number of components that interact with each other, typically in a nonlinear fashion. Complex systems may arise and evolve through self-organization such that they are neither completely regular nor completely random, permitting the development of emergent behavior at macroscopic scales [38].

As such, corruption in public procurement manifests as a non-separable or intertwined activity that takes places within a complex procurement system defined by contracting institutions, suppliers, public servants, and business people interacting through tendering processes, awarded contracts, labor relationships, and through hidden ties or connections [8, 12, 15]. At the level of ownership and management of the companies involved in procurement, this definition allowed us to hypothesize the criminal conspiracy (as a proxy for corruption) of corporations as an underlying behavior that creates connections among companies with shared personnel and that could be identified and quantified through the macroscopic properties of the corporate networks. First, we found that the companies involved in the scandals have unusual corporate profiles full of multi-role personnel, especially Case 1 (**Figure 1**). Also, the rate of creation of companies compared to the rate of addition of new personnel displayed some common patterns, except for Case 1 (**Figure 2**). Then, the analysis was performed under an empirical bipartite company–personnel network approach that allowed us to do the following: i) show how companies indeed nucleate into networks, with different properties, due to the existence of shared multi-role personnel, pointing toward potential criminal conspiracy activity (**Figure 3**); ii) describe and quantify the subtle differences and similarities in the structure and dynamics of the corporate networks for each corruption scandal (**Figure 4**); iii) show the degree to which the companies behave in a conspiracy fashion by introducing a conspiracy indicator, ϕ_C, based on the fraction of connected components of each network (**Eq. 5**); and iv) identify some elements that contribute to the description of corporate criminal conspiracy based on network-based metrics, specifically, the conspiracy-clustering description (**Figure 5**). The conspiracy-clustering description (see **Figure 5F**) generated a relevant classification of the cases that are in an interesting correspondence with their general features, such as scale of operation and percentage of shell companies involved: Veracruz, Case 1 (local scale of operation with less than 50% of shell companies involved) evolves in time toward the low clustering and high conspiracy region I; Puebla y Guanajuato, Cases 2 and 3 (local scale of operation with close to 100% of shell companies involved) remain within the low clustering and low conspiracy region III; and Federation, Case 4 (federal scale of operation with less than 50% of shell companies involved) falls into the high clustering and low conspiracy region IV.

We would like to remark that although our analysis is able to differentiate the corrupt cases in terms of the conspiracy and other metrics (such as the clustering) of the corporate networks, it still needs to be applied in a context that also includes non-corrupt companies in order to test its capabilities to classify corrupt from non-corrupt patterns. Also, one downside of our analysis is that so far, we have no evidence to conclude that these network structures are features of the whole ecosystem. As such, our research does not replace or substitute traditional corruption approaches or research on the procurement environment, but brings more elements to consider regarding corporate risk indicators based on the network properties of companies connected at the ownership and management level before any tendering process or contract is awarded.

4.1.2 Professional Practice and Further Research

Corruption practices in procurement processes have been extensively explored in the literature [8, 39], and it has been relevant to promote standards, legislation, and enforcement mechanisms in order to prevent opportunities for behaviors such as collusion, bid-rigging, and cartel agreements [17, 20, 40, 41], often by looking at the economic incentives and punishment rules [42, 43]. Global strategies seek to increase transparency, convenience, higher revenue in procurement through electronic government procurement [2, 4, 44, 45], and improve the value of reputation of the suppliers with compliance policies [46–48]. The underlying presumption seems to be that governments' purchasing decisions are influenced by legitimate concerns and governments are willing to work with firms with a reputation, where public–private partnership will be endorsed within the compliance of domestic and international legal framework and codes of ethics [9]. Nevertheless, the assessments and risk metrics in procurement are becoming more comprehensive in order to fully ascertain the presence of organized crime acting in conspiracy and other offenses such as collusion and fraud in a way that they promote competition on the bases of law enforcement scenarios [49, 50]. Governments are the controllers and the responsible parties for enforcing anti-corruption legislation, but fragmentation, electoral interests, or private gain often diverts public officials from public purpose, and corruption is more difficult to detect [51].

The cases presented in this study represent examples of corruption where the government officials such as governors; heads of federal and state agencies, companies, and notaries; and natural persons undertake corruption action plans allegedly for electoral and self-profit ends. In contexts where governments are acting unlawfully, buyers and/or suppliers are acting in conspiracy with the obvious intent to remain undetected, and public officials protect or act in collusion with corporate networks; information traces hidden in journalistic investigations remain one of the best tools to expose criminal activities. Using corruption scandals data precisely benefits scientific analysis by often proving that regardless of a comprehensive legal framework, procurement practices overtake normative rationality, to serve a practical realpolitik system that poses a challenge to the design of corruption proxies and network analysis. Also, by analyzing corporate corruption under a network's perspective, it is possible to avoid biases, such as taking for granted honesty, firm's reputation [18], or public officials' lawfulness, that go beyond the risks posed by single author's offenses into patterns of racketeering activity [37].

Further research should delve into which corporate characteristics are likely to maximize connectivity by exploring the contribution of each personnel layer to the macro properties of the network together with the heuristic function for alleged conspiracy introduced in this article. Also, with more official information about the cases regarding government participation and financial transactions or taxpayers activities, extra analysis could assess the role of government in perpetuating and protecting corruption networks in procurement. Likewise, the more corruption judicial cases there could be before courts, the better understanding of the wrongfulness of the conspiracy activities could be categorized.

DATA AVAILABILITY STATEMENT

For privacy protection, we have held in anonymity all the information regarding the identification of the companies and associated people. The data is available at the Supporting Data for "Conspiracy of Corporate Networks in Corruption Scandals" repository, https://figshare.com/s/1efd3d6ce33f289044fa.

AUTHOR CONTRIBUTIONS

ILP reviewed the criminal networks and legal literature, connecting concepts and identifying the contribution of the analysis of cases here presented. JRNC developed the datasets, designed the methodology, and performed the network and statistical analyses. Both authors contributed equally to the interdisciplinary analysis of the results and manuscript preparation and approval.

FUNDING

Research funded by Universidad Nacional Autónoma de México (UNAM) through the "Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica" (PAPIIT), Grant No. PAPIIT-IV300217.

ACKNOWLEDGMENTS

We thank Arturo Ángel, Raúl Olmos and Yosune Chamizo, members of the investigative journalism group, Animal Político, and the non-governmental organization, Mexicanos Contra la Corrupción e Impunidad, for the access granted to their data records on the "Veracruz" and "Federal" cases. We thank René Valencia e Israel Velázquez, members of the investigative journalism group, Datamos, for the access granted to their data records on the "Puebla" case. We are thankful for the important work of the Comité de Participación Ciudadana of the state of Guanajuato, México, that lead to the data for the "Guanajuato" case. We also are thankful to all the students that contributed to this research, particularly Alfredo Andrés López Rivas, Ricardo Martínez Morales, Diana Gabriela Baena Rivera, Ana Laura Chávez García, Miguel Ángel Enríquez Gutiérrez and Linda Angulo.

REFERENCES

1. UN. *United Nations Convention against Corruption*. United Nations (2003).
2. UNECE. *UNECE Standard on a Zero Tolerance Approach to Corruption in PPP Procurement. Implementing the United Nations Agenda for Sustainable Development through Effective*. United Nations (2017).
3. Schultz J, and Søreide T. Corruption in Emergency Procurement. *Disasters* (2008) 32(4):516–36. doi:10.1111/j.1467-7717.2008.01053.x
4. WB. *Enhancing Government Effectiveness and Transparency. The Fight against Corruption. Global Report*. Washington, DC: World Bank (2020). Technical report.
5. DeAses AJ. Developing Countries: Increasing Transparency and Other Methods of Eliminating Corruption in the Public Procurement Process. *Public Contract L J* (2005) 34(3):553–72.

6. Jain AK. Corruption: A Review. *J Econ Surv* (2001) 15(1):71–121. doi:10.1111/1467-6419.00133
7. Rose-Ackerman S, and Palifka B. *Corruption and Government: Causes, Consequences, and Reform*. Cambridge: Cambridge University Press (1999).
8. Fazekas M, Cingolani L, and Tóth B. A Comprehensive Review of Objective Corruption Proxies in Public Procurement: Risky Actors, Transactions, and Vehicles of Rent Extraction. In: GTI-WP/2016 (2016). p. 3.
9. OECD. *Preventing Corruption in Public Procurement*. Paris: Organization for Economic Co-operation and Development (2016). Technical report.
10. Reeves-Latour M, and Morselli C. Bid-rigging Networks and State-Corporate Crime in the Construction Industry. *Social Networks* (2017) 51:158–70. doi:10.1016/j.socnet.2016.10.003
11. Lambert-Mogiliansky A, and Sonin K. Collusive Market Sharing and Corruption in Procurement. *J Econ Manage Strategy* (2006) 15(4):883–908. doi:10.1111/j.1530-9134.2006.00121.x
12. Kertész J, and Wachs J. Complexity Science Approach to Economic Crime. *Nat Rev Phys* (2020) 1–2. doi:10.1038/s42254-020-0238-9
13. Fazekas M, and Tóth B. *Proxy Indicators for the Corrupt Misuse of Corporations*. Bergen, Norway: Michelsen Institute (2017).
14. Jeude JAv. Ld., Aste T, and Caldarelli G. The Multilayer Structure of Corporate Networks. *New J Phys* (2019) 21(2):025002. doi:10.1088/1367-2630/ab022d
15. Luna-Pla I, and Nicolás-Carlock J. Corruption and Complexity: a Scientific Framework for the Analysis of Corruption Networks. *Appl Netw Sci* (2020) 5(1):13. doi:10.1007/s41109-020-00258-2
16. Ribeiro HV, Alves LGA, Martins AF, Lenzi EK, and Perc M. The Dynamical Structure of Political Corruption Networks. *J Complex Networks* (2018) 6(6):989–1003. doi:10.1093/comnet/cny002
17. Wachs J, Fazekas M, and Kertész J. Corruption Risk in Contracting Markets: a Network Science Perspective. *Int J Data Sci Analytics* (2020), 1–16. doi:10.1007/s41060-019-00204-1
18. Kadens E. Cheating Pays. *Columbia L Rev* (2019) 119: 527–590.
19. Ferrali R. Partners in Crime? Corruption as a Criminal Network. *Games Econ Behav* (2020) 124:319–53. doi:10.1016/j.geb.2020.08.013
20. Lyra M, Curado A, Damásio B, Bação F, and Pinheiro FL. Characterization of the Firm-Firm Public Procurement Co-bidding Network from the State of Cear'a (brazil) Municipalities. *arXiv* 2021). preprint arXiv:2104.08547.
21. IMCO. *Compras públicas en méxico. competencia: la gran ausente*. Mexico: Instituto Mexicano para la Competitividad A.C (2020). Technical report.
22. OECD. *Review of Practices of Local Public Procurement in Five Mexican States*. Paris: Organization for Economic Cooperation and Development (2015). Technical report.
23. Animal-Político. *Animal-Político (2016). El caso de las empresas fantasma de veracruz*. Mexico: Animal Político (2016). Technical report.
24. DellaPosta D. Network Closure and Integration in the Mid-20th century American Mafia. *Soc Networks* (2017) 51:148–57. Crime and Networks. doi:10.1016/j.socnet.2016.11.005
25. FATF. *Concealment of Beneficial Ownership*. FATF (2018). Technical report.
26. Valencia R, and Velázquez M. *Puebla, fábrica de empresas fantasma*. Puebla: Datamos (2020). Technical report.
27. Pizano C, and Lab P. Dos empresas contratadas por autoridades de guanajuato están ligadas a la estafa maestra. *Anim Político* (2020).
28. León G. *Dependencias guanajuatenses habrían pagado hasta 193 mdp a empresas fantasmas: Cpc*. Guanajuato: Zona Franca (2020).
29. Mcca A. *La estafa maestra graduados en desaparecer dinero público*. Mexico: Mexicanos Contra la Corrupción and Animal Político (2017). Technical report.
30. Menczer F, Fortunato S, and Davis CA. *A First Course in Network Science*. Cambridge: Cambridge University Press (2020). doi:10.1017/9781108653947
31. Robins G, and Alexander M. Small Worlds Among Interlocking Directors: Network Structure and Distance in Bipartite Graphs. *Comput Math Organ Theor* (2004) 10(1):69–94. doi:10.1023/b:cmot.0000032580.12184.c0
32. Hagberg A, Swart P, and S Chult D. *Exploring Network Structure, Dynamics, and Function Using Networkx*. Los Alamos, NM (United States): Los Alamos National Lab.(LANL) (2008). Technical report.
33. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, and Ramage D. Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res* (2003) 13(11):2498–504. doi:10.1101/gr.1239303
34. Steffensmeier D, and Martens F. Crime: Organized. In: NJ Smelser and PB Baltes, editors. *International Encyclopedia of the Social and Behavioral Sciences*. Oxford: Pergamon (2001). p. 2928–34. doi:10.1016/b0-08-043076-7/01850-7
35. Tumminello M, Petruzzella F, Ferrara C, and Miccichè S. Anagraphical Relationships and Crime Specialization within Cosa Nostra. *Soc Networks* (2021) 64:29–41. doi:10.1016/j.socnet.2020.07.011
36. Sargiacomo M, Ianni L, D'Andreamatteo A, and Servalli S. Accounting and the Fight against Corruption in Italian Government Procurement: A Longitudinal Critical Analysis (1992-2014). *Crit Perspect Account* (2015) 28:89–96. doi:10.1016/j.cpa.2015.01.006
37. Sergi A. Organised Crime in English Criminal Law. *J Money Laundering Control* (2015) 18(2):182–201. doi:10.1108/jmlc-10-2014-0038
38. Sayama H. *Introduction to the Modeling and Analysis of Complex Systems*. New York: Open SUNY Textbooks (2015). doi:10.7551/978-0-262-33027-5-ch009
39. Fazekas M, and Kocsis G. Uncovering High-Level Corruption: Cross-National Objective Corruption Risk Indicators Using Public Procurement Data. *Br J Polit Sci* (2020) 50(1):155–64. doi:10.1017/s0007123417000461
40. Porter RH, and Zona JD. Detection of Bid Rigging in Procurement Auctions. *J Polit Economy* (1993) 101(3):518–38. doi:10.1086/261885
41. Conley TG, and Decarolis F. Detecting Bidders Groups in Collusive Auctions. *Am Econ J Microeconomics* (2016) 8(2):1–38. doi:10.1257/mic.20130254
42. Olken BA, and Pande R. Corruption in Developing Countries. *Annu Rev Econ* (2012) 4(1):479–509. doi:10.1146/annurev-economics-080511-110917
43. Kwon I. Motivation, Discretion, and Corruption. *J Public Adm Res Theor* (2012) 24(3):765–94. doi:10.1093/jopart/mus062
44. Concha G, Astudillo H, Porrúa M, and Pimenta C. E-government Procurement Observatory, Maturity Model and Early Measurements. *Government Inf Q* (2012) 29:S43–S50. Government Information Networks. doi:10.1016/j.giq.2011.08.005
45. OECD. *Follow up Report on mexico's Compranet Reform Improving E-Procurement through Stakeholder Engagement*. Paris: Organization for Economic Co-operation and Development (2019). Technical report.
46. Klabi H, Mellouli S, and Rekik M. A Reputation Based Electronic Government Procurement Model. *Government Inf Q* (2018) 35(4Supplement):S43–S53. Platform Governance for Sustainable Development. doi:10.1016/j.giq.2016.01.001
47. Haberbush KL. Limiting the Government's Exposure to Bid Rigging Schemes: A Critical Look at the Sealed Bidding Regime. *Public Contract L J* (2000) 30(1):97–122.
48. UNCITRAL. *Uncitral Model Law on Public Procurement*. New York: UNCITRAL (2014). Technical report.
49. LaCasse C. Bid Rigging and the Threat of Government Prosecution. *RAND J Econ* (1995) 26(3):398–417. doi:10.2307/2555995
50. Velasco RB, Carpanese I, Interian R, Paulo Neto OCG, and Ribeiro CC. A Decision Support System for Fraud Detection in Public Procurement. *Intl Trans Op Res* (2021) 28:27–47. doi:10.1111/itor.12811
51. Knack S, Biletska N, and Kacker K. Deterring Kickbacks and Encouraging Entry in Public Procurement Markets: Evidence from Firm Surveys in 88 Developing Countries. In: Policy Research working paper no. WPS 8078. Washington, D.C:. World Bank Group (2017). doi:10.1596/1813-9450-8078

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

PERSPECTIVE
published: 21 June 2021
doi: 10.3389/fphy.2021.666386

Check for updates

Agent-Based Modeling as a Legal Theory Tool

*Sebastian Benthall and Katherine J. Strandburg**

[1]School of Law and Information Law Institute, New York University, New York, NY, United States

Agent-based modeling (ABM) is a versatile social scientific research tool that adapts insights from sociology and physics to study complex social systems. Currently, ABM is nearly absent from legal literature that evaluates and proposes laws and regulations to achieve various social goals. Rather, quantitative legal scholarship is currently most characterized by the Law and Economics (L&E) approach, which relies on a more limited modeling framework. The time is ripe for more use of ABM in this scholarship. Recent developments in legal theory have highlighted the complexity of society and law's structural and systemic effects on it. ABM's wide adoption as a method in the social sciences, including recently in economics, demonstrates its ability to address precisely these regulatory design issues.

Keywords: agent-based modeling, simulation, complexity, law and economics, law and macroeconomics, law and political economy

OPEN ACCESS

Edited by:
Pierpaolo Vivo,
King's College London, United Kingdom

Reviewed by:
Giorgio Fagiolo,
Sant'Anna School of Advanced Studies, Italy
Dominik Klein,
Utrecht University, Netherlands

****Correspondence:***
Katherine J. Strandburg
katherine.strandburg@nyu.edu

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *10 February 2021*
Accepted: *28 May 2021*
Published: *21 June 2021*

Citation:
Benthall S and Strandburg KJ (2021) Agent-Based Modeling as a Legal Theory Tool. Front. Phys. 9:666386. doi: 10.3389/fphy.2021.666386

INTRODUCTION

The time is ripe for legal scholars to use agent-based modeling (ABM) to produce actionable theoretical insights. One major strand of legal scholarship attempts to design or evaluate potential regulatory approaches based on their anticipated effectiveness at achieving societal goals [1, 2, 3]. This is an inherently normative project: both the overall consequentialist perspective and particular choices of appropriate goals are contestable. This sort of legal scholarly project also has social scientific underpinnings, however, because its success depends on the quality of its predictions about how society will respond to legal changes. Indeed, legal scholarship of this sort is often explicitly interdisciplinary, relying on theoretical concepts, models and methods from various fields of social science fields to inform those predictions. Microeconomics has been particularly influential, spawning a sub-field known as "law and economics," ("L&E") or sometimes "economic analysis of law" [4–7]. (As one rough measure of this impact, the LEXIS database of law journal articles contains more than 35,000 articles mentioning "law and economics" or "economic analysis of law," just over 3,000 articles mentioning "law and sociology" or "sociology of law" and just about 1,500 mentioning each of "law and political science" and "law and psychology.")

ABM is a computer simulation approach that has been increasingly deployed in social science to study the societal implications of various specifications of agents (who can be modeled as individuals, firms or other entities), their incentives and decision-making strategies, the interactions between those agents, and the social frameworks in which they interact. The computer simulation approach allows agent-based models to incorporate heterogeneity, nonlinearity and feedback effects in ways that are not possible with more traditional analytical solutions and approximation techniques [8, 9]. Using ABM's bottom-up "generative social science" approach, "fundamental social structures and group behaviors emerge from the interaction of individuals operating in artificial environments under rules that place only bounded demands on each agent's information and computational capacity" [10].

ABM's strengths would appear to make it an attractive approach for exploring the potential societal implications of proposed changes to laws and regulations, particularly in light of the difficulty of employing experimental methods to probe these issues. Moreover, laws and regulations intersect with many aspects of social life that have been studied with ABMs. Yet the legal literature seems surprisingly oblivious to ABM's potential to inform the evaluation of proposals for legal and regulatory change. Because ABM is an increasingly important social scientific tool, its lackluster uptake by legal scholars who aim to predict the effects of regulatory proposals is disappointing. Nonetheless, we believe that several developments make a more robust incorporation of ABM into legal scholarship possible now. A first set relates to legal scholarly demand for less individualistic and more systemic, structural, and political approaches to regulatory design, while a second set relates to the legal academy's capacity for and openness to computational modeling. Both sets are usefully understood in relation to "law and economics," which has been one of the most (arguably the most) influential–and controversial – strands of legal scholarship since the seminal work by Posner and others in the mid-1980s [6, 7, 11].

AGENT BASED MODELING AND ITS ABSENCE FROM THE LEGAL LITERATURE

Arguably rooted in mathematical sociology [12, 13], ABM is an alternative to analytical calculation that leverages computational resources to permit a wide and flexible range of specifications of agents, their incentives and decisionmaking strategies, and the interactions between them [8, 9]. The modeler is also free to specify features of the social, legal and policy frameworks in which the agents interact.

ABMs have been used in many social scientific domains of public policy relevance. We mention only a few examples. Notably, economists have begun to use ABM, especially since the 2008 financial crisis [14, 15, 16, 17, 18]. ABMs have also been used to study housing segregation beginning with models based on ethnic homophily and over time introducing models of market dynamics [13, 19–24]. Another line of research investigates labor markets, referrer networks, segregation, and affirmative action [25–28]. These studies are obviously relevant to questions in legal theory about antidiscrimination law. ABM models of the formation and adoption of norms [29–31] have general relevance to theories about the need for and effects of regulation.

Though Picker [29] article noted the potential usefulness of ABMs in understanding the interaction between norms and law, other legal scholars have not, for the most part, followed his lead. For example, a search of the LEXIS database of law journals for mentions of "agent-based model" or "agent-based modeling" turned up only 80 articles in total. Of those, 61 mention ABM only in footnotes (52) or in passing (9). Of the 19 articles that do more, seven merely propose ABM as a potentially useful technique for addressing policy issues related to: market panic [32]; water pollution in a river basin [33]; "exploratory analysis of policy options" in telecommunications [34, 35]; administrative rulemaking [36]; telecommunications complexity in comparative law [37]; and taxpayer behavior [38]. Seven others cite results from previously published ABM studies: [39] (citing ABM results about dependence of social structure on initial conditions to refute argument for genetic determinism of societal differences); [40] (citing ABM studies in discussion of environmental justice); [41] (citing ABM studies of crime displacement); [42]; (using results of ABM of the effects of economic-based college admissions criteria on racial diversity); [43] (further analysis based on the residential segregation ABM described below); [44] (relating ABM results about collective behavior formation to the formation of customary international law); [45] (using concepts from ABM to consider diffusion of innovations).

Thus, only five articles uncovered by our LEXIS search report new ABM studies designed to address a legal issue. [46], uses ABM to test and critique the assumptions underlying the Supreme Court's invocation of "critical mass" in its affirmative action jurisprudence. [47], uses a detailed agent-based model of the Lake Champlain Basic as part of a larger project aimed at designing policies to ensure clean water. [48], reports the results of an ABM model designed to test theories about what factors influence tax compliance. [29], uses ABM to explore social norm formation as part of an inquiry into when legal regulation is required. [22], uses an ABM to explore how residential segregation can be locked in by historical events that create wealth and social disparities even in the absence of intentional discrimination or any preference for racial homophily. She then argues that current housing discrimination law cannot overcome these lock-in effects. Another seven articles apply the results of a previous ABM study to analyze a legal issue.

Of course, this quick survey is not an exhaustive search of the legal literature: for example, some articles may rely on agent-based modeling without using the term. These numbers also do not capture publications by legal scholars in non-law journals or other venues not included in the LEXIS database. Nonetheless, one can only conclude that ABM has made few inroads into legal scholarship, despite earlier discussions of its potential benefits [49–52].

LAW AND ECONOMICS, ITS CRITICS AND THE ABM OPPORTUNITY

Both the strengths and the weaknesses of Law and Economics make it an important backdrop for understanding the current potential for a more robust incorporation of ABM into legal scholarship. L&E adopts the consequentialist goal of steering the social system toward a desirable state and conceptualizes law primarily in terms of the incentives for individual behavior that it provides (rather than, for example, primarily as a means for compensating harms or providing "just deserts"). It thus requires some method for predicting how society will respond to a legal regime and some way to normatively evaluate the states of the world that are likely to result. Normatively, L&E adopts the goal of designing legal rules that will maximize social welfare, ordinarily defined as the sum of individual utility functions. It traditionally leaves distributional concerns to be addressed (if at all) through the tax system. To predict how society will respond to a proposed legal regime, L&E analyses often employ mathematical models and game theory, adopting simplifying assumptions from neoclassical microeconomics that favor analytical tractability over realistic portrayal of social dynamics and complexity. L&E analysis also tends to focus on designing rules to incentivize transaction-by-transaction efficiency, under the assumption that the cumulative effect will be social welfare maximizing.

L&E is a powerful methodology because, once one accepts its simplifying assumptions, it can be used to explore many questions about legal design, such as whether a negligence rule or a strict liability rule will induce the socially optimal level of precautions in a particular context, what combination of punishment severity and enforcement certainty most effectively deters crime, and whether the costs of implementing a regulatory regime outweigh its benefits. The legal academy's experience with L&E also demonstrates the value of mathematical models for sharpening analysis and clarifying assumptions, as well as (for better or for worse) the persuasive force of simple models. Beyond L&E's appeal as a basis for detailed models, its foundational assumptions, such as that incentives matter and that rational self-interest drives much of human behavior, can usefully be deployed to make "hand-waving," but plausible, informal assessments of incentives and trade-offs. This flexibility means that the L&E approach can be used not only by L&E scholars, but also by a wide range of legally relevant actors, including judges, attorneys, legislators and administrative policymakers.

To predict how affected individuals might respond to a legal change, traditional L&E models individuals as self-interested, informed and rational actors, whose goals are to maximize individual utility by satisfying their preferences. Because utility cannot be easily determined, compared or summed across individuals, L&E typically focuses on arranging society's rules to facilitate "efficient" voluntary transactions, in which participants decide for themselves whether they are made better off. When, as is often the case, transactions have spillover effects ("externalities") on the utility of outsiders, however, the L&E analyst must somehow account for those effects. One approach is to design regulatory mechanisms aimed at forcing the transactors to internalize–or at least account for–the social costs of the externalities. Often, however, L&E analyses simply redefine efficiency in terms of "Kaldor-Hicks" improvement [53, 54], in which "state A is to be preferred to state B if those who gain from the move to A gain enough to compensate those who lose" [55]. This approach requires an interpersonal comparison of utilities that is ordinarily performed using a monetary metric. L&E analyses thus often boil down to transaction-by-transaction marginal cost-benefit analyses. Because the goal is to maximize *total* social utility, the analysis need not consider whether the Kaldor-Hicks losers are actually compensated.

Of course, predicting how individuals will react to legal changes and how the state of the world will evolve as a result is a Herculean (or perhaps Sisyphean) task. Legal systems are thus designed to facilitate revisions when existing law begins to have socially undesirable effects. The marginal, transaction-based approach of L&E implicitly assumes that when this happens, problems with current laws can be detected and incremental course corrections can effectively guide society toward an optimal legal regime. This expectation has led some L&E scholars to embrace the common law system, in which legal doctrine shifts gradually as cases come to court.

The assumptions made in traditional L&E analyses are advantageous for tractability, but have well-known practical, conceptual and normative weaknesses [11]. L&E has always been criticized for its simplified, rational actor model of human motivations and behavior. The subfield of "behavioral economics" arose to develop and implement more realistic models of human beings as boundedly rational, subject to cognitive biases and not fully informed [56]. While such more realistic representations of individual behavior can sometimes be incorporated into L&E's traditional analytical methods, this is more easily done in an ABM, which can incorporate heterogeneity, limited information and various forms of motivations and behavioral rules without the need for analytical tractability. Nonetheless, better modeling of individual behavior is not ABM's most important contribution. Indeed, while ABM can be used for highly detailed and realistic modeling of specific situations (see [33, 47]), ABM's main strength is its ability to demonstrate and explore the ways in which unexpected system-level properties can arise from relatively simple models of individual behavior. In this respect, most applications of ABM to legal theory are likely to retain both the advantages and some of the disadvantages of L&E's simple modeling of individual behavior.

Defining social welfare in terms of total wealth maximization is normatively troubling, especially because the traditional L&E proposal to deal with the unfairness of Kaldor-Hicks efficiency through the add-on of redistributive taxation is both practically and politically infeasible. In practice, traditional L&E tolerates large disparities in its single-minded pursuit of increasing the size of the pie. Some scholars have suggested ways to tweak standard L&E analyses to account for other normatively important considerations such as equality. For example, one might posit individual preferences for altruism, equality, or reciprocity or for biodiversity or clean air [6] or introduce quantitative and qualitative mechanisms for taking non-monetary and distributional social values into account in cost-benefit analysis [57, 58]. These proposals certainly have some practical merit (indeed some of Sunstein's proposals for modified cost-benefit analysis were implemented during the Obama administration). Nonetheless, these attempts to maintain the tractability of L&E's utility maximization by representing other values in terms either of individual preferences or of add-ons to total social utility are difficult to calibrate empirically, do not direct account for the nuanced ways in which people care about the utility of others, and tend to frame policy debates in terms of "trade-offs" between total utility and other values.

Overall, by attempting to maintain L&E's basic framework–and analytical tractability–these tweaks continue to prioritize wealth maximization over other normative considerations. Moreover, adopting a more realistic model of individual behavior or a more complicated social welfare function deprives L&E of some of its main selling points–analytical tractability and simplicity–requiring further approximations elsewhere and reducing the range of questions which the model can address.

One example of an L&E regulatory design approach that has not aged well comes from privacy law. Early L&E literature argued that laws restricting use of personal information are suspect because they prevent allocative efficiency in the market [7]. While later work in Economics has challenged this view and provided more sophisticated models [51, 59, 60], these upgrades have not been sufficient to deal with the changing conditions of e-commerce and the breadth of social concerns implicated in digital technology's use of personal information. Alternative frameworks for understanding privacy in terms of contextually defined norms [61] and systemic financial risk [62] suggest that legal scholarship could benefit from ABMs, perhaps adapted from those modeling norm formation [63–66] and bank stress testing [67, 68, 69], and enforcement policies [70].

[71], provides a direct comparison of ABM and traditional L&E approaches to the problem of devising tort law standards for accident compensation. While some findings are consistent between the two, the ABM demonstrates that when agents learn the rules from experience, they behave differently from the neoclassical predictions. Among other observations, results vary depending on how the duty of care is specified. Agents sometimes are careful even when they would not be liable under the applicable standard of negligence, but sometimes continue to behave negligently for long periods of time; and individual agents generally continue to experiment with safer and riskier strategies long after overall system variables appear to have settled at equilibrium values. This sort of heterogeneous experience of individual agents may affect long-term wealth distributions and might plausibly be relevant to the design of tort law.

ABM AND THE COMPLEXITY CRITIQUE

Fundamentally, even the most "souped up" L&E approaches tend to take an individualistic transaction-by-transaction perspective, implicitly assuming that society's response to legal rules will approximate a linear cumulation of these assessments, where deviations can be handled by incremental course corrections. Unfortunately, complexity theory suggests that L&E's faith in linearity and incrementalism is likely to be misguided. The networked nature of social systems makes non-linear cumulative effects and feedback between transactions not only likely, but increasingly so as technology increases global interconnectedness. As Miller and Page [72] explain:

Complexity arises when the dependencies among the elements become important. In such a system, removing [or altering] one such element destroys system behavior to an extent that goes well beyond what is embodied by the particular element that is removed [or altered]. (p. 9).

Complex systems are known to exhibit phenomena, such as phase transitions, tipping points and metastability, that have dramatic non-linear effects. Computational simulation has been the primary tool for studying these systems. When the world is complex, a head in the sand insistence on locally tractable models simply will not do. Indeed, the increasing reliance on ABM in the social sciences is a direct consequence of the recognition of complexity's importance in social systems.

Our unfortunate recent experiences with the mismanagement of the COVID-19 pandemic [73] and the spread of electoral misinformation highlight the importance of devising regulatory approaches responsive to complex phenomena. It is thus increasingly urgent for legal scholars to attend to those who have long been sounding the alarm about the challenges that complexity poses for law. (For overviews of complexity science and its applications in law, see [4, 74–81]). To date, the legal scholarly response to "complexity science" (also called "complex adaptive systems") has focused mostly on environmental law and financial systems, where the focus has understandably been on designing legal structures and institutions to avoid disastrous tipping points. Other systemic problems, such as the persistence of racial inequality, suggest that society has become stuck in a metastable state that cannot be escaped through incremental improvements. Though such metastability is to be expected in the social welfare landscape of a complex society, even less attention has been paid so far to the question of how law might be used to facilitate socially desirable systemic change. Moreover, as our above survey of the literature suggests, to date the legal literature addressing complex systems has mostly tried to spin out the implications of general observations about the nature of complex systems, rather than ABMs tailored to the problem under consideration. This translational work is extremely valuable, but more could be learned from more targeted ABM projects.

For example, Malcai and Shur-Ofry [74] point out that the conceptual toolbox of complexity theory can illuminate a longstanding, and polarized, debate about whether to apply cost-benefit analysis or a more constraining "precautionary principle" in shaping environmental regulation, particularly with regard to climate change. They contend that a complexity-based approach can alleviate concerns that the precautionary principle is insufficiently sensitive to the costs of environmental precautions by "delineating several factors, which may serve as guidelines for the principle's application: phenomena that spread exponentially, in short time-scales, and pose systemic, existential, risk." With that basic insight in hand, more specifically tailored ABM could help to further delineate these (and possibly other) factors and provide further guidance as to when they are likely to arise in real-world systems.

The stream of applications of ABM to residential segregation, discussed earlier, illustrates the value of this approach, as well as the unique perspective that legal scholars can bring to these questions. While sociologists have focused on understanding how small amounts of racial bias and preference for homophily can result in drastic segregation, legal scholars' contributions [22, 43] emphasize the lasting effects of prior legal tolerance of enforced segregation and critique the way that current anti-discrimination law fails to address these "lock-in" effects. These two effects both arise from feedback effects characteristic of complex social systems, but they are quite different and might suggest different regulatory responses. To understand and address current housing segregation problems, both of these perspectives (and others) are undoubtedly needed.

DISCUSSION

While the number of scholarly articles discussing law and complexity is dwarfed by the law and economics literature, it is by no means negligible today. Nonetheless, as discussed above, ABM remains largely absent from the legal literature. As noted, however, we see two sorts of reasons—theoretical and practical–to be optimistic about the potential for growth in legal scholarly attention to ABM methods and results.

Theoretical Demand for Consideration of Law's Structural and Systemic Effects

Two recent developments in legal theory, framed as critiques of L&E and highlighting current societal problems, draw attention to precisely the sort of systemic, structural and dynamic effects that ABM may be able to model. These movements, termed "Law and Macroeconomics" and "Law and Political Economy" by their proponents, are still relatively nascent, but seem to be gaining traction among legal scholars. Moreover, both movements have so far lodged their critiques at a relatively abstract and theoretical level. ABM provides a method for translating at least some of these systemic critiques of L&E into actionable insights about legal and regulatory design that can be compared to and juxtaposed with those of traditional L&E. While ABM does require somewhat simplified models, it can avoid many of the most troublesome simplifying assumptions of traditional L&E, while maintaining the many advantages of using well-specified models to test and deepen qualitative insights.

Law and Macroeconomics

[82] proposal for "Law and Macroeconomics" argues that traditional L&E "should really be called 'law and microeconomics,'" because of its inability to reckon with aggregate level constraints, such as financial recessions and the business cycle, that appear in macroeconomics. He argues that legal analysis should be more responsive to macroeconomic considerations, at least under some conditions. Meanwhile, the 2008 financial crisis was a wake-up call for macroeconomics, which has since revisited its assumptions about the connections between individual behavior and larger patterns and effects. Whereas macroeconomic models formerly depended on a "representative agent" that was somehow both a single agent and an average of all agents in the economy, now the field is moving toward heterogeneous agent modeling (HAM), explicitly modeling a variety of agents with ex ante and ex post differences [83]. HAM methods combine the classic economic tool of dynamic stochastic programming [84] with the flexibility of ABMs. ABM macro methods have found some traction in federal policy-making through research aimed at studying financial risk [68]. Further incorporation of ABM into relevant legal scholarship would be both a natural outgrowth of Listokin's call for a "Law and Macroeconomics" approach and a method for conducting such studies.

Law and Political Economy

A new scholarly movement calling itself Law and Political Economy (LPE) has recently combined a number of earlier critiques of L&E to contend that L&E is reflective of a "[n]eoliberal political economy, with its underlying commitments to efficiency, neutrality, and antipolitics, [that] helped animate, shape, and legitimate a twentieth-century consensus that erased power, encased the market, and reinscribed racialized, economic, and gendered inequities." [11] LPE scholars see a need to correct L&E's erosion of antitrust, intellectual property, and environmental law (for example) through "a legal imaginary of democratic political economy, that takes seriously underlying concepts of power, equality, and democracy" to "amplify and accelerate [recent] movements for structural reform."

One need not believe that computational models can account for all of these scholars' criticisms of L&E to expect that the wider modeling scope made possible through ABM techniques can help to illuminate the effects of separated markets and power imbalances, take into account the endogenous effects of the market on the law, and model various ways in which law and legal institutions can promote social values such as equality and democracy, rather than optimizing a linear representation of the atomistic preferences of individuals. These models would necessarily have a different view of social structure, taking less for granted and acknowledging new forms of social (in)stability and transformation.

Practical Developments Favoring Law and ABM

In addition to the current demand by legal scholars for ways to account for a broader set of normative values and systemic effects in evaluating legal and regulatory proposals, several more mundane developments favor greater use of ABM in legal scholars. Whatever L&E's failings and limitations, several decades of L&E scholarship, along with other influences such as the emergence of the discipline of technology law, have created a significant cadre of legal scholars with the capacity to engage in mathematical and computational modeling, either alone or in interdisciplinary collaboration, as well as a much larger group that is now prepared to read, discuss and critique the resulting applications to legal questions. While some legal scholars are equipped to take on ABM projects alone, bringing such efforts to the wide range of legal arenas involving significant complexity is likely to require interdisciplinary collaborations between legal scholars and social scientists. Fortunately, as a result of various "law and … " approaches and of the growing importance of technological understanding for regulatory design, interdisciplinary collaboration is now entirely unremarkable for legal scholars (at least in the U.S.) Moreover, the growing use of ABM methods in social science means that many new legal scholars will have been exposed to these techniques during their undergraduate or graduate studies prior to entering law school. The spread of ABM methods also means that open source software packages are now widely available for running fairly sophisticated simulations without deep programming expertize. This development, along with the widespread availability of cheap computing resources that are powerful enough to run meaningful simulations (indeed, a laptop will often suffice) reduces barriers to entry for this type of research.

In sum, the potential for ABM to contribute to progress on important issues in the evaluation and design of proposals for law and regulation is high in light of the complex problems confronting today's society. Fortunately, for both theoretical and practical reasons, the ground is much more fertile for adoption of ABM methods than it has been in the past.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

FUNDING

One of the authors of this article is supported by the NYU Information Law Institute's Fellows program, which is funded in part by Microsoft Corporation. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

ACKNOWLEDGMENTS

KS acknowledges summer research funding from the Filomen D. Agostino and Max E. Greenberg Research Fund. This work benefitted greatly from conversations with Joshua Epstein, Erez Hatna, and Salome Viljoen and participants in the Physics of Law conference (November, 2020, online).

REFERENCES

1. Citron DK, and West R. *On Legal Scholarship, Current Issues in Legal Education* (2014). Available at: https://scholarship.law.bu.edu/shorter_works/75. doi:10.4159/harvard.9780674735613
2. West R. The Contested Value of Normative Legal Scholarship. *J Leg Edu* (2016) 66(1):6–17.
3. Muchmore AI. Uncertainty, Complexity, and Regulatory Design. *Hous L Rev* (2015) 53:1321.
4. Posner RA. *Economic Analysis of Law*. Seventh ed. Austin, TX: Wolters Kluwer (2007).
5. Calabresi G. *The Future of Law and Economics: Essays in Reform and Recollection*. Yale University Press (2016). doi:10.12987/yale/9780300195897.001.0001
6. Cooter R, and Ulen T. *Law & Economics*. 6 Ed. Addison-Wesley (2016).
7. Epstein JM. *Generative Social Science: Studies in Agent-Based Computational Modeling*, 13. Princeton University Press (2006).
8. Wilensky U, and Rand W. An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo. *Mit Press* (2015).
9. Gilbert N, and Troitzsch K. *Simulation for the Social Scientist*. UK: McGraw-Hill Education (2005).
10. Britton-Purdy J, Grewal DS, Kapczynski A, and Rahman KS. Building a Law-And-Political-Economy Framework: Beyond the Twentieth-century Synthesis. *Yale LJ* (2019) 129:1784.
11. Posner RA. The Economics of Privacy. *Am Econ Rev* (1981) 71(2):405–9.
12. Hägerstrand T. A Monte Carlo Approach to Diffusion. *Arch Europ Sociol* (1965) 6(1):43–67. doi:10.1017/s0003975600001132
13. Schelling TC. Dynamic Models of Segregation†. *J Math Sociol* (1971) 1(2): 143–86. doi:10.1080/0022250x.1971.9989794
14. Hamill L, and Gilbert GN. *Agent-based Modelling in Economics*. Chichester: Wiley (2016).
15. Fagiolo G, and Roventini A (2016). Macroeconomic Policy in DSGE and Agent-Based Models Redux: New Developments and Challenges Ahead. SSRN 2763735.
16. Gatti DD, Fagiolo G, Gallegati M, Richiardi M, and Russo A, editors. *Agent-based Models in Economics: A Toolkit*. Cambridge University Press (2018).
17. L Tesfatsion and KL Judd, editors. Handbook of Computational Economics. *Agent Based Computational Economics*, Vol. 22. Elsevier (2006). p. 829–1660.
18. Dawid H, Harting P, and Neugart M. Economic Convergence: Policy Implications from a Heterogeneous Agent Model. *J Econ Dyn Control* (2014) 44:54–80. doi:10.1016/j.jedc.2014.04.004
19. Hatna E, and Benenson I. Combining Segregation and Integration: Schelling Model Dynamics for Heterogeneous Population. *J Artif Societies Soc Simulation* (2015) 18(4):15. Available at: http://jasss.soc.surrey.ac.uk/18/4/15.html. doi:10.18564/jasss.2824
20. Hatna E, and Benenson I. The Schelling Model of Ethnic Residential Dynamics: Beyond the Integrated - Segregated Dichotomy of Patterns. *J Artif Societies Soc Simulation* (2012) 15(1):6. Available at: http://jasss.soc.surrey.ac.uk/15/1/6.html. doi:10.18564/jasss.1873
21. Benenson I, Omer I, and Hatna E. Entity-Based Modeling of Urban Residential Dynamics: The Case of Yaffo, Tel Aviv. *Environ Plann B Plann Des* (2002) 29(4):491–512. doi:10.1068/b1287
22. Roithmayr D. Locked in Segregation. *Va. J Soc Pol'y L* (2004) 12:197. doi:10.2139/ssrn.587963
23. Sahasranaman A, and Jensen HJ. Ethnicity and Wealth: The Dynamics of Dual Segregation. *PloS one* (2018) 13(10):e0204307. doi:10.1371/journal.pone.0204307
24. Benenson I, and Hatna E. Minority-Majority Relations in the Schelling Model of Residential Dynamics. *Geographical Anal* (2011) 43(2011):287–305. doi:10.1111/j.1538-4632.2011.00820.x
25. Tassier T, and Menczer F. Social Network Structure, Segregation, and equality in a Labor Market with Referral Hiring. *J Econ Behav Organ* (2008) 66(3-4): 514–28. doi:10.1016/j.jebo.2006.07.003
26. Neugart M, and Richiardi M. Agent-based Models of the Labor Market. *Laboratorio R Revelli Working Papers Series* (2012) 125:164–212. doi:10.1093/oxfordhb/9780199844371.013.44
27. Rubineau B, and Fernandez RM. Missing Links: Referrer Behavior and Job Segregation. *Manag Sci* (2013) 59(11):2470–89. doi:10.1287/mnsc.2013.1717
28. Reardon SF, Baker R, Kasman M, Klasik D, and Townsend JB. What Levels of Racial Diversity Can Be Achieved with Socioeconomic-Based Affirmative Action? Evidence from a Simulation Model. *J Pol Anal Manage* (2018) 37(3):630–57. doi:10.1002/pam.22056
29. Picker RC. Simple Games in a Complex World: A Generative Approach to the Adoption of Norms. *Univ Chicago L Rev* (1997) 64(4):1225–88. doi:10.2307/1600216
30. Axtell RL, Epstein JM, and Young HP. The Emergence of Classes in a Multiagent Bargaining Model. *Soc Dyn* (2001) 27:191–211. doi:10.7551/mitpress/6294.003.0009
31. Weisbuch G. Persistence of Discrimination: Revisiting Axtell, Epstein and Young. *Physica A: Stat Mech its Appl* (2018) 492:39–49. doi:10.1016/j.physa.2017.09.053
32. Allen HJ. The SEC as Financial Stability Regulator. *Iowa J Corp L* (2018) 43:715.
33. Amos AL. Developing the Law of the River: The Integration of Law and Policy into Hydrologic and Socio-Economic Modeling Efforts in the Willamette River Basin. *U Kan L Rev* (2014) 62:1091. doi:10.17161/1808.20266
34. Cherry BA. The Telecommunications Economy and Regulation as Coevolving Adaptive Systems: Implications for Federalism, Fed. *Comm L.J* (2007) 59:369. doi:10.1353/tj.2007.0134
35. DeVries P. The Resilience Principles: A Framework for New ICT Governance. *J Telecomm High Tech L* (2011) 9:137. doi:10.2139/ssrn.1732798
36. Coglianese C, and Lehr D. Regulating by Robot: Administrative Decision Making in the Machine Learning Era. *Geo L.J* (2017) 105:1147.
37. Gerger DJ. Method, Community & Comparative Law: An Encounter with Complexity Science. *Roger Williams U L Rev* (2011) 16:110.
38. Olson NE. Minding the Gap: A Ten-step Program for Better Tax Compliance. *Stan L Pol'y Rev* (2009) 20:7.
39. Cohen PN. A Review of Genetics and Race in the Social Sciences. *Annals* (2015) 661:51. doi:10.1177/0002716215587673
40. Diaz RS. Getting to the Root of Environmental Injustice: Evaluating Claims, Causes and Solutions. *Geo Envtl L Rev* (2017) 29:767.
41. Hodgkinson T, Saville G, and Andresen MA. The Diffusion of Detriment: Tracking Displacement Using a City-wide Mixed Methods Approach. *Br J Criminol* (2020) 60(1):198–218. doi:10.1093/bjc/azz025
42. Kidder WC. How Workable Are Class-Based and Race-Neutral Alternatives at Leading American Universities? *UCLA L Rev Disc* (2016) 64:100.
43. Roithmayr D. Them that Has, Gets. *Miss C.L Rev* (2007) 27:373.
44. Worster WT. The Transformation of Quantity into Quality: Critical Mass in the Formation of Customary International Law. *B.U Int'l L.J* (2013) 31:1. doi:10.2139/ssrn.2012452
45. Wylly P. Evaluating the Costs of Technology Neutrality in Light of the Importance of Social Network Influences and Bandwagon Effects for Innovation. *N.Y.U Entl L.J* (2015) 23:298.
46. Kalbfeld JR. Critical Mass for Affirmative Action: Dispersing the Critical Cloud. *L Soc Rev* (2019) 53:1266–304. doi:10.1111/lasr.12441

47. Koliba C, Zia A, Schroth A, Bomblies A, Van Houten J, and Rizzo D. Lake Champlain Basin as a Complex Adaptive System: Insights from the Research on Adaptation to Climate Change ("RACC") Project. *Vt J Envtl L* (2016) 17:533.
48. Manhire JT. There Is No Spoon: Reconsidering the Tax Compliance Puzzle. Fla. *Tax Rev* (2015) 17:623. doi:10.2139/ssrn.2459644
49. Ota S. The Perspectives of Agent Based Models in Law. *Sociological Theor Methods* (2004) 19:53–65.
50. Schwartz A. Agent-based Modeling for Legal Studies. In: *Computational Legal Studies*. Edward Elgar Publishing (2020).
51. Hui KL, and Png IP. Economics of Privacy. In: HANDBOOK OF INFORMATION SYSTEMS AND ECONOMICS (2005).
52. Johnson SD, and Groff ER. Strengthening Theoretical Testing in Criminology Using Agent-Based Modeling. *J Res Crime Delinquency* (2014) 51(4):509–25. doi:10.1177/0022427814531490
53. Hicks JR. The Foundations of Welfare Economics. *Econ J* (1939) 49(196): 696–712. doi:10.2307/2225023
54. Kaldor N. Welfare Propositions of Economics and Interpersonal Comparisons of Utility. *Econ J* (1939) 49(195):549–52. doi:10.2307/2224835
55. Parisi F. *The Language of Law and Economics: A Dictionary*. Cambridge University Press (2013). doi:10.1017/cbo9781139034043
56. Frantz R., Chen S. H., Dopfer K., Heukelom F., and Mousavi S. *Routledge handbook of behavioral economics*. Abingdon: Routledge (2017).
57. Sunstein CR. The Real World of Cost-Benefit Analysis: Thirty-Six Questions (And Almost as Many Answers). *Colum L Rev* (2014) 114:167. doi:10.2139/ssrn.2199112
58. Sunstein CR. The Limits of Quantification. *Cal L Rev* (2014) 103:1369. doi:10.2139/ssrn.2424878
59. Hermalin BE, and Katz ML. Privacy, Property Rights and Efficiency: The Economics of Privacy as Secrecy. *Quant Market Econ* (2006) 4(3):209–39. doi:10.1007/s11129-005-9004-7
60. Cofone IN. Nothing to Hide, but Something to Lose. *U Toronto L.J* (2020) 70: 64–90. doi:10.3138/utlj.2018-0118
61. Nissenbaum H. *Privacy in Context*. Stanford University Press (2020).
62. Benthall S, and Viljoen S. Data Market Discipline: From Financial Regulation to Data Governance. *J Int'l Comp L* (2021). doi:10.2139/ssrn.3774418
63. Axelrod R. The Dissemination of Culture. *J conflict resolution* (1997) 41(2): 203–26. doi:10.1177/0022002797041002001
64. Hoffmann MJ. *Entrepreneurs and Norm Dynamics: An Agent-Based Model of the Norm Life Cycle.* Penn's Agent-Based Modeling Laboratory (PAMLA) (2003). Available at: http://www. polisci. upenn. edu/abir/private/Pamla/Hoffmann norms. Doc (Accessed June 7, 2021)
65. Wang Y, Xue H, Chen HJ, and Igusa T. Examining Social Norm Impacts on Obesity and Eating Behaviors Among US School Children Based on Agent-Based Model. *BMC public health* (2014) 14(1):1–11. doi:10.1186/1471-2458-14-923
66. Hokamp S. Dynamics of Tax Evasion with Back Auditing, Social Norm Updating, and Public Goods Provision - an Agent-Based Simulation. *J Econ Psychol* (2014) 40:187–99. doi:10.1016/j.joep.2013.01.006
67. Aymanns C, Farmer JD, Kleinnijenhuis AM, and Wetzer T. Models of Financial Stability and Their Application in Stress Tests. In: *Handbook of Computational Economics*, 4. Elsevier (2018). p. 329–91. doi:10.1016/bs.hescom.2018.04.001
68. Balke T, De Vos M, and Padget J. I-ABM: Combining Institutional Frameworks and Agent-Based Modelling for the Design of Enforcement Policies. *Artif Intell L* (2013) 21:371–98. doi:10.1007/s10506-013-91410.1007/s10506-013-9143-1
69. Iori G, and Mantegna RN. Empirical Analyses of Networks in Finance. In: *Handbook of Computational Economics*, 4. Elsevier (2018). p. 637–85. doi:10.1016/bs.hescom.2018.02.005
70. Bookstaber R, Paddrik M, and Tivnan B. An Agent-Based Model for Financial Vulnerability. *J Econ Interact Coord* (2018) 13(2):433–66. doi:10.1007/s11403-017-0188-1
71. Seagren CW. *Evolutionary Agents and Dynamic Analysis of Accident Law* (2014). Available at https://core.ac.uk/download/pdf/36740722.pdf (Accessed October 7, 2021).
72. Miller JH, and Page SE. *Complex Adaptive Systems: An Introduction to Computational Models of Social Life*, 17. Princeton university press (2009).
73. Malcai O, and Shur-Ofry M. Using Complexity to Calibrate the Legal Response to COVID-19. *Front Phys Soc Phys* (2021). doi:10.2139/ssrn.3763376
74. Ruhl JB. Complexity Theory as a Paradigm for the Dynamical Law-And-Society System: A Wake-Up Call for Legal Reductionism and the Modern Administrative State. *Duke L J* (1996). 45:849. doi:10.2307/1372975
75. Ruhl JB. Governing cascade Failures in Complex Social-Ecological Technological Systems: Framing Context, Strategies, and Challenges. *Vanderbilt J Ent Tech L* (2020) 22(2):407.
76. Ruhl JB. Law's Complexity: a Primer. *Ga St. L. Rev.* (2008) 24:885.
77. Ruhl JB. Measuring, Monitoring and Managing Legal Complexity. *Iowa L Rev* (2015) 100:191.
78. Shur-Ofry M. IP and the Lens of Complexity. *IDEA* (2013) 54:55.
79. Shur-Ofry M, and Malcai O. Institutions for Collective Action and Social Contagion – Community Garden Case Study. *Regul Governance* (2021) 15(1): 63. doi:10.1111/rego.12256
80. Ruhl JB. Managing Systemic Risk in Legal Systems. *Indiana L.J.* (2014) 89(2): 559. doi:10.2139/ssrn.2212212
81. Wheatley SC, Murray J, and Webb T. *Complexity Theory and Law, Mapping an Emergent Jurisprudence*. Taylor & Francis (2018).
82. Listokin Y. *Law and Macroeconomics: Legal Remedies to Recessions*. Harvard University Press (2019). doi:10.4159/9780674239838
83. C Hommes and L Blake, editors. Handbook of Computational Economics, Vol 4: Heterogeneous Agent Modeling. *Handbook of Computational Economics*, 4. Elsevier (2018). p. 2–796.
84. Stachurski J. *Economic Dynamics: Theory and Computation*. MIT Press (2009). doi:10.1163/9789042026452 CrossRef Full Text

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 21 June 2021
doi: 10.3389/fphy.2021.665054

Check for updates

Quantifying Legal Entropy

*Ted Sichelman**

School of Law, University of San Diego, San Diego, CA, United States

Many scholars have employed the term "entropy" in the context of law and legal systems to roughly refer to the amount of "uncertainty" present in a given law, doctrine, or legal system. Just a few of these scholars have attempted to formulate a quantitative definition of legal entropy, and none have provided a precise formula usable across a variety of legal contexts. Here, relying upon Claude Shannon's definition of entropy in the context of information theory, I provide a quantitative formalization of entropy in delineating, interpreting, and applying the law. In addition to offering a precise quantification of uncertainty and the information content of the law, the approach offered here provides other benefits. For example, it offers a more comprehensive account of the uses and limits of "modularity" in the law—namely, using the terminology of Henry Smith, the use of legal "boundaries" (be they spatial or intangible) that "economize on information costs" by "hiding" classes of information "behind" those boundaries. In general, much of the "work" performed by the legal system is to reduce legal entropy by delineating, interpreting, and applying the law, a process that can in principle be quantified.

Keywords: entropy, indeterminacy, legal entropy, information theory, modularity, Hohfeld, legal complexity, legal uncertainty

OPEN ACCESS

Edited by:
Daniel Martin Katz,
Illinois Institute of Technology,
United States

Reviewed by:
Marika Taylor,
University of Southampton,
United Kingdom
Michael Bommarito,
Consultant, Lansing, United States

***Correspondence:**
Ted Sichelman
tsichelman@sandiego.edu

Specialty section:
This article was submitted to
Social Physics,
a section of the journal
Frontiers in Physics

Received: *07 February 2021*
Accepted: *27 April 2021*
Published: *21 June 2021*

Citation:
Sichelman T (2021) Quantifying Legal Entropy. Front. Phys. 9:665054.
doi: 10.3389/fphy.2021.665054

INTRODUCTION

It goes without saying that the law and legal systems are uncertain to a significant degree. Several scholars (e.g., Katz and Bommarito [1]; Friedrich et al. [2]) have attempted to determine the uncertainty (and related complexity) of legal systems by formulating measures of the "entropy" of words in legal texts, including statutes and other legal authorities. Although measuring the ambiguity of words in texts can be valuable in many situations, it does not provide a comprehensive measure of the uncertainty in interpreting legal rules, much less a "system-wide" measure of the uncertainty of the law and legal system and subsystems more generally. Other scholars (e.g., Dworkin [3], Parisi [4, 5], Ruhl and Ruhl [6]) have focused their efforts on more general notions of legal entropy and related concepts, but have done little to nothing to formalize those notions in mathematical terms.

This article provides several important contributions to the literature by formalizing the notion of legal entropy. First, it offers a conceptual framework to quantify the entropy of legal systems that extends beyond legal text to capture how the law actually functions in real-world situations, including not only legal interpretation, but also the entropy and related information costs in formulating and applying the law. Second, although some previous works have foreshadowed the possibility of a quantitative description of legal entropy (e.g., D'Amato [7]), the formalization offered here provides a fully mathematical formulation as it applies to legal systems and disputes. Third, the mathematical model proposed here offers a potential template for how legal AI systems can measure and store information about the uncertainty of legal systems. Fourth, the model helps to explain more fully the nature and function of important concepts in the law, including the so-called "modularization" of the law and legal concepts, as proposed in the seminal works on the topic by Smith [8–10] and follow-on works by others (e.g., Newman [11]), as a well as the Coase Theorem [12] and the indeterminacy of legal rules [13].

The article proceeds as follows. Part 2 provides a brief background of the notion of entropy in physics and information theory, particularly Shannon's [14] formulation of information entropy. Part 3 describes previous attempts to describe legal entropy, including descriptive notions of legal entropy and measures of the word entropy found in legal texts. It explains the limitations inherent in these previous treatments. In Part 4, relying on Shannon [14] and the theoretical work of Hohfeld [15], it introduces a formal mathematical description of legal entropy, as it applies to a particular legal issues and disputes as well as across legal systems and subsystems. Specifically, Part 4 proposes models for quantifying entropy in formulating and interpreting the law, as well as applying the law to a set of facts. In Part 5, the article applies its formal model to important theoretical and practical issues in the law, including legal indeterminacy, modularity, and the Coase Theorem. In so doing, it discusses practical problems in "measuring" legal entropy. Finally, the article concludes with some suggestions for further research.

ENTROPY IN PHYSICS AND INFORMATION THEORY

The concept of entropy in physics traces to the work of Clausius [16] in the mid-nineteenth century to describe a property of the transfer of heat, ΔQ, from a heat source at a certain temperature, T, to an idealized engine in a so-called reversible process.[1] In this situation, according to Clausius, the entropy of the system increases by $\Delta Q/T$. Similarly, entropy decreases by such an amount when an idealized engine loses heat ΔQ to a heat sink at temperature T. In other words, as heat enters a thermodynamic system, entropy increases—particularly, if the system is cold, less if the system is already hot.

In the 1870s, Boltzmann [17] offered a molecular (i.e., microscopic) description of Clausius's notion of entropy. Specifically, Boltzmann [17] postulated that Clausius's macroscopic description of entropy could be explained in relation to microscopic states. Because heat at a macroscopic level is essentially a "disordered" collection of microscopic particles, the exact behavior or which is unknown at the macroscopic level, the entropy of the system can be viewed roughly as a measure of macroscopic disorder. As a cold system becomes hotter, its ordered, stable microscopic state of particles in fixed positions yields to a frenzy of quickly moving particles. Although in a classical system, the position and momentum of microscopic particles is measurable in principle, merely measuring the temperature and other macroscopic properties of a system would be insufficient to determine the precise position and momentum of each and every particle. As more heat enters a system, the more difficult it becomes to use macroscopic measures to determine the position and momentum of each particle that the system comprises. This increasing uncertainty results because the microscopic particles could be in a greater number of potential states (i.e., of position and momentum) in an increasingly hotter system than an increasingly colder system, where particles are relatively motionless. If a system is already hot, introducing a bit more heat increases the uncertainty of the microscopic states much less than if the system begins cold.

Boltzmann [17] was able to formulate a microscopic definition of entropy along these lines, $S = k_b \ln W$, that explained Clausius's macroscopic definition. According to Boltzmann's Equation, the entropy S of an ideal gas is simply the natural logarithm (introduced for mathematical convenience) of the number of microstates, W, of the system corresponding to the gas's macrostate multiplied by a constant k_b (Boltzmann's constant).[2] In other words, the number of different position and momentum arrangements the microscopic particles may occupy for a given macroscopic state effectively explains the macroscopic entropy of the system.

Boltzmann [18] and later Gibbs [19] generalized his equation to take into account that certain microstates are more or less probable than others. In this case, weighting is necessary to take account of the variable probability of certain microstates that correspond to a given macrostate. In this case, using a well-known mathematical approximation, the so-called Gibbs entropy becomes:

$$S = -k_B \sum p_i \ln p_i \tag{1}$$

Roughly speaking, the Gibbs entropy reflects a weighted average of the number of microstates corresponding to a given macrostate, where p_i is the probability for a certain microstate to occur. In this regard, note that $\ln p_i$ (for $p_i < 1$) is negative and decreases as p_i approaches zero. Thus, a state i with a seemingly small probability of occurrence may significantly contribute to overall entropy.

The Gibbs entropy is in effect a special case of a more general phenomena in which some "macroscopic" state of a generalized system, call it M, may be instantiated by W different "microscopic" states of the generalized system (Jaynes [20]). For instance, the "macroscopic" state of having 10 cents in one's hand can be instantiated by four "microscopic" states: (1) 10 pennies; (2) a nickel and 5 pennies; (3) two nickels; or (4) one dime. In other words, knowing the "macroscopic" state (here, the total monetary value) generally will be insufficient to specify the "microscopic" state (here, the precise coins used to achieve the total monetary value).

The greater the uncertainty in microscopic configuration, the greater the entropy. In other words (and using log base 2 to capture the number of bits of entropy), $S = k \log_2 W$. Setting $k = 1$ for simplicity, and reducing W in the same manner as

[1] The discussion in this section is designed to offer a concise and simplified qualitative background of the notion of entropy in physics and information theory in order to set the stage for the following discussion of legal entropy, and thus should not be viewed as a precise technical account.

[2] Because the $\log(AB) = \log A + \log B$, by defining the entropy in terms of a logarithm (such as the natural logarithm), it becomes simpler to calculate entropy as the number of microstates increases, particularly when two systems are combined.

the Gibbs entropy, one arrives at the formula for Shannon [14] entropy in bits:

$$S = -\sum p_i \log_2 p_i \quad (2)$$

If we suppose that each of the four microstates in our example are equally likely, the Shannon entropy of 10 cents is $-4 \times 0.25 \times \log_2 (0.25) = 2$ bits of information.[3] This is sensible since there are only four choices, which we can label 00, 01, 10, or 11. This more generalized notion of entropy as the uncertainty over a range of informational microstates, as reflected in the Shannon entropy, will play a central role in quantifying legal entropy herein.

PREVIOUS TREATMENTS OF ENTROPY IN LEGAL SYSTEMS

Many scholars have applied the concept of entropy to legal systems. All of these treatments can be classified into two categories: (1) metaphorical uses of the concept of entropy; and (2) uses of formal mathematical and physical definitions to measure the "entropy" of legal texts.

Although the first category of scholarship (metaphor) can often be useful in thinking about the disorder, complexity, and uncertainty present in legal systems, it fails to offer any formal quantification of legal entropy. For instance, in a well-known article, Parisi [4] contends that real property is subject to a fundamental law of entropy that leads to increasing fragmentation of property interests, but fails to quantify the notion. Lewis [22] applies thermodynamic principles, including entropy, to the explain corporate reorganizations but, like other treatments, does not extend his notions beyond the level of metaphor. Ferrara and Gagliotti [23] purport to develop a conceptual "mathematical" approach to the law, including a notion that has "somewhat to do with the concept" of entropy in information theory, but their scheme is devoid of formal definition and thus reduces to metaphor. Ultimately, all previous treatments of the broad of concept of legal entropy (see also Berg [24], Edgar [25], Fromer [26], King [27], Moran [28], Stephan [29], as examples) fail to quantify the notion.[4]

Perhaps the treatment that comes closest to any quantification of legal entropy is that of D'Amato [7], who recognizes that entropy is at a maximum when the outcome of a legal dispute is equally likely for each part, further remarking that "[i]n order to use entropy in law directly, the legal scientist would have to embed the collection of predictions we call law into an abstract space that exhibited the variations in the level of uncertainty of the predictions." Yet, immediately following this insight, D'Amato [7] states that "Since law cannot be completely transcribed into words, it cannot be transcribed into symbols and spaces either." D'Amato's [7] statement is deficient in two important respects. First, to the extent one is concerned about the probabilities of outcomes in legal disputes, as is D'Amato [7], although it may be practically difficult to "transcribe" disputes into probability spaces, it is not impossible. Indeed, attorneys regularly the estimates the odds of winning and losing cases. Moreover, recent developments in legal artificial intelligence have vastly expanded the promise of more automated approaches to predicting legal outcomes (e.g., Katz et al. [32]; Branting et al. [33]). Second, aside from practical interest, it may be theoretically illuminating to devise mathematical models of the operation of the law. In this regard, such theoretical modeling is in turn arguably critical to practical advances in legal artificial intelligence.

The second category of articles relies on measures of entropy from computational linguistics and related fields, typically derived from Shannon [14, 34], to measure the uncertainty or ambiguity that is present in the text of statutes, regulations, and legal documents. For instance, Katz and Bommarito [1] measure legal complexity based on linguistic entropy present in U.S. federal statutes. In a similar vein, Friedrich et al. [2] examine the word and document entropy of opinions from the U.S. Supreme Court and the German Bundegerichtshof in order to measure and compare the textual ambiguity present in the courts' opinions.[5]

Although these text-based endeavors are important contributions to the literature, especially by formalizing previous metaphorical treatments, they are limited to what I term the "interpretive entropy" of legal systems—namely, the entropy and associated information costs involved in interpreting the law prior to its application to a particular set of facts. Moreover, because this scholarship tends to focus on the language of statutes and regulations, it measures only a portion of the interpretive entropy, because interpretation also involves the consultation of authoritative legal opinions, administrative interpretations, legislative and regulatory history, the text of other statutes and regulations, and not infrequently, general facts about the world (e.g., social norms, scientific facts, etc.).[6]

[3] Another way to conceptualize Shannon entropy is in terms of "surprise," which typically is defined as the unlikelihood of an event occurring, i.e., 1/p [21]. Since the log (1/p) = –log p, we can rewrite Shannon entropy as proportional to the sum over states of $p_i \log_2 (1/p_i)$. Thus, information entropy is driven by a combination of the logarithm of the level of surprise (i.e., improbability) of a given microstate and the probability of the microstate occurring, summed across all microstates.

[4] Loevinger [30] cites Shannon and Weaver [31] and offers an "equation," which is best characterized as tongue-and-cheek. Namely, Loevinger [30] states: "The second law of sociodynamics is the law of the conservation of entropy. Entropy, in social as in physical phenomena, is a measure of disorder, uncertainty or confusion. The law of the conservation of entropy in sociodynamics states that the amount of entropy concerning any social problem remains constant regardless of the number of agencies or entities to which it is referred while the time required for decision or action on the problem increases in geometrical proportion to the number of agencies or entities whose concurrence is required. This law can be expressed as $T = NC^2$, where "W" is the time required for decision or action and 'NC' is the number of agencies or entities whose concurrence is required."

[5] Other studies use legal documents as inputs and measure entropy unrelated to legal entropy. For instance, Zhang et al. [35] extend the application of Shannon entropy from text to patent indicators, including citation counts, number of patent families, and similar indicators, to measure the importance of particular patents in technological innovation. Although such approaches may be useful for determining the economic "information" content and, hence, the economic importance signified by a particular legal document, they do not measure legal entropy, that is, the uncertainty or ambiguity of a legal document or broader legal relation within the legal system.

[6] As D'Amato [7] insightfully remarks, "For example, a statute that seemed to mean one thing may be construed by a court to mean something different. Although the court will usually say that it is clarifying the statute, it does not always do so. It may create an exception, an exemption, a privilege; it might construe the rule

The present article adds to the literature by formalizing the metaphorical treatments in the first category. Like the articles examining the entropy of legal texts, it relies on formal mathematical and physical definitions, but it extends beyond the mere words of laws to provide more general, quantitative definitions of legal entropy.[7]

FORMALIZING LEGAL ENTROPY

This section relies on the work of Shannon [14], Hohfeld [15], and others to introduce a basic mathematical formalization of the entropy involved in the formulation, interpretation, and application of a law to a given set of facts involving a single legal actor, as well as system-wide entropy across multiple laws and facts concerning many legal actors. In so doing, it begins to overcome the theoretical limitations of the prior literature described earlier.

The Entropy of Legal Systems

As noted, although formal measures of the ambiguity of words in legal documents through measures of word entropy is useful to analyze and parse legal texts, it does not measure the entire extent of interpretive entropy, much less the entropy of legal systems more generally. Rather, one would like to quantify the ambiguity across the entire range of the formulation, interpretation, and application of particular laws to particular behavior.

For instance, the tax laws are notorious for being uncertain in delineation, interpretation, and application (Osofsky [37]). Regarding application, just small variations in the underlying facts relating to a particular tax provision can lead to large changes in the likelihood that the applicable legal actor is obligated to pay taxes or not. Similarly, patent infringement disputes are often difficult to predict, and like tax issues, are sensitive to small variations in the underlying law and facts (Sichelman [38]). Moreover, even if one can quantify the "entropy" of a particular application of law to facts involving a single legal actor, is it possible to quantify the entropy of a legal system and its subsystems encompassing many laws and many legal actors?

Thus, it becomes incumbent to conceptualize the different domains of entropy that arise in legal systems. *Delineative entropy* involves the ambiguity and related information costs in formulating the law in the first instance, typically into written symbols, in a constitution, statutes, regulations, judicial decisions, and the like. As noted earlier, *interpretive entropy* concerns the ambiguity in interpreting the written symbols in legal documents, including not only constitutions, statutes, and regulations, but also judicial decisions. Such an endeavor is not merely textual in nature, but will often involve relying upon institutional and social norms, which themselves can be uncertain. Finally, *applicative entropy* is roughly the uncertainty involved in applying an interpreted law to a given set of facts.[8] Each type of legal entropy is considered in turn, along with a proposed formal quantification of each.[9]

Delineative Entropy

Formulating the law involves many different types of transaction costs. For instance, legislators must be paid to meet, investigate, negotiate, deliberate, and so forth. Similar transaction costs are borne by regulatory agencies and judges in formulating the law. Political scientists and economists have regularly modeled the delineation of law, including related transaction costs, in terms of public choice, game theoretic, and related models (e.g., Benson and Engen [42], Crump [43]). Yet, some of these transaction costs involve information costs that potentially relate to reducing the legal entropy involved in formulating the law, and scholars have yet to provide quantitative measures of such.[10]

What information costs reduce entropy? According to Shannon [14], the entropy of a system is precisely the number of informational bits needed to encode the microstates of the system. When the microstates themselves are unknown, such encoding involves an information cost in determining the precise informational bits of the microstates. As explained more fully in Part 5, actors within the legal system—such as lawyers, lawmakers, judges, law enforcement and others—regularly perform work by incurring information costs to encode microstates—typically, in order to reduce legal entropy by selecting one of the microstates, or at least by reducing the uncertainty in which microstate will be selected. In general, all information costs that genuinely generate new "legal" information about the microstates of some legal system or subsystem will reduce that system's or subsystem's legal entropy (see Part 5). Delineating the law involves many activities that generate new legal information and, in turn, reduce legal entropy. Here, I examine just a portion of those activities. A fuller account of *delineative entropy* would systematically review each and every one of them.

To begin with, lawmakers must determine which actions, roughly speaking, of legal actors not subject to law should be subject to law and, at conversely, which actions currently subject

narrowly to avoid constitutional problems, or broadly to give effect to an unnoticed legislative intent buried in the legislative history. The court's decision becomes a part of the meaning of the rule, so that the rule now becomes more complex—it is a statute plus a judicial decision. The more complex rule may invite further adjudication and more inventive subsequent constructions by courts."

[7] Lee et al. [36] propose a statistical mechanics-based model of voting within groups using a maximum entropy model, applying it to the U.S. Supreme Court. This approach is more in the vein of political science than law per se and, as such, is somewhat orthogonal to the discussion here, but it could be useful in quantifying applicative entropy for disputes to be resolved by a group of adjudicators (e.g., on appeal).

[8] In this regard, contracts also may be considered as a form of "private lawmaking" [39], subject to delineative, interpretative, and applicative entropy. In other instances, law may be formulated in unwritten ways, such as through oral tradition or even social symbols (e.g., Weyrach and Bell [40]), again, subject to all forms of legal entropy.

[9] Another type of legal entropy is enforcement entropy, which stems from the uncertainty in the enforcement of a given law. I abstract away from enforcement entropy in this treatment for simplicity, but the same types of approaches discussed herein would apply to enforcement entropy (see generally Lederman and Sichelman [41]).

[10] Dinga et al. [44] provide a conceptual model of "social entropy," briefly addressing legislation, but do not quantify social entropy in any manner. Although there is a substantial literature regarding "entropy economics" and "economic complexity," that literature does not address the legal domain (see, e.g., Golan [45], Gold and Smith [46], Hausmann et al. [47]).

to law should not be subject to law (or should be subject to amended laws). Ex ante, there will be uncertainty in which categories of human behavior should be subject to law, not subject to law, or subject to legal amendment. Ultimately, this boils down to whether a category of human behavior should be subject to a general change in state with respect to the law.

Astute observers, perhaps aided by legal artificial intelligence, could estimate the probability that some given area of human behavior, especially those relating to newly arising technologies, should be subject to a change in legal state. Suppose it is first day of Nakamoto's [48] now-famous article on bitcoin, and just a few observers know about it. Eventually, information spreads—including from information costs incurred generally unrelated to legal entropy—and lawmakers learn about it but want to know more. At this time, whether activities related to cryptocurrency will be regulated is highly uncertain, probably close to 50%. The lawmakers expend information costs, including formal hearings with experts, to learn more about the economic and social ramifications of bitcoin and whether it should be subject to law.

After much expenditure, legal observers estimate that the probability is 80% that it will be subject to regulation in the near-term within some legal jurisdiction. At each stage during the investigative and deliberative process, one can in theory construct a "near-term" Shannon entropy related to whether some category of human behavior will be subject to law, call it "deontic" entropy,[11] which will be one form of overall delineative entropy.[12]

Specifically, for a "single" human behavior (e.g., whether cryptocurrency usage will be regulated), one can calculate the near-term Shannon deontic entropy by summing across the probability that the category will be subject to law. So, assuming 80%, we arrive at

$$S = -\sum p_i \log_2 p_i = -[(0.8)\, log_2\,(0.8) + (0.2)\, log_2\,(0.2)]$$
$$= 0.72\ bits$$

Such entropy may appear at first blush low, but one must sum across all categories of human behavior to determine the total deontic entropy. Indeed, the level of abstraction at which we categorize human behavior will have a substantial effect on the total deontic entropy. For instance, within the category of cryptocurrency usage, there may be hundreds of distinct behaviors, of which may be subject to law. Thus, conditional on the category being subject to law, one may want to know the probabilities and associated entropy with specific behaviors within the category being subject to law, and for some areas of human behavior, even more fine-grained analyses. Here, in theory, one can use the formal notion of conditional entropy—described further below—to quantify the amount of entropy at every level of abstraction.

Of course, in practice, perfectly quantifying the entropy of hundreds of distinct behaviors at multiple levels of abstraction will be impossible. But contrary to D'Amato's [7] pessimistic suggestion that we give up on the endeavor entirely, as noted earlier, conceptualizing how we should formulate various types of legal entropy, even if practically difficult, serves important purposes. First, it provides an ideal model and, thus, a proper roadmap as to how one should build out a practical approach to modeling entropy. For instance, with respect to cryptocurrency, one may build out a rough model of various behaviors and categories potentially subject to law to generate a rough quantification of deontic entropy. Increasingly, these processes may be aided by legal artificial intelligence and related mechanisms. Second, building out these models may help us to understand law, from a theoretical perspective, in a more coherent and precise fashion. Indeed, the history of science is rife with models that are initially rough and thus difficult to solve, model, or explain, but later become subject to rigorous modeling and application. Notable examples are the development of the atomic theory of matter [50], the theory of evolution [51], and the theory of gravitation [52].

Once lawmakers determine that a given category of human behavior should be subject to law, the question becomes how to specifically draft statutes, regulations, or judicial decisions, particularly at a conceptual level, but even at a textual level, that instantiate the aims of the lawmakers. Specifically, if one begins with a concept (or set of concepts) that lawmakers or judges seek to instantiate into written law (or some other more concrete expression), there will exist multiple ways of instantiating the concept into written law.

Often, this involves the particularization of a more general concept (e.g., "Though shall not kill") into a main rule that specifies the conditions under which the law is violated (e.g., "The killing of another person with malice aforethought," etc., etc.) and sets of exceptions (e.g., except in self-defense, legitimately as a soldier in war, etc.). In other situations, it involves aggregating multiple concepts (e.g., the types of prohibited interferences of third parties against a landowner's permitted uses) into a single rule (e.g., the law against trespass). Quantifying the amount of uncertainty in how to specify the law, generally in text, is useful to understand how delineative entropy and the associated information costs in reducing that entropy play a role in formulating the law—particularly to illustrate how that process differs across different legal domains (e.g., torts vs. real property). In general, one can term this form of delineative entropy as "specificative" entropy.

Suppose there is a single, general concept of interest (e.g., "Thou shall not kill") and astute observers have determined that there are roughly 50 different ways in which the lawmakers could conceptually instantiate the law. These 50 different ways may represent the various degrees of crimes (e.g., 1st vs. 2nd degree murder, voluntary and involuntary homicide), mens rea

[11] Deontic logic concerns the logic of normative concepts, including obligations (see generally Hilpinen [49]).

[12] Alternatively, one might imagine a hypothetical in which lawmakers investigate whether some behavior previously subject to legal restrictions should no longer be subject to them—for instance, the use of a previously illegal drug—resulting in similar deontic probabilities. More generally, deontic entropy concerns the uncertainty regarding whether some set of human behavior should be subject to a change in the law, from no regulation to regulation, from regulation to no regulation, or some intermediate set of changes.

requirements (e.g., malice aforethought, intent, recklessness, and gross negligence), available defenses (e.g., self-defense and duress), and other potentially relevant aspects of the crime (e.g., transferred intent). In general, there may be a large number of permutations of how to instantiate a single concept, either as a single criminal rule, or a large set of related criminal rules.

The line between deontic entropy and specificative entropy is not bright. Although arriving at the applicable concept (e.g., "Thou shall not kill") is squarely in the deontic entropy box, determining the permutations of instantiations under consideration (e.g., 50 forms) and the viability of each of those instantiations can be related both to deontic and specificative entropy. Once the likelihood of each instantiation is estimated, we can calculate the residual specificative entropy, again using Shannon entropy:

$$S = -\sum p_i \log_2 p_i \tag{3}$$

For instance, returning to the bitcoin example, suppose that lawmakers are debating three different bills to regulate bitcoin and other cryptocurrencies. Experts estimate that the first bill has a 50% chance of passage, the second bill a 30% chance, and the third bill, a 20% chance, all subject to the earlier 20% chance that no bill passes.

In this case, we can first calculate the specificative entropy as:

$$S = -\sum p_i \, log_2 \, p_i = -[0.5 \, log_2 \, (0.5) + (0.3) \, log_2 \, (0.3) + 0.2 \, log_2(0.2)] = 2.006 \, bits$$

Note that the total delineative entropy will be a form of conditional entropy, on which the specificative entropy is conditional upon the deontic entropy, for if a given category of human behavior is not going to be subject to regulation, then lawmakers need not expend any effort to regulate it. In this bitcoin example, there is only an 80% chance that any bill will pass, which resulted in a deontic entropy of 0.72 bits. How should one combine the deontic and specificative entropy into an overall delineative entropy value?

Because the specificative entropy is conditional on the deontic entropy, one cannot simply add them together. Rather, the chain rule for conditional information entropy [53] applies:

$$H(X, Y) = H(X) + H(Y|X) \tag{4}$$

where H(X,Y) is the joint (or combined) entropy of two random variables conditional upon one another, H(X) is the entropy solely due to random variable X and $H(Y|X)$ is entropy of Y, conditional upon some specific X (i.e., X = x) occurring.[13]

In other words, the joint (or combined) entropy of a second random variable (Y) that is conditional on a first random variable (X) is the entropy of X plus the entropy of Y conditional on X.[14] In our bitcoin example this results in the following total delineative entropy:

$$H(X, Y) = H(X) + H(Y|X) = 0.72 \, bits + (0.8) * (2.006 \, bits) = 2.33 \, bits$$

Note that this is less than simply adding the deontic and specificative entropy together. This is because the specificative entropy only plays a role conditional upon the deontic entropy resulting in the passage of a bill, which happens 80% of the time. Thus, in this simple example, one effectively adds the deontic entropy to chance that the specificative entropy will be meaningful.

Interpretive Entropy

As noted earlier, once a law is formulated—in a constitution, statute, regulation, judicial decision, or some other legal text—it must typically be interpreted to understand its scope and applicability. In this regard, other types of intermediate legal documents, such as contracts and patents, must be interpreted to determine their legal effect. Legal interpretation is fraught with ambiguity, which can be conceptualized in the framework of information entropy as *interpretive entropy*. Specifically, if one considers the legal rule under consideration the legal "macrostate," then the "microstates" are all of the possible interpretations of the legal rule.[15] Thus, the interpretive entropy is again expressed by the Shannon information entropy, where each state i is a potential interpretation, and the probability of i being the interpretation adopted by the legal institution of interest (e.g., a court or regulatory agency), p_i:

$$S = -\sum p_i \, log_2 \, p_i$$

If the only step involved in this process were to interpret the express text of a legal rule with a standard dictionary, then techniques using word and related forms of linguistic entropy would provide a fairly accurate value of the interpretive entropy. For instance, one could measure the word entropy of legal texts that measures the ambiguity inherent in each word using a standard corpus (cf. Piantadosi et al. [54]).

Yet, legal interpretation extends well-beyond textual interpretation with a standard dictionary. As an initial matter, many legal terms are "terms of art," requiring interpretation by specialized, legal dictionaries. Quantifying interpretive entropy using text-based measures of entropy must rely therefore not on a standard corpus, but a specialized one. More problematic, legal interpretation typically draws upon the

[13] In this regard, the term $p(y|x)$ means the probability of event y given that event × occurs [12].

[14] Formally, the joint entropy H(x, y) of a pair of discrete random variables (X, Y) with a joint distribution p(x, y) is defined [12] as $H(X, Y) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \, log \, p(x, y)$. The conditional entropy is formally defined [12] as $H(Y|X) = -\sum_{x \in X} \sum_{y \in Y} p(x, y) \, log \, p(y|x)$.

[15] Of course, one must also determine where a legal rule begins and ends to interpret the rule. Sometimes, this process is fraught with difficulty, which itself may introduce a form of interpretive entropy. Here, I abstract away from this potentially additional layer of entropy.

language of other legal rules, which necessitates determining how those other rules' affect the entropy of the rule-at-issue. And even more problematic, interpreting rules draw on yet more disconnected sources, such as legislative history, judicial decisions, and even general policies and social norms. Not to mention that interpretive entropy applies not only to the interpretation of legal rules, but also legal documents more generally, such as patents and contracts, which introduce further interpretive issues.

In the face of such complexity, following D'Amato [7], one might throw up one's hands and abandon the quantitative endeavor entirely. However, while complex, legal rules ultimately are interpreted, and experienced attorneys regularly estimate the likelihood of a court interpreting a rule in one fashion or another. Indeed, so much has been recognized since at least Holmes [55]. Moreover, new approaches in legal analytics—like those in sports analytics—are likely to be paradigm-shifting in the ability to predict the outcomes of disputes and related aspects of the law more generally (cf. Katz et al. [32]; Branting et al. [33]).

Finally, the notions of joint and conditional entropy described in the context of delineative entropy similarly allow for interpretive entropy in principle to be broken into discrete parts and recomposed. For instance, courts and others interpreting legal documents usually only turn to sources other than the words when there is some ambiguity in specific words or phrases. If a word or phrase is entirely clear on its face, then typically legal interpreters will adopt a textual interpretation. Thus, the text-based techniques described earlier (e.g., Friedrich et al. [2]) and others can be used as initial cut to determine those words and phrases subject to some latent ambiguity. Words without such ambiguity can either be assumed to have zero entropy or simply the entropy calculated on the basis of the text-based methods. For those words or phrases with latent ambiguity, the text-based score may roughly be considered a primary variable upon which other sources for interpretation (e.g., other statutes or regulations, case decisions, social norms, and the like) can be considered secondary variables conditioned on the primary variables, allowing for the use of joint and conditional entropy as explained earlier. Although the precise nature of this staged approach is beyond the scope of this paper, the general contours sketched here should provide the beginnings of a more comprehensive and realizable method to quantifying interpretive entropy.

Applicative Entropy

Once a legal rule has been interpreted, in order to understand how it specifically regulates human behavior, it must be applied to a specific situation, or set of facts.[16] Even though the applicable legal rule has been fully interpreted, residual indeterminacy in the application of the law may remain and can be quantified by *applicative entropy*. This indeterminacy can arise from "uncertainty as to the impact evidence will have on the decisionmaker," idiosyncratic behavior in adjudication by a decisionmaker such as a judge or jury, and the influence of extra-legal factors on the regulatory and judicial process [56].

Applying the law typically results in the imposition of liability (or no liability), plus some form of remedy in the event liability is imposed. Again, Shannon entropy can be used to measure the entropy of liability and the conditional entropy to measure the entropy present in the range of remedies in the event liability is imposed. Before turning to these specifics, it is instructive to examine the typology of Hohfeld [15], as it offers a sound, quantitative conceptual basis to describe the entropy of composite legal systems and subsystems, which is illustrated well by applicative entropy.

Hohfeld's (Probabilistic) Typology

Rather than try to describe the entire formalism of Hohfeld [15], for purposes of this article, is straightforward enough to explain two Hohfeldian relations: a Hohfeldian right (that is, a right in the strict sense, hereinafter "strict-right") and a Hohfeldian power (see generally Sichelman [57] for a detailed exposition).

Specifically, a legal actor, X, who holds a positive strict-right vis-à-vis Y, with respect to some action A, implies that Y is legally obligated vis-à-vis X to perform that action. For instance, X may hold a contractual strict-right that Y deliver to X's warehouse 100 widgets by the following Wednesday. If X holds a negative strict-right vis-à-vis X with respect to some action A, then Y is obligated to refrain from performing that action (in other words, Y is prohibited from performing the action). For instance, X may hold strict-right in tort that Y not punch X on the nose without justification (e.g., in self-defense).

In Hohfeld [15], whether a first actor X holds a strict-right vis-a-vis a second actor Y can be answered only in a binary fashion such as by a classical bit of information.[17] In other words, if the strict-right is positive in nature, Y either has an obligation to perform some action A or not. Adjudication in the sense of Hohfeld [15] thus involves a determination by the court (or other adjudicatory body) if the application of a law to a given set of facts results in a strict-right/obligation for X and Y or a no-right/no-obligation for X and Y.

For convenience, we will label a strict-right as r_1 and the absence of a strict-right as $\sim r_{1.}$ In this fashion, one can represent a strict-right as an "on-bit" (in binary notation, the number "1") and a no-right as an "off-bit" (in binary notation, the number "0"). In order to more easily manipulate these bits mathematically, it is useful to adopt an equivalent vector formalism, wherein:

$$r_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{and} \quad \sim r_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{(matrix notation)}$$

Hohfeldian powers alter, terminate, or create other legal relations. For instance, by changing the applicable law, the legislature in effect may change X's right vis-à-vis Y that Y perform some action A to no obligation for Y perform the action. In mathematical

[16] Of course, some laws apply to the process of lawmaking, enforcement, and adjudication itself. The discussion of applicative entropy here extends to these legal rules as well. Relatedly, it also applies to the application of an interpreted contract, patent, or similar legal document to a set of facts.

[17] See Marinescu and Marinescu [58] for a detailed discussion of classical information theory.

terms, the legislature's power would be akin to second-rank permutation tensor (here, a 2 × 2 permutation matrix) that changes X's right vector into its negation (a Hohfeldian "no-right") corresponding to the lack of any obligation (a Hohfeldian "privilege") on the part of Y (Sichelman [59]).

Similarly, higher-order powers may change lower-order powers. For instance, a new constitutional amendment may eliminate a previously held second-order power of the legislature. One such example is the passage of the controversial Proposition 209, which amended the California Constitution to prohibit the government, including the legislature, from "discriminat[ing] against, or grant[ing] preferential treatment to, any individual or group on the basis of race, sex, color, ethnicity, or national origin in the operation of public employment, public education, or public contracting" (Cal. Const. art. I, § 31).

In this instance, the amendment would act as a higher-rank tensor that flips the legislature's second-order permutation tensor to a second-rank identity tensor. This is because an identity tensor that operates on first-order right/obligation relations would have no effect on these relations. In other words, a second-order identity tensor is emblematic of the lack of (a second-order) power (a Hohfeldian "disability"). More generally, higher-order powers can be described mathematically as higher-rank tensors that operate on lower-rank tensors (lower-order legal relations) (Sichelman [59]). In this regard, it is important to recognize that vectors are simply lower-rank tensors—thus, a first-order Hohfeldian strict-right is simply a lower-order Hohfeldian power. Thus, all Hohfeldian relations can simply be expressed in terms of legal powers and well-known mathematical relations.

A probabilistic version of Hohfeld's [15] schema may also be developed (Sichelman [59]). Here, instead of legal relations being described by a classical binary bit, the states can exist in probabilistic superpositions, better described by a quantum bit (i.e., qubit). The probabilistic nature of the legal relation may be a result of lack of knowledge of the underlying system or due to inherent indeterminacy in the system itself prior to judgment (a form of system measurement), or a combination of both reasons.[18] Using the qubit formalism, one can specify a probabilistic Hohfeldian relation in the following form:

$$|j>_{\mathrm{n}} = a_n|j_{\mathrm{r}}>_{\mathrm{n}} + b_n\,|j_{\sim\mathrm{r}}>_{\mathrm{n}} \tag{5}$$

Where $|j_{\mathrm{r}}>_{\mathrm{n}}$ is a legal power (or right, to first-order), $|j_{\sim\mathrm{r}}>_{\mathrm{n}}$ is the negation of a legal power, $\mathrm{P}(j_{\mathrm{r}})_{\mathrm{n}} = |a_n|^2$ is the probability of a legal power obtaining upon judgment, $\mathrm{P}(j_{\sim\mathrm{r}})_{\mathrm{n}} = |b_n|^2$ is the probability of the negation, $|a_n|^2 + |b_n|^2 = 1$, and n is the order of the legal relation. In other words, probabilistic Hohfeldian relations, be they first-order strict-rights and obligations or higher-order powers and liabilities, can be characterized by the qubit formalism,[19] where there is a probability $|a_n|^2$ that the legal relation will be measured (i.e., adjudicated) in the "power" state and $1-|a_n|^2 = |b_n|^2$ in the negation (or lack) of the "power" state. The indeterminacy regarding the state of the system prior to judgment can be quantified as a form of applicative legal entropy.[20]

Quantifying Applicative Entropy

As noted earlier, a legal judgment with respect to a first-order relation (strict-right/obligation) will either result in a finding that the defendant had an obligation (is liable) or not. Less frequently, a judgment may concern whether a legal actor holds a higher-order relation (power) or not. Prior to this judgment, the indeterminacy in the judgment again can be quantified using the Shannon entropy:

$$S = -\sum p_i \, log_2 \, p_i$$

As there are only two potential outcomes in judgment, this reduces to the binary Shannon entropy:

$$S = -plog_2 p - (1-p) log_2 (1-p) \tag{6}$$

where p is the probability that a court finds a power (right to first-order) and corresponding liability (obligation to first-order) on the part of the defendant. In general, as discussed earlier, this entropy is maximum when there is a 50% chance of liability and a 50% of chance of no liability (see **Figure 1**).

Conditional upon a finding of liability, Shannon entropy can be applied to the range of potential remedies. In the event the remedies are discrete in nature, the ordinary Shannon entropy formula may be used. However, because remedies are usually continuous,[21] it becomes necessary to use the differential or continuous entropy [53]:

$$h(Y) = -\int_Y f(y) \, lnf(y) \, dy \tag{7}$$

where f(y) is a probability density function of the potential remedies. For instance, f(y) may represent the likelihoods

[18]In the event the indeterminacy results from mere lack of knowledge, the state of the system can be described wholly by classical probability theory. However, the qubit formalism can easily model both the lack of knowledge and inherent indeterminacy in a given system prior to measurement. Like the difference between classical and quantum mechanics (i.e., in physics), the conceptual difference between mere lack of knowledge and inherent indeterminacy for legal relation states may—in certain conceptions—lead to quantum-like legal effects, such as legal "entanglement" of states, that can distinguish "classical" legal indeterminacy (i.e., lack of knowledge) from "quantum" legal indeterminacy (i.e., inherent indeterminacy). For simplicity, however, I rely upon the classical entropy formulas herein.

[19]Note that the for the higher-order relations, the states $|j_{\mathrm{r}}>_{\mathrm{n}}$ and $|j_{\sim\mathrm{r}}>_{\mathrm{n}}$ are 2nd-rank and higher-rank tensors, not simply vectors, as in the standard quantum formalism.

[20]In this regard, note that while the Von Neumann entropy—which in effect measures the indeterminacy of a mixed quantum state with respect to its entangled substates—is zero for a pure quantum state, there is nonetheless Shannon information entropy for a pure state with respect to the indeterminacy of its potential measurement outcomes prior to a measurement [60].

[21]In general, outcomes in legal contexts are discrete and small in number, e.g., one of a small number of potential formulations, interpretations, and applications of the law. Even potential remedies are often discrete, though as noted, in some cases, remedies may form a continuous distribution. Enforcement likelihoods may also form a continuous distribution, but as noted, this paper abstracts away from enforcement for simplicity. Of course, empirical studies of legal systems may construct effectively continuous distributions from large datasets (e.g., of words in statutes), but this paper explores legal entropy from an internal perspective of the legal system itself. Of course, it should be straightforward given the discussion here to apply its concepts to such external empirical studies.

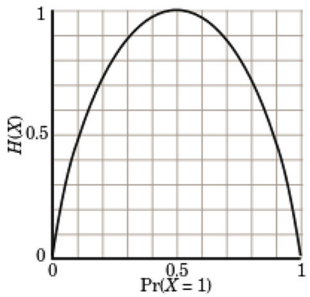


FIGURE 1 | Binary shannon entropy vs. probability of an event occurring.

of various lengths of prison sentences or the amount of damages owed.

For example, suppose that the likelihood of sentences follows a typical bell curve function:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \tag{8}$$

where μ is the mean of the distribution and σ is the standard deviation of the curve. In this instance, the differential entropy is:

$$h(Y) = -\int_Y \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} \ln \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-\mu)^2}{2\sigma^2}} dy \tag{9}$$

With some rearranging, this reduces to the simple expression [61]:

$$h(Y) = 2\ln(2\pi e\sigma^2) \tag{10}$$

The total applicative entropy of a given dispute can then be calculated by the chain rule discussed in Part 4.2:

$$H(X,\ Y) = H(X) + H(Y|X) \tag{11}$$

Here, *H(X)* is applicative entropy related to whether liability will be found or not, and $H(Y|X)$ is the applicative entropy related to the remedy, conditional upon liability being found.

Systemwide Entropy

The previous discussion of legal entropy, be it delineative, interpretive, or applicative entropy, has involved one or a narrow set of rules, interpretations, and applications. An immediate question arises as to how to derive the entropy of a legal system or subsystem that encompasses numerous legal relations (in the Hohfeldian sense). Although a full treatment of systemwide legal entropy is beyond the scope of this article, some preliminary remarks may be made.

First, the Hohfeldian mathematical formalization of relations into vectors and tensors is particularly helpful in conceptualizing the quantitative state and related entropy of a system. Recall that the general probabilistic state of a legal relation can be described as the following:

$$|j>_{\mathrm{n}} = a_n|j_{\mathrm{r}}>_{\mathrm{n}} + b_n|j_{\sim\mathrm{r}}>_{\mathrm{n}} \tag{12}$$

For a first-order relation (a Hohfeldian strict-right), the state of a single legal relation is simply a 2-D vector in a "Hohfeldian" state space. As the vector rotates in the state space, the a_n and b_n—and, hence, the relative probabilities of a right being found—change in time. This rotation may be due to the operation on the vector by a power (i.e., a tensor) or simply by the change in external circumstances (e.g., changing underlying facts). For the higher-order relations (a Hohfeldian power), the state of a single relation is a multi-dimensional tensor in the Hohfeldian state space. Like the vector, these tensors may rotate (in a higher-dimensional sense) in the state space, with corresponding changes in the relative probabilities of liability being found (or not).

In theoretical terms, one can imagine the complete Hohfeldian state space as a collection of a multitude of vectors and tensors corresponding to every possible action and states of the world affected by law (and a complement space of all of those actions and states not so affected). In other words, any action or state of the world with a non-zero probability of being subject to a power (including a first-order, power, i.e., strict-right) has an associated vector (or tensor) in the state space representing that specific probability.[22] The complement state space represents all actions and states of the world with a zero probability of being subject to a power (including a strict-right).

Of course, listing every possible action and state of the world potentially subject to law and determining how those states change over time is essentially an impossible task.[23] However, for discrete subspaces, it is certainly possible to construct such a space, measuring entropy and other useful properties of the system. For instance, a patent typically will contain multiple, separate claims, each of which describes a slightly different instantiation of the invention. In this sense, each claim provides a separate legal right to prevent third parties from making, using, and selling the corresponding claimed invention. One claim in a patent may relate to a product, the sale of which infringes that claim, while another claim in a patent may relate to the performance of a method with the product that infringes the claim. Although the two claims are related, actions that infringe one claim (e.g., the product claim) may not infringe the other claim (e.g., the method claim).

The legal subspace of interest may be all of the claims of the patent as they apply to the activity of a potential infringer. For an accused infringer to be found liable, the claim must both be valid and enforceable (which typically does not depend on the specific activity of the infringer[24]) and infringed (by the particular activity of the infringer). Thus, using Shannon entropy, one can first determine the probability that each claim will be found valid and enforceable, using that to determine the entropy related to validity/enforceability, and conditional upon a positive

[22] Cf. Gold and Smith [45, 46] ("In other words, we pick the costlier-to-provide legal relation where Shannon entropy is higher and the cheaper signal (sometimes doing almost nothing) for where it is lower.").

[23] For a discussion of how one might model the changing information nature of legal rights and related interests over time using an evolutionary approach, see Alston and Mueller [62] and Ruhl [63].

[24] Sometimes the enforceability of a patent claim turns on facts unique to a given dispute, but for simplicity, I assume here that it is a general determination—i.e., that it applies across all disputes.

determination, the probability that the claim is infringed by the accused infringer, which then can be used to calculate the conditional entropy and, ultimately, total entropy of each claim (again, according to the chain rule described earlier).

If the decision for each patent claim is statistically independent, then one can simply add the entropy for each claim together. However, whether a given patent claim is valid, enforceable, and infringed is typically correlated to other patent claims in the same patent. When the judgment of particular claims is correlated, the joint entropy of the claims in combination can be used to determine the total entropy.

Specifically, the joint entropy of a set of multiple, random, discrete variables that are potentially correlated can be written as [53][25]:

$$H\left(X_1,\ \ldots,\ X_n\right) = -\sum_{x_1} \ldots \sum_{x_n} P(x_1,\ \ldots,\ x_n) \log_2 P(x_1,\ \ldots, x_n)$$

Here, P $(x_1, \ldots x_n)$ is the joint probability that each event occurs together. For instance, if there are only two patent claims, X and Y, then the possible outcomes are defendant is liable on X and Y (outcome 1); liable on X but not Y (outcome 2); liable on Y but not X (outcome 3); and liable on neither (outcome 4). If the probability of outcome 1 is 1/4, outcome 2 is 1/3, outcome 3 is 1/6, and outcome 4 is 1/4, then the joint entropy for the two patent claims is:

$$-\ (1/4)\log_2(1/4) + (1/3)\log_2(1/3) + (1/6)\log_2(1/6) \\ +(1/4)\log_2(1/4) = 0.96\text{ bits}$$

Note that the underlying probabilities of liability—and, hence, legal entropy—may change in time due to the exercise of a second-order power by a legislature or some external set of circumstances. For instance, Congress passed the America Invents Act in 2011, which effectively changed the probability that a given patent claim would be found invalid. In this instance, the legislative change would immediately rotate all patent claim vectors in state space, resulting in different probabilities and, in turn, a different subsystem entropy. Similarly, exogenous changes in societal norms, technology, economics, and so forth, may affect the underlying probabilities of legal claims (e.g., Cooter [64]), again, rotating vectors in state space as these changes take effect. For instance, the advent of the Internet arguably changed how judges view software patent claims as a whole, which in turn led to a diminished role for software patents more generally (e.g., Barzel [65]).

These concepts can be extended to the legal system as a whole. First, divide the legal system into independent legal subsystems (To the extent the law is truly a "seamless web," skip this step.) The entropy of any legal subsystem can be constructed in principle by using the joint and conditional entropies of individual states within the subsystem. (Even if the law is a seamless web, at some point, the correlations among states is so low, they can be ignored and the legal system treated as if it is composed of independent subsystems.)

Again, in practice, this will be nearly impossible, but for certain subsystems of interest—e.g., a patent—certainly possible, especially with improvements in AI approaches for modeling the law. And, again, in the very least, it provides a conceptual framework for richer jurisprudential understandings of the law. The next part considers the beginnings of these richer accounts in a few notable areas.

PRACTICAL USES OF LEGAL ENTROPY

Legal Indeterminacy

There is an extensive literature on the notion of legal indeterminacy (see Solum [13] for a discussion). One camp, particularly those in the critical legal studies vein, argue for radical indeterminacy of legal doctrine and judicial decisionmaking (e.g., Kennedy [66], D'Amato [7], Singer [67]). Another argues for minimal determinacy, at least in principle (e.g., Dworkin [3]). And the last camp takes an intermediate position (e.g., Kress [68]). Yet, despite the numerous articles on the topic of legal indeterminacy, only a handful of pieces attempt to quantify it—some by examining the ambiguity of legal language using measures from computational linguistics (e.g., Katz and Bommarito [1]) and others by analyzing reversal rates and dissents as a possible proxy of indeterminacy (e.g., Lefstin [69]). Yet, none attempt a wholesale quantification of the amount of indeterminacy present in legal rules and adjudication. This lacuna is notable, because filling it may help to solve many of the recurring debates and disagreements regarding legal indeterminacy in the literature.

For instance, Kress [68] contends, "The pervasiveness of easy cases undercuts critical scholars' claim of radical indeterminacy." Interpretive and applicative entropy measured across numerous legal rules and related disputes provide a quantitative test of this assertion. Of course, practically quantifying these types of entropy is no simple feat, but a combination of human-coded and automated methods—including those using advances in AI (e.g., Branting et al. [70]; Katz et al. [32]; Branting et al. [33])[26]—could certainly provide a precise quantitative metric for at least a particular field or doctrine in the law. As machine learning and other automated techniques in legal document classification and analysis continue to improve, arguably, the previous impasse among scholars regarding the "indeterminacy" of the law should

[25]For continuous variables, such as remedies, the differential joint entropy may be used [53].

[26]Branting et al. [33] uses "maximum entropy" classification models to predict outcomes of a variety motions in federal district court. Although Branting et al. [33] does not discuss legal entropy as that term is used here, maximum entropy models as applied to legal disputes implicitly concern interpretive and applicative entropy. Specifically, Branting et al. [33] attempt to predict outcomes for three different types of motions using the following features: "the party filing the motion, the judge ruling on the motion, the sub-type of motion, and alphanumeric character sequences having non-alphanumeric characters on both the left and right sequence borders that occur in the text of the motion." As Berger et al. [50] explain, a maximum entropy approach "model[s] all that is known and assume[s] nothing about that which is unknown." Thus, if one has no information about the result of a motion in court, one assumes the "maximum entropy," which would give each side a 50% chance of winning. In modeling legal outcomes with maximum entropy models, one begins with one bit of interpretive and applicative entropy per decision and works to reduce the overall entropy—and, thus, increase predictability—by incorporating more and more (training) information by fitting that information to the known data through a series of logistic regressions (see, e.g., Yu et al. [71]).

yield to at least a modicum of agreement. Regardless, quantitative approaches to legal entropy can provide a deeper understanding of this core issue in the law.

Legal Entropy, Modularity, and Work

Refining Legal Modularity With Legal Temperature and Work

As noted earlier, the original concept of entropy in thermodynamics resulted from investigations regarding an ideal engine and how the transfer of heat, ΔQ, from a heat source at a certain temperature, T, to an idealized engine in a so-called reversible process, increased the entropy of the system by an amount $\Delta Q/T$. Recall that Boltzmann [17] provided a microscopic picture of the macroscopic entropy, whereby heat is essentially a "disordered" collection of microscopic particles, the introduction of which increases the system disorder by a measure of $\Delta Q/T$. According to Boltzmann [17] this ratio can be captured by the total number of microstates of a system corresponding to a given macrostate.

To gain a deeper appreciation of legal entropy, it is useful to construct the notion of legal heat and legal temperature. For simplicity, consider the context of applicative legal entropy with respect to first-order Hohfeldian relations (i.e., strict-right/obligation), where a fully interpreted legal rule is subject to judgment. Recall that each legal relation can be depicted as a vector in state space, such that if the vector does not lie upon one of the axes, there is indeterminacy in the judgment, such that:

$$|j>= a|j_{\mathrm{r}}> +b|j_{\sim\mathrm{r}}> \tag{13}$$

where $|j_r>$ is a legal right, $|j_{\sim r}>$ is the negation of a right, $P(j_r) = |a|^2$ is the probability of a legal right obtaining upon judgment, $P(j_{\sim r}) = |b|^2$ is the probability of the negation obtaining, and $|a|^2 + |b|^2 = 1$. Also recall that the entropy in this situation is:

$$S = -p\, log_2\, p - (1-p)\, log_2\,(1-p) = -|a|^2\, log_2\, |a|^2 \\ -|b|^2\, log_2\, |b|^2$$

As legal "heat" enters a legal system, the entropy of the system increases by increasing the underlying uncertainty in outcome. For instance, in the context of applicative entropy (i.e., judgments), increasing uncertainty in the underlying facts constitutes legal "heat" that shifts the Hohfeldian state vector away from vertical or horizontal and into a diagonal position, maximizing entropy when the state vector is at a perfect diagonal, $|a|^2 = |b|^2$, with the corresponding result that judgment is a coin flip (50/50).

Incoming legal heat will have less effect the higher the legal temperature. For instance, if a legal system is at its maximum applicative entropy (50/50), the introduction of more heat cannot increase the entropy of the system. For instance, suppose that the adjudicator—the judge or jury—has already decided to flip a coin to determine the outcome of a dispute. Thus, increasing uncertainty in the underlying facts will have no effect on the ultimate outcome. Because the *change* in entropy is proportional to the change in heat divided by the background system temperature, at least for applicative entropy, we can see that *background* system temperature is directly proportional to the *background* entropy. In other words, when the background entropy is very low, the background temperature is low, and the introduction of legal heat will be more meaningful.

The notion of legal heat, temperature, and entropy can be useful in explaining important theoretical concepts in the law at a more quantitative and arguably deeper level. Importantly, the insightful work of Smith [8–10] on the role of modularity in legal systems can be refined using these concepts. Specifically, Smith [8–10] posits that information costs play an integral role in the "modularity" of legal systems—namely, the use of "boundaries" in the law (be they spatial or intangible) to "economize on information costs" by "hiding" classes of information "behind" the boundaries.

Smith's concept can also be understood in terms of legal entropy. Specifically, as the entropy of a legal subsystem increases, more information—and, hence, more information costs—are required to encode and resolve disputes concerning the legal subsystem. For instance, consider a piece of real property, and the potential uses of an "owner" and "third parties" with respect to the property. One could list out every potential use of the owner and third parties with respect to the property, determining whether such use "improperly" interferes with the owner's or a third party's "rights," where the rights and interference thereof are defined by some set of background laws and principles. Each step in this use-by-use analysis would be fraught with substantial indeterminacy, generating high entropy and hence large information costs to resolve whether each use is "rightful" (cf. Smith [8–10]). Similarly, a use-by-use approach would involve large costs in delineating and interpreting the law.

As an alternative, the boundary of the property may be used as a proxy to define rightful and wrongful uses to substantially reduce systemwide entropy and, hence, information costs in delineating, interpreting, and applying the law. In other words, in the terms of Smith [8–10], the boundary effectively hides the owner's (unspecified) interests in using the land from legal consideration in the investigation of actions by a third party. To determine if a third party unreasonably interfered with the owner's interests, instead of examining whether a particular action on the part of the third party interfered with particular uses of the owner, the law generally assumes that when a third party unjustifiably crosses the boundary, an interference occurs. This assumption economizes on information costs by using the boundary as a reliable proxy for actual interference with the owner's and third parties' specific interests.

Of course, erecting boundaries as proxies can introduce error costs in allocating rights and duties, so it is important to place some constraints on the modularization of law. The notions of legal temperature and entropy can also perform important work in imposing such constraints. Namely, it is only when the legal system inside the boundary has relatively low temperature and entropy, especially when compared to the temperature and entropy near or outside the boundary, that modularity will serve its role to reduce information costs without imposing significant error costs. If, on the other hand, entropy and temperature were to rise inside the boundary—for instance, as the result of substantial, ever-changing and indeterminate State regulation as to the uses that the owner could undertake—then the modularity of an "exclusionary" approach to property becomes less attractive, instead yielding to a more particularized "governance" approach (Smith [8–10]). Using the information-theoretic concepts of legal entropy and temperature not only helps to more fully explain modularity, but also provides a means to quantify how modularity functions, and when exclusionary regimes should yield to governance regimes [see generally (Sichelman and Smith, unpublished)[27]].

Reducing Legal Entropy and the "Work" of the Legal System

As the discussion of modularity shows, the legal system can in effect reduce legal entropy by reducing the uncertainty and related information costs in delineating, interpreting, and applying the law. More generally, lawyers and the legal system expend "work" by drafting and interpreting constitutions, laws, regulations, contracts, patents, and other legal documents to reduce the amount of uncertainty in whether particular actions that legal actors may perform are permitted, forbidden, or obligated (in deontic terms) and in whether particular laws are valid or not (i.e., can effectuate a power in Hohfeldian terms).

Legal "work" expends transaction costs in the time and effort required to draft, interpret, and apply the law, which often encompasses time and effort in negotiation, the collection of facts, the investigation of background law, and so forth.[28] A portion of these transaction costs are "information costs" in the sense of Smith [8–10]. The efficient level of information costs can be specifically quantified as the amount of entropy reduction in a legal system or subsystem performed by legal work.[29] Specifically:

$$\Delta W_I = -\Delta S_L \tag{14}$$

In other words, the amount of legal work efficiently expended in information costs directly reduces the legal entropy by the same amount.[30] The legal entropy is the sum of the delineative, interpretive, and applicative entropy defined in earlier, as well as other types of legal entropy not discussed here (e.g., enforcement entropy).

The direct relationship between legal work, information costs, and entropy reduction provides a direct linkage between legal entropy and the economic theory of law, importantly including Coase's Theorem. Specifically, because information costs are a class of transaction costs, one can postulate a Coasean world in which the only transaction costs are information costs.

In a world solely of information costs, the efficient allocation of legal rights will depend on the amount of legal work the legal system must expend on reducing the entropy of the system from one in which both actors hold Hohfeldian privileges (i.e., the absence of a duty) to one in which the actors are subject to one or more duties. As Parisi [4] notes, it is only in the hypothetical absence of transaction costs that delineation is costless and entropy may effectively be disregarded.

Importantly, Coase [12] abstracted away from the fact that the initial delineation of legal entitlements between legal actors itself expends transaction (including information) costs (Lee and Smith [76]). Instead, Coase posited an artificial world of zero transaction costs *only after* the initial assignment of "property rights," assuming such assignment is costless. Indeed, as Lee and Smith [76] properly recognize, Coase's [12] notion of "property rights" is more akin to "thin," costly-to-delineate contractual rights than the usually "lumpy," less-costly property rights. In this sense, Coase [12] obscured an important aspect of the relationship between transaction costs and the assignment of legal rights. Namely, because "pre-Coasean" transaction costs must be expended in the delineation of the law itself, it will only be efficient to randomly delineate and assign a right to one legal actor or another when these pre-Coasean transaction (including information) costs in so doing are symmetric and, thus, equal.

However, in many, if not most, cases, the transaction costs of delineation are not symmetric in the assignment of entitlements, which implies that an efficient allocation of entitlements will *not* occur even if transaction costs following that allocation are zero.[31] This is particularly so for *well-defined* legal entitlements, a key assumption of the Coasean, post-assignment, transaction cost-less world.[32]

[27] Sichelman T, Smith HE. *Measuring Legal Modularity*. On file with author (2021).

[28] Another set of information costs (and associated entropy) arises from enforcement of the law. As noted earlier, this article abstracts away from such concerns for simplicity, but they are certainly important, and susceptible to the approaches described herein.

[29] Here, the "efficient level of information costs" assumes that it is efficient to increase legal certainty; in some situations, legal uncertainty may be economically efficient. Cf. Kaplow [72] (describing potential benefits of uncertainty in legal rules).

[30] Parisi [5] is apparently the earliest work to associate legal entropy with positive, asymmetric transaction and strategic costs. The model offered in the present article provides a more precise relationship between legal entropy and transaction costs; namely, it posits that transaction costs *arise from activities that reduce legal entropy*. In other words, information (a form of transaction) costs are expended to make a legal system or set of entitlements more predictable. Cf. Yang [73] (noting that costs to acquire information in the context of an economic coordination game reduce informational entropy). However, in economic parlance, transaction costs may be viewed as a form of economic "friction," burning up surplus in a metaphorical manner that reduces the amount of available energy (i.e., analogous to surplus) with a concomitant increase in systemwide entropy. More precisely, the burning of surplus in the form of transaction costs will lead to economic uncertainty in the sense of Shannon entropy if the number of microstates corresponding to suboptimal, high-transaction cost welfare regimes is higher than the number of microstates corresponding to relatively optimal, low-transaction cost welfare regimes. (For general reflections on the notion of entropy in economics, see Rosser [74].) Moreover, economic transactions are achieved by physical activities that generate transaction costs, including information costs, and may result in real-world, physical entropy (e.g., the use of a computer) (see, e.g., Georgescu-Roegen [75]). Nonetheless, economic and physical frictions are not the type of "frictions" of concern to *legal* entropy in the sense used in the present article. In other words, each form of entropy (e.g., legal, economic, physical) is an independent instantiation of the general notion of entropy embodied in the Shannon entropy formula (cf. Jaynes [20]).

[31] Parisi [77] offers a revision to the Coase Theorem based on a conceptual approach to entropy that focuses on how asymmetric transaction costs in the transfer of property rights affect remedies, rather than the asymmetric costs in the delineation of legal entitlements in the first instance (see also Luppi and Parisi [78]).

[32] Barzel [65] notes that "[i]t is evident, however, that costless transacting results in the perfect delineation of rights and that it is redundant to also require that rights are well defined," but does not extend this observation to its effects on the Coase Theorem in the first instance (see also Barzel [79]). Cheung [80] makes a similar observation, stating "private property rights cannot coexist with zero transaction costs." Cheung [80] criticizes the Coase Theorem on these and other grounds, but does not discuss how these costs might be incorporated into an extended notion of the Coase Theorem.

In sum, even before the Coasean world comes into being, transaction costs typically play a fundamental role. Thus, postulating a world with no transaction costs *only subsequent* to the allocation of entitlements does not necessarily imply that the initial assignment is always efficient. An understanding of the delineation entropy involved in the assignment of rights prior to the Coasean world of zero transaction costs is central to a deeper understanding of how transaction costs bear on the efficiency of the legal system.

CONCLUSION

Numerous legal scholars have discussed the notion of legal entropy, but few have attempted to quantify it. Those attempts to quantify the notion have been limited to analyzing the ambiguity of legal texts by measuring the entropy of words. Although certainly useful, these approaches fail to capture the multifaceted nature of legal entropy. In this article, relying upon the work of Shannon [14, 31] and Hohfeld [15], I have proposed the beginnings of a mathematical framework to quantify legal entropy more broadly. The model proposed offers several useful benefits. First, it offers a potential template for how legal AI systems can measure and store information about the uncertainty of legal systems. Second, the model helps to explain more fully the nature and function of so-called legal indeterminacy as well as the "modularization" of the law and Coasean notions of how transaction costs affect the allocation of legal entitlements. To be certain, the model fails to address important practical details concerning how to assess the underlying probabilities necessary to calculate legal entropy, but hopefully increasing advances in legal AI will lead to the wide-scale realization of such a model in the near future.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of this work and has approved it for publication.

ACKNOWLEDGMENTS

I am grateful to my research assistant, Millicent Whitemore, and my research associate, Yuksel Calli, for assembling sources and their detailed review of an earlier version of the manuscript. I also thank Michael Bommarito, Philip Glass, Daniel Martin Katz, Francesco Parisi, Anita Parkash, Henry Smith, and Marika Taylor for helpful comments and questions. I thank Daniel Martin Katz, J. B. Ruhl, and Pierpaolo Vivo for hosting and inviting me to present this article at *The Physics of the Law: Legal Systems Through the Prism of Complexity Science*.

REFERENCES

1. Katz DM, Bommarito MJ. Measuring the complexity of the law: the United States Code. *Artif Intell Law*. (2014) 22:337–74. doi: 10.1007/s10506-014-9160-8
2. Friedrich R, Luzzatto M, Ash E. Entropy in legal language. In: *CEUR Workshop Proceedings*. (2020). p. 25–30.
3. Dworkin R. *Law's Empire*. Belknap Press (1986).
4. Parisi F. Entropy in property. *Am J Comparat Law*. (2002) 50:595–632. doi: 10.2307/841063
5. Parisi F. Freedom of contract and the laws of entropy. *Supreme Court Econ Rev*. (2003) 10:65–90. doi: 10.1086/scer.10.1147138
6. Ruhl JB, Ruhl HJ. The arrow of the law in modern administrative states: using complexity theory to reveal the diminishing returns and increasing risks the burgeoning of law poses to society. *UC Davis Law Rev*. (1997) 30:407–82.
7. D'Amato A. Legal uncertainty. *Cal Law Rev*. (1983) 71:1–55. doi: 10.2307/3480139
8. Smith HE. Modularity in contracts: boilerplate and information flow. *Mich Law Rev*. (2006) 104:1175–222.
9. Smith HE. Intellectual property as property: delineating entitlements in information. *Yale Law J*. (2007) 116:1742–822. doi: 10.2307/20455776
10. Smith HE. Property as the law of things. *Harv Law Rev*. (2012) 125:1691–726.
11. Newman CM. A license is not a 'contract not to sue': disentangling property and contract in the law of copyright licenses". *Iowa Law Rev*. (2013) 98:1101–62. doi: 10.2139/ssrn.2010853
12. Coase RH. The problem of social cost. *J Law Econ*. (1960) 3:1–44. doi: 10.1086/466560
13. Solum L. Indeterminacy. In: Patterson D, editor. *A Companion to Philosophy of Law and Legal Theory*. Oxford: Blackwell Publishing, Ltd (1999). p. 488–502. doi: 10.1111/b.9780631213291.1999.00034.x
14. Shannon CE. A mathematical theory of communication. *Bell Syst Techn J*. (1948) 27:379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x
15. Hohfeld W. Some fundamental legal conceptions as applied in judicial reasoning. *Yale Law J*. (1913) 23:16–59. doi: 10.2307/785533
16. Clausius R. Über eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmertheorie. *Poggendorffs Annalen Phys Chem*. (1854) 93:481–506. doi: 10.1002/andp.18541691202
17. Boltzmann L. Weitere Studien uber das Warmegleichgewicht unter Gasmolekulen. *Sitzungsb Akad Wissen Wien Ber*. (1872) 66:275–370.
18. Boltzmann L. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wien Ber*. (1877) 76:373–435.
19. Gibbs JW. *On The Equilibrium of Heterogeneous Substances*. Connecticut Academy of Arts and Sciences (1874). doi: 10.5479/sil.421748.39088007099781
20. Jaynes ET. Information theory and statistical mechanics. *Phys Rev*. (1957) 106:620–30. doi: 10.1103/PhysRev.106.620
21. Stone J. *Information Theory: A Tutorial Introduction*. Tutorial Introductions (2015).
22. Lewis PB. Bankruptcy and thermodynamics. *Florida Law Rev*. (1998) 50:329–84.
23. Ferrara M, Gagliotti RA. Legal values and legal entropy: a suggested mathematical model. *Int J Math Models Methods Appl Sci*. (2012) 3:490–8.
24. Berg L. System of legal impact from social entropy to legal development. *Rev Quid*. (2018) 2 (Special Issue):133–7.
25. Edgar C. Struggling against entropy: monetary patent infringement damages after eBay. *Drake Law Rev*. (2018) 66:45–67.
26. Fromer J. An information theory of copyright law. *Emory Law J*. (2014) 64:71–128.
27. King JH Jr. Deus ex machina and the unfulfilled promise of New York Times v. Sullivan: applying the times for all seasons. *Ky Law J*. (2007) 95: 650-714.
28. Moran J. Postmodernism's misguided place in legal scholarship: chaos theory, deconstruction, and some insights from Thomas Pynchon's fiction. *South Calif Interdiscip Law J*. (1997):155–99.

29. Stephan PB. One voice in foreign relations and federal common law. *Va J Int Law*. (2019) 60:1–79.
30. Loevinger L. Politics and the regulatory agencies. *Columb Law Rev*. (1968) 68:371–83. doi: 10.2307/1120891
31. Shannon CE, Weaver W. *The Mathematical Theory Of Communication*. Urbana, IL: The University of Illinois Press (1949).
32. Katz B, Bommarito MJ, Blackman M. A general approach for predicting the behavior of the supreme court of the United States. *PLoS ONE*. (2017) 12:e0174698. doi: 10.1371/journal.pone.0174698
33. Branting LK, Yeh A, Weiss B, Merkhofer EM, Brown B. Inducing predictive models for decision support in administrative adjudication. In: *AICOL International Workshops. AI Approaches to the Complexity of Legal Systems, Vol. 10791*. Springer (2017). p. 465–77. doi: 10.1007/978-3-030-00178-0_32
34. Shannon CE. Prediction and entropy of printed english. *Bell Syst Techn J*. (1951) 30:50–64. doi: 10.1002/j.1538-7305.1951.tb01366.x
35. Zhang Y, Qian Y, Huang Y, Guo Y, Zhang G, Lu J. An entropy-based indicator system for measuring the potential of patents in technological innovation: rejecting moderation. *Scientometrics*. (2017) 111:1925–46. doi: 10.1007/s11192-017-2337-7
36. Lee ED, Broedersz CP, Bialek W. Statistical mechanics of the US supreme court. *J Stat Phys*. (2015) 160:275–301. doi: 10.1007/s10955-015-1253-6
37. Osofsky L. The case against strategic tax law uncertainty. *Law Rev*. (2011) 64:489–538.
38. Sichelman T. Myths of (un)certainty at the federal circuit. *Loyola Los Angeles Law Rev*. (2010) 43:1161–94.
39. Slawson WD. Standard form contracts and democratic control of lawmaking power. *Harv Law Rev*. (1971) 84:529–66. doi: 10.2307/1339552
40. Weyrach WO, Bell MA. Autonomous lawmaking: the case of the “Gypsies”. *Yale Law J*. (1993) 103:323–99. doi: 10.2307/797098
41. Lederman L, Sichelman T. Enforcement as substance in tax compliance. *Wash Lee Law Rev*. (2013) 70:1679–749.
42. Benson BL, Engen EM. The market for laws: an economic analysis of legislation. *South Econ J*. (1988) 54:732–45. doi: 10.2307/1059016
43. Crump D. Game theory, legislation, and the multiple meanings of equality. *Harv J Legis*. (2001) 38:331.
44. Dinga E, Tanaescu CR, Ionescu GM. Social entropy and normative network. *Entropy*. (2020) 22:1051. doi: 10.3390/e22091051
45. Golan A. Information and entropy econometrics — a review and synthesis. *Foundat Trends Econometr*. (2008) 2:1–145. doi: 10.1561/08000 00004
46. Gold A, Smith HE. Scaling up legal relations. In: Balganesh S, Sichelman T, Smith H, editors. *The Legacy of Wesley Hohfeld: Edited Major Works, Select Personal Papers, and Original Commentaries*. Cambridge, MA: Cambridge University Press (2021).
47. Hausmann R, Hidalgo C, Bustos S, Coscia M, Simoes A. *The Atlas of Economic ComplexityL Mapping Paths to Prosperity*. Cambridge, MA: MIT Press (2014). doi: 10.7551/mitpress/9647.001.0001
48. Nakamoto S. *Bitcoin: A Peer-to-Peer Electronic Cash System*. (2008). Available online at: http://www.bitcoin.org/bitcoin.pdf
49. Hilpinen R (ed). Deontic logic: introductory systematic readings. In: *Synthese Library, Vol. 33*. Dordrecht: Springer (1971). doi: 10.1007/978-94-010-3146-2
50. Berger AL, Della Pietra SA, Della Pietra VJ. A maximum entropy approach to natural language processing. *Comput Linguist*. (1996) 22:39–71.
51. Smith JM. *The Theory of Evolution*. Cambridge, MA: Cambridge University Press (1993).
52. Sachs M. *Quantum Mechanics and Gravity*. Berlin: Springer (2004) doi: 10.1007/978-3-662-09640-6
53. Cover TM, Thomas JA. *Elements of Information Theory*. Wiley (2012).
54. Piantadosi ST, Tily H, Gibson E. The communicative function of ambiguity in language. *Cognition*. (2012) 122:280–91. doi: 10.1016/j.cognition.2011.10.004
55. Holmes OW. The path of the law. *Harv Law Rev*. (1897) 10:457–78. doi: 10.2307/1322028
56. Yablon CM. The good, the bad, and the frivolous case: an essay on probability and rule 11. *UCLA Law Rev*. (1996) 44:65–107.
57. Sichelman T. Some fundamental conceptions as applied in judicial reasoning. In: Hohfeld WN, editors. *Wesley Hohfeld One Hundred Years Later: Edited Major Work*. Cambridge: Cambridge University Press (2021).
58. Marinescu DC, Marinescu GM. *Classical and Quantum Information*. Burlington, MA: Academic Press (2011) doi: 10.1016/B978-0-12-383874-2.00003-5
59. Sichelman T. *The Mathematical Structure of the Law* (2021).
60. Bjork G. Information gain when measuring an unknown qubit. *Eur J Phys*. (2018) 39:015403. doi: 10.1088/1361-6404/aa94e4
61. Norwich K. *Information, Sensation, and Perception*. Biopsychology.org (2003).
62. Alston L, Mueller B. Towards a more evolutionary theory of property rights. *Iowa Law Rev*. (2015) 100:2255–73.
63. Ruhl JB. The fitness of law: using complexity theory to describe the evolution of law and society and its practical meaning for democracy. *Vanderbilt Law Rev*. (1996) 49:1407–90.
64. Cooter R. Structural adjudication and the new law merchant: a model of decentralized law. *Int Rev Law Econ*. (1994) 14:215–31. doi: 10.1016/0144-8188(94)90020-5
65. Barzel Y. The capture of wealth by monopolists and the protection of property rights. *Int Rev Law Econ*. (1994) 14:393–409. doi: 10.1016/0144-8188(94)90023-X
66. Kennedy D. Form and substance in private law adjudication. *Harv Law Rev*. (1976) 89:1685–778. doi: 10.2307/1340104
67. Singer JW. The player and the cards: nihilism and legal theory. *Yale Law J*. (1984) 94:3–70. doi: 10.2307/796315
68. Kress K. Legal indeterminacy. *Calif Law Rev*. (1989) 77:283–337. doi: 10.2307/3480606
69. Lefstin JA. The measure of the doubt: dissent, indeterminacy, and interpretation at the federal circuit. *Hastings Law J*. (2007) 58:1025–1094.
70. Branting K, Weiss B, Brown B, Pfeifer C, Chakraborty A, Ferro L, et al. Semi-supervised methods for explainable legal prediction. In: *Seventeenth International Conference on Artificial Intelligence and Law*. (2019). p. 22–31. doi: 10.1145/3322640.3326723
71. Yu HF, Huang FL, Lin CJ. Dual coordinate descent methods for logistic regression and maximum entropy models. *Mach Learn*. (2011) 85:41–75. doi: 10.1007/s10994-010-5221-8
72. Kaplow L. A model of the optimal complexity of legal rules. *J Law Econ Organ*. (1995) 11:150–63.
73. Yang M. Coordination with flexible information acquisition. *J Econ Theory*. (2015) 158:721–38. doi: 10.1016/j.jet.2014.11.017
74. Rosser JB. Entropy and econophysics. *Eur Phys J Special Topics*. (2016) 225:3091–104. doi: 10.1140/epjst/e2016-60166-y
75. Georgescu-Roegen N. *The Entropy Law and the Economic Process*. Cambridge, MA: Harvard University Press (1971). doi: 10.4159/harvard.9780674281653
76. Lee BA, Smith H. The nature of coasean property. *Int Rev Econ*. (2012) 59:145–55. doi: 10.1007/s12232-012-0152-2
77. Parisi F. *The Asymmetric Coase Theorem: Dual Remedies for Unified Property*. SSRN (2001). doi: 10.2139/ssrn.264314
78. Luppi B, Parisi F. Toward an asymmetric coase theorem. *Euro J Law Econ*. (2011) 31:111–22. doi: 10.1007/s10657-010-9201-z
79. Barzel Y. *Economic Analysis of Property Rights, 2nd ed*. New York, NY: Cambridge University Press (1997). doi: 10.1017/CBO9780511609398
80. Cheung SNS. The transaction costs paradigm. *Econ Inq*. (1998) 36:514–21. doi: 10.1111/j.1465-7295.1998.tb01733.x

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

MINI REVIEW
published: 25 June 2021
doi: 10.3389/fphy.2021.665727

Negotiating the Descriptive–Normative Frontier of Complexity Research in the Anthropocene

Michael C. Leach *

Tilburg Law School, Tilburg University, Tilburg, Netherlands

This mini-review article offers a commentary on a singular analytical problem faced by legal scholars who use complexity theory and methods in legal research on climate change and the "Anthropocene". It positions such research as a subset of complexity scholarship in law, which is generally faced with the methodological and analytical challenge of negotiating and reconciling empirical description with normative prescription. It argues that this challenge is particularly acute for legal scholars writing on climate change and the Anthropocene. Using examples from scholars writing about "Earth systems law," it demonstrates how a heavy reliance on complexity-based empirical data as a source material for normative claim-making can distract scholars from important but difficult questions about normative legitimacy and how legal change happens at multiple levels. The special epistemological challenges posed by climate change and the Anthropocene should demand that scholars writing in this domain be especially mindful and explicit on how they link complexity descriptions to the normative claims they make, both for the sake of scientific credibility as well as for the legitimacy and viability of their propositions.

Keywords: complexity, Anthropocence, law, climate change, methodology and approaches

OPEN ACCESS

Edited by:
J. B. Ruhl,
Vanderbilt University, United States

Reviewed by:
Niko Soininen,
University of Eastern Finland, Finland
Barbara Cosens,
University of Idaho, United States

Correspondence:
Michael C. Leach
m.c.leach@tilburguniversity.edu

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: *08 February 2021*
Accepted: *24 May 2021*
Published: *25 June 2021*

Citation:
Leach MC (2021) Negotiating the Descriptive–Normative Frontier of Complexity Research in the Anthropocene. Front. Phys. 9:665727. doi: 10.3389/fphy.2021.665727

INTRODUCTION

Climate change and the "Anthropocene" [1] pose new challenges for understanding how complexity can usefully contribute to research in environmental law as legal discussions increasingly anticipate existential challenges to law in the face of heightened risks in the uncertain future of a rapidly changing planet [2, 3]. The advent of the Anthropocene has spawned calls for better and new understandings of humanity's relationship(s) to the planet along with considerable debate about how law should change in the future [4–6]. How complexity might contribute to these discussions is not self-evident. While it is uncontroversial that complexity-based analytical frameworks can provide insights into the dynamics of law's constitutive responsibility for both enabling and mitigating climate change, it is far less certain what normative contribution complexity models and analysis can or should make. This short article will focus attention on the boundary zone where the descriptive meets the normative in complexity studies in law to consider problems that arise when the two are not clearly distinguished in legal discussions about the Anthropocene. It will argue that muddying the descriptive and the normative leaves insufficient room for considerations about how law is to be socially legitimized as a part of any adaptive response to climate change. While there are opportunities for complexity research to serve as a frame of reference for law in the Anthropocene, its use comes with some dangers that demand the careful delineation of normative questions for which complexity on its own cannot be expected to provide answers.

COMPLEXITY'S NORMATIVITY FOR LEGAL RESEARCH

Legal scholars have been imagining complexity's utility for law and legal analysis for at least the past quarter century. Initially, it was seen as a compelling antidote to the limitations of traditional, doctrinal legal analysis. One of its earlier advocates, JB Ruhl, envisioned complexity as a key to a better understanding of law and society interactions and hoped that complexity would be a way to break free from the descriptive limitations of "reductionist" legal reasoning and the increasingly complicated legal architecture that such reasoning misguidedly constructed in pursuit of "absolute system predictability" [7]. By analogizing law to a dynamic, interactive complex system, it became possible to describe in new ways for how and why law and legal phenomena change over time as outcomes of processes of systemic self-assembly and emergence, without having to rely on mechanistic and linear doctrinal or institutional explanations that often struggle to account for observable reality [8].

Ever since then, complexity theorists have offered new perspectives and vocabularies for understanding a variety of legal phenomena over the past decades, often demonstrating its novel utility for certain kinds of legal research. However, the application of complexity to law generally has never been clear-cut, and its utility has always been bounded by the difficulty of distinguishing and negotiating the relationship between its descriptive capacity with the normative concerns of law and legal reasoning. It is not clear, for instance, how to best describe law as a systemic object of complexity analysis: is law a system of interacting norms [9]; or an institutional administrative system governing complex societies [10]; or a site of interaction between norms and larger social or geophysical systems [11]; or something else? Describing law in complex terms as an evolving, nonlinear complex system that produces stochastic and probabilistic outcomes [12] can run against the grain of law's self-image as a stabilizing force that ensures future certainty in social relations or that anchors political–legal orders around goals and policy objectives in the legitimate exercise of authority [13]. Complexity's fit with law, in other words, is imperfect.

This imperfection can pose risks when certain realities are given normative readings when they are described as complex. As social and legal constructs [14], legal institutions and rule systems present thorny normative problems that complexity-based descriptions about law struggle to engage with. While there is little that is objectionable with using complexity to understand how normative goals are pursued in legal systems [15], the same cannot be said of deriving normative conclusions from complex descriptions of law or other empirical realities. Complexity-based proposals that legal systems should have as their goal system stability or robustness teeter on that line [16]. Propositions that legal systems benefit from being more "resilient," or by increasing their adaptative capacities by promoting lower level systemic innovation [7], etc., are unproblematic only when they do not presume to know what normative outcomes such as systemic features should produce. One cannot simply assume that any legal or legislative intervention that is inspired by complexity will be "good" or appropriate simply because it respects complex system dynamics [17]. Complexity does not lend itself easily to normative claim-making because it is not designed for that purpose [8].

This awkwardness of complexity's fit parallels rich ongoing debates about the fluid and difficult relationship between law and science generally, much of which tends to emphasize that they are distinct realms [18]. One can consider how legal systems in some ways maintain this distinction, such as with mechanisms like the precautionary principle in environmental law that recognizes the limits of scientific certainty in predicting future risks and serves as a counterweight to narrow science-based decision-making processes by experts [19]. On the other hand, Sheila Jasanoff has pointed out how legal scholarship typically has a "deep rooted commitment to the existence of objective facts" [20, 21] and Robin Feldman has cautioned against law and policymakers relying too heavily on science as an "authoritative other" whose "superior reliability allows us to indulge in the fantasy that it can reduce us from the discomfort of legal dilemmas and deliver the certainty we crave" [22]. "The appropriate role for science," Feldman argues, "should be to test the assumptions that underlie legal rules, rather than prescribe them" [22]. The inverse has also been argued that it is inappropriate for the law and courts to act as arbiters of scientific truth, and Katalin Sulyok has observed how international courts are grappling with this problem and increasingly ruling on the legitimacy of competing scientific evidentiary claims only from the restricted perspective of legal rules [23]. The ambiguity of this law/science divide becomes similarly problematic when complexity is used in legal scholarship and commentary about climate change and the Anthropocene. Even though the bread and butter of much complexity research is systemic uncertainty and nonlinearity, descriptions of largely positivist complex realities in legal scholarship can offer temptations of "superior reliability" and material certainties that can distract from the equally complex social and political structures, processes, and mechanisms by which law and norms are legitimized in the human society.

COMPLEXITY IN ENVIRONMENTAL AND CLIMATE CHANGE LEGAL SCHOLARSHIP

Environmental law has attracted the attention of complexity scholars since as early as the 1990s [24], in part because its relative incoherence [25] makes it fertile ground for complexity research: it is methodologically plural, interdisciplinary, and concerns itself with intricately interconnected legislative and treaty regimes that address difficult transboundary and multi-jurisdictional problems. Finding a home within this domain, scholars have turned to complexity as a theoretical frame to describe environmental governance regimes [26–28], to anticipate how law will change as social values and interests compete in the advent of climate change [29], and to critically reflect on law's present inadequacy at tackling the magnitude of environmental problems around the world [30, 31]. In arguing that international environmental law "exhibits some key characteristics of a complex adaptive system," Kim and Mackey, for instance, propose that in spite of its capacity for self-organization, international environmental law is mal- or mis-adaptive and needs to be "more appropriately align[ed] with the functioning of the Earth system itself" [32].

Legal scholars writing in the "Earth systems" domain are emblematic of a certain temptation in environmental law and complexity scholarship to intermingle the normative and the descriptive in writing about the climate change and the Anthropocene. The Earth systems approach treats the Earth as a single complex and dynamic system [33] that is characterized by a limited number of biochemical subsystems that human activity is pushing out of balance in ways that threaten the biosphere's capacity to maintain life [34, 35]. The paradigm of "Earth system law" seeks out normative determinations about what coordinated human behavior would best manipulate the constitutive flows of energy and matte of those subsystems in ways that can maintain biospheric stability for human and nonhuman life. In this way, law should better "reflect" Earth system transformations and be "simultaneously reformative or prescriptive … to proactively enable and govern human-dominated Earth-system transformations for sustainability" [30]. Such an approach implies that the structural benefits of normative centralization around the "Earth system" would be enough to legitimize the global normative project of "aligning with the functioning of the Earth system."

From a normative standpoint, the complexity-based description of the "Earth system" produces a powerful and unambiguous narrative that often reduces climate science down to specific reference points like "planetary boundaries" [36] or "tipping points" [37]. The presentation of these reference points as certainties places any normative objective other than optimizing planetary biochemical systems into stark relief as concerns of secondary importance, thereby attempting to limit the scope and range of political solutions that should be available to policymakers in the present [38]. While using environmental and biochemical complexity as a reference or base for lawmaking is rhetorically compelling, assuming that it can translate into a self-evident global legal project disregards the complexity of the political mechanisms by which human polities around the world determine what should be done, what constitutes acceptable risk, or what relationship science should have with those questions [19, 39]. It also disregards how choices made about how to empirically describe climate change as a particular kind of complex process itself produces specific normative framings that can conflict with those offered by different models. Consider, for instance, how historical accounts of the complex human processes responsible for global warming today might offer very different prescriptions for what can or should be done on the basis of retrospective responsibility and distributive compensatory justice compared to future-oriented Earth system perspectives that aim to adjust human behaviors in ways that manipulate planetary biochemical systems [40, 41].

Things like reparations or debts owed by those who have benefitted from industrialization to present and future generations who will suffer its consequences do not fit the specific focus of the Earth system model. Furthermore, the planetary scale of its abstractions are difficult to translate down to middle- or microlevels where policy and legal actors must engage in compromise and horse trading among the many conflicting norms, values, opinions, needs, and interests that come to bear around climate change in different societies. In this sense, while laudable, the ambition of devising a coherent global normative framework around a particular complexity-based model of the world understates the difficult problem of how law is legitimized while overstating the universality of the problem it addresses. The call for law to be "align(ed) with the functioning of the Earth system" equates to a call to reorder national legal and political priorities and processes through some kind of a global planetary regulatory goal-setting process. The powerful rhetorical emphasis that the model provides that this must be done overshadows questions about how and in what form it could be done, who would benefit or lose the most in so doing, and who or what could have sufficient authority and legitimacy to accomplish it in the absence of any genuinely self-aware "global" public [19, 42, 43].

None of this is to say that Earth system models of climate change are wrong, nor that using them for legal and normative purposes is also wrong. Rather, the concern here is with the reliance on complexity-based descriptions to make normative prescriptions in ways that obscure the norm-making role that modelers and scholars play whenever they interpret a given empirical reality as "complex." It is quite artificial to separate such models from the equally complex sociopolitical contexts in which they are made and those to which they are expected to be applied. Victor Galaz has referred to this as the analytical challenge of "double complexity" [44], which scholars in the "socio-ecological systems" subfield have tried to reconcile in singular system frames, although with no consensus on how exactly to do this [45]. Scholars have also tried to delicately navigate this terrain by limiting themselves to asking questions on how law plays a facilitating or obfuscating role in processes of legal change and reform in the face of climate change, and evaluating whether or to what degree law contributes to or weakens the capacity of social and governance systems to adapt or be "resilient" (i.e., capable of absorbing disturbances without having to induce structural systemic changes) [46, 47]. Importantly, however, such perspectives cannot rely on the descriptive systemic frames they deploy to address difficult normative questions of whether or to what degree social and political change and/or stability and continuity are good or desirable in the first place [48]. Engagement along these lines will always have to circle back authorial choices about what is or is not desirable and beg bigger questions about who benefits or loses when a given legal system is made more or less "resilient" or "adaptive" [49, 50], or what kind of "resilience," if any, is desirable for a given legal system [51], or what is worth keeping and preserving, and what can "we" afford to lose and change [52]? These are substantive and often deeply political questions that descriptive complexity models of planetary systems cannot provide answers to on their own, largely because they are not designed to account for how social factors like knowledge, power, agency, and conflict factor into how different human societies determine their normative pRefs. [48, 53].

RECONSIDERING COMPLEXITY'S USES FOR LEGAL SCHOLARSHIP IN THE ANTHROPOCENE

The brief discussion above cautions against relying on complexity models of climate change to meaningfully prescribe legal decision-making because doing so can miss or deflect attention away from the normative dimensions of the social and political contexts in which lawmaking occurs. These concerns are augmented further by orders of magnitude if the vastness of the Anthropocene is appreciated as being a kind of "hyper-object" that may be beyond our powers to describe accurately [54]. In this light, the role of climate change system modelers then becomes one of authoritatively reducing the unimaginable to the (merely) complex but in ways that are normatively consequential and should be treated with care by legal scholars. While one should be skeptical of the ability of complexity analyses to provide specific normative answers to the Anthropocene, they can nevertheless play a significant role in improving what Timothy Clark has called "scalar literacy" around law and climate change [55] and to offer insights into the ranges of options and possibilities available for decision-makers to choose from. Furthermore, if the Anthropocene truly does lie beyond our descriptive capacities, then complexity may be more useful to help elucidate the limits of human knowledge [56] and to draw the attention of lawyers and politicians to difficult problems where normative choice-making is needed in the face of uncertainty [26]. In such a role, complexity's utility comes not from its ability to portray the world more accurately or with greater normative truth, but from how it can assist with thinking about conflicting and contradictory environmental and social phenomena that coexist on multiple levels and scales. Clark offers the global market for biofuels as an example, something that is simultaneously beneficial and destructive, depending on which overlapping narrative about emissions reductions and deforestation one ascribes to [55].

What complexity models cannot be relied upon to do on their own, however, is to speak to the normative dimensions of the empirical findings they generate. This is because what makes them useful and meaningful for law is determined not by the strength of their data, but by the nature of the socialized domains in which modelers and authors use them as interventions for particular legal purposes [45]. Legal scholars interested in complexity therefore have an ethical responsibility to resist the temptation that the empirical strengths of their complexity models give them any special methodological ability or authority to better determine what is normatively appropriate or better for the society [57]. Taking complex social and environmental realities into account may better help policymakers, but it cannot guarantee that they will be more able to discern what is normatively desirable. Indeed, with their emphases on empirical uncertainty and contingent variability, they may actually make it harder for them.

Nothing in this article should be construed as proposing that complexity scholars should never make normative claims about law, of course. Rather, the point is that it would be inappropriate for them to use the complexity of environmental realities as a rationale for de-complexifying political choice-making in ways that disregard the equally complex realities of the human political society and legal systems through which such realities are made meaningful. In effect, this is a call to remember that law is a politicized realm, and that scholars who are interested in complexity-based understandings of law in the Anthropocene cannot afford to disregard how the normative–descriptive interface affects the scientific credibility of their analyses and the political legitimacy of their prescriptions.

AUTHOR CONTRIBUTIONS

ML is the sole author of this manuscript.

FUNDING

This research was conducted as a part of the "Constitutionalizing in the Anthropocene" Project at Tilburg Law School, Department of Public Law and Governance, funded through the "Sectorplan Social Sciences and Humanities" of the Ministry of Education, Culture and Science (Ministerie van Onderwijs, Cultuur en Wetenschaap) of the Government of The Netherlands. Funds for open access fees were provided by Tilburg Law School, Tilburg University.

REFERENCES

1. Crutzen PJ. Geology of Mankind. *Nature* (2002) 415:23. doi:10.1038/415023a
2. Burdon PD. Ecological Law in the Anthropocene. *Transnational Leg Theor* (2020) 11(1-2):33. doi:10.1080/20414005.2020.1748483
3. Birrell K, and Matthews D. Laws for the Anthropocene: Orientations, Encounters, Imaginaries. *L Critique* (2020) 31:233–8. doi:10.1007/s10978-020-09282-8
4. Latour B. *Facing Gaia: Eight Lectures on the New Climatic Regime*. Cambridge: Polity Press (2017).
5. Haraway DJ. *Staying with the Trouble: Making Kin in the Chthulucene*. Durham NC: Duke University Press (2016).
6. Biermann F, and Lövbrand E. *Anthropocene Encounters: New Directions in Green Political Thinking*. Cambridge: Cambridge University Press (2019).
7. Ruhl JB. Complexity Theory as a Paradigm for the Dynamical Law-And-Society System: A Wake-Up Call for Legal Reductionism and the Modern Administrative State. *Duke L J* (1996) 45(5):849. doi:10.2307/1372975
8. Hornstein DT. Complexity Theory, Adaptation, and Administrative Law. *Duke L J* (2005) 54(4):913. doi:10.2307/40040504
9. Post DG, and Eisen MB. How Long Is the Coastline of the Law? Thoughts on the Fractal Nature of Legal Systems. *J Leg Stud* (2000) 29:545–84. doi:10.1086/468085
10. Leach M, Scoones I, and Stirling A. Governing Epidemics in an Age of Complexity: Narratives, Politics and Pathways to Sustainability. *Glob Environ Change* (2010) 20:369–77. doi:10.1016/j.gloenvcha.2009.11.008
11. Ebbesson J, and Hey E. Introduction: Where in Law Is Social-Ecological Resilience? *Ecol Soc* (2013) 18(3):24. doi:10.5751/es-05750-180325
12. Auyang SY. *Foundations of Complex-System Theories*. Cambridge: Cambridge University Press (1998).

13. Ebbesson J. The Rule of Law in Governance of Complex Socio-Ecological Changes. *Glob Environ Change* (2010) 20(3):414–22. doi:10.1016/j.gloenvcha.2009.10.009
14. Room G. Complexity, Power and Policy. In: P Cairney and R Geyer, editors. *Handbook on Complexity and Public Policy*. Cheltenham: Edward Elgar (2015).
15. Ruhl J, and Katz DM. Measuring, Monitoring, and Managing Legal Complexity. *Iowa L Rev* (2015-2016) 101:191.
16. Farber DA. Probabilities Behaving Badly: Complexity Theory and Environmental Uncertainty. *U C Davis L Rev* (2003-2004) 37:145.
17. Le QB, Seidl R, and Scholz RW. Feedback Loops and Types of Adaptation in the Modelling of Land-Use Decisions in an Agent-Based Simulation. *Environ Model Softw* (2012) 27-28:83–96. doi:10.1016/j.envsoft.2011.09.002
18. Jasanoff S. A Constitutional Moment: Science and Social Order at the Millenium. In: B Joerges and H Nowotny, editors. *Social Studies of Sciences and Technology: Looking Back, Ahead*. Dordrecht: Kluwer Academic Publishers (2003).
19. Peel J. *Science and Risk Regulation in International Law*. Cambridge: Cambridge University Press (2010).
20. Jasanoff S. The Eye of Everyman: Witnessing DNA in the Simpson Trial. *Soc Stud Sci* (1998) 28(5-6):713. doi:10.1177/030631298028005003
21. Jasanoff S. *Science at the Bar: Law, Science, and Technology in America*. New York: Oxford University Press (1995).
22. Feldman R. *The Role of Science in Law*. Abingdon: Oxford University Press (2009).
23. Sulyok K. *Science and Judicial Reasoning: The Legitimacy of International Environmental Adjudication*. Cambridge: Cambridge University Press (2021).
24. Ruhl J. Thinking of Environmental Law as a Complex Adaptive System: How to Clean up the Environment by Making a Mess of Environmental Law Houston Law Review. Houston Law Review (1997-1998) 34:933.
25. Fisher E, Lange B, Scotford E, and Carlarne C. Maturity and Methodology: Starting a Debate about Environmental Law Scholarship. *J Environ L* (2009) 21(2):213–50. doi:10.1093/jel/eqp012
26. J Norberg and G Gumming. *Complexity Theory for a Sustainable Future*. New York City: Columbia University Press (2008).
27. Rockström J, Steffen W, Noone K, Persson A, Chapin FS, Lambin EF, et al. A Safe Operating Space for Humanity. *Nature* (2009) 461(2):472–5. doi:10.1038/461472a
28. Craig RK. "Stationarity Is Dead" - Long Live Transformation: Five Principles for Climate Change Adaptation Law. *Harv Environ L Rev* (2010) 34(1):9.
29. Ruhl J. Climate Change Adaptation and the Structural Transformation of Environmental Law. *Environ L* (2010) 40:363.
30. Kotzé LJ, and Kim RE. Earth System Law: The Juridical Dimensions of Earth System Governance. *Earth Syst Governance* (2019) 1:100003. doi:10.1016/j.esg.2019.100003
31. Young OR. *Governing Complex Systems: Social Capital for the Anthropocene*. Cambridge, MA: The MIT Press (2017).
32. Kim RE, and Mackey B. International Environmental Law as a Complex Adaptive System. *Int Environ Agreements* (2014) 14:5–24. doi:10.1007/s10784-013-9225-2
33. Steffen W, Richardson K, Rockström J, Schellnhuber HJ, Dube OP, Dutreuil S, et al. The Emergence and Evolution of Earth System Science. *Nat Rev Earth Environ* (2020) 1:54–63. doi:10.1038/s43017-019-0005-6
34. Bridgewater P, Kim RE, and Bosselmann K. Ecological Integrity: A Relevant Concept for International Environmental Law in the Anthropocene? *Yearb Int Environ L* (2014) 25(1):61–78. doi:10.1093/yiel/yvv059
35. Petersmann M-C. *Earth System Law and Sympoiesis: The Earth, its Subjects and the Law* (Forthcoming 2021).
36. Biermann F. Planetary Boundaries and Earth System Governance: Exploring the Links. *Ecol Econ* (2012) 81:4–9. doi:10.1016/j.ecolecon.2012.02.016
37. Eisenhauer DC. Tipping Points in the Anthropocene: Crafting a Just and Sustainable Earth. *J Extreme Events* (2017) 4(1):1750004–1. doi:10.1142/s234573761750004x
38. Kennedy D. Challenging Expert Rule: The Politics of Global Governance. *Sydney L Rev* (2005) 27:5.
39. Webb J. Law, Ethics, and Complexity: Complexity Theory & (And) the Normative Construction of Law. *Cleveland State L Rev* (2004) 52(1 & 2):227.
40. Bonneuil C, and Fressoz J-B. *The Shock of the Anthropocene*. London: Verso (2016).
41. Moore JW. *Capitalism in the Web of Life: Ecology and the Accumulation of Capital*. London: Verso (2015).
42. Zürn M. Global Governance and Legitimacy Problems. *Government and Opposition* (2004) 39(2):268. doi:10.1111/j.1477-7053.2004.00123.x
43. Esty DC. Good Governance at the Supranational Scale: Globalizing Administrative Law. *Yale L J* (2005-2006) 115:1490.
44. Galaz V. Double Complexity: Information Technology and Reconfigurations in Adaptive Governance. In: E Boyd and C Folke, editors. *Adapting Institutions: Governance, Complexity and Social-Ecological Resilience*. Cambridge: Cambridge University Press (2012).
45. Cumming GS. *Spatial Resilience in Social-Ecological Systems*. Heidelberg: Springer (2008).
46. B Cosens and LH Gunderson. *Practical Panarchy for Adaptive Water Governance: Linking Law to Social-Ecological Resilience*. Cham: Springer Nature (2018).
47. Holling CS, and Gunderson LH. Resilience and Adaptive Cycles. In: LH Gunderson and CS Holling, editors. *Panarchy: Understanding Transformations in Human and Natural Systems*. Washington D.C.: Island Press (2002).
48. Olsson L, Jerneck A, Thoren H, Persson J, and O'Byrne D. Why Resilience Is Unappealing to Social Science: Theoretical and Empirical Investigations of the Scientific Use of Resilience. *Sci Adv* (2015) 1(4):e1400217. doi:10.1126/sciadv.1400217
49. Cote M, and Nightingale AJ. Resilience Thinking Meets Social Theory. *Prog Hum Geogr* (2012) 36(4):475–89. doi:10.1177/0309132511425708
50. Chandler D. Resilience and the End(s) of the Politics of Adaptation. *Resilience* (2019) 7(3):304–13. doi:10.1080/21693293.2019.1605660
51. Powell NS, Larsen RK, and van Bommel S. Meeting the "Anthropocene" in the Context of Intractability and Complexity: Infusing Resilience Narratives with Intersubjectivity. *Resilience* (2014) 2(3):135–50. doi:10.1080/21693293.2014.948324
52. Duit A, Galaz V, Eckerberg K, and Ebbesson J. Governance, Complexity, and Resilience. *Glob Environ Change* (2010) 20(3):363–8. doi:10.1016/j.gloenvcha.2010.04.006
53. Davidson DJ. The Applicability of the Concept of Resilience to Social Systems: Some Sources of Optimism and Nagging Doubts. *Soc Nat Resour* (2010) 23:1135–49. doi:10.1080/08941921003652940
54. Morton T. *Hyperobjects: Philosophy and Ecology after the End of the World*. Minneapolis: University of Minnesota Press (2013).
55. Clark T. *The Value of Ecocriticism*. Cambridge: University of Cambridge Press (2019).
56. Smil V. *The Earth's Biosphere: Evolution, Dynamics, and Change*. 2nd edn. Cambridge MA: MIT Press (2003).
57. Cilliers P. Complexity, Ethics and Justice. In: R Preiser, Editor. *Paul Cilliers Critical Complexity: Collected Essays*. Berlin: Walter de Gruyter GmbH (2016).

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 06 July 2021
doi: 10.3389/fphy.2021.665719

A Global Community of Courts? Modelling the Use of Persuasive Authority as a Complex Network

D. Hoadley[1]*, M. Bartolo[2], R. Chesterman[3], A. Faus[3], W. Hernandez[1], B. Kultys[1], A. P. Moore[1,2,4], E. Nemsic[1], N. Roche[1,2,4], J. Shangguan[1], B. Steer[5], K. Tylinski[1] and N. West[1]

[1]Mishcon de Reya LLP, London, United Kingdom, [2]Department of Computer Science, UCL, London, United Kingdom, [3]vLex Justis Limited, London, United Kingdom, [4]School of Management, UCL, London, United Kingdom, [5]Department of Computer Science, Queen Mary University, London, United Kingdom

There is a growing discussion in the legal literature of an emerging global community of courts composed of a network of increasing judicial dialogue across national borders. We investigate the use of foreign persuasive authority in common law countries by analyzing the network of citations to case law in a corpus of over 1.5 million judgments given by the senior courts of twenty-six common law countries. Our corpus of judgments is derived from data available in the vLex Justis database. In this paper we aim to quantify the flow of jurisprudence across the countries in our corpus and to explore the factors that may influence a judge's selection of foreign jurisprudence. Utilization of foreign case law varies across the countries in our data, with the courts of some countries presenting higher engagement with foreign jurisprudence than others. Our analysis shows that there has been an upward trend in the use of foreign case law over time, with a marked increase in citations across national borders from the 1990s onward, potentially indicating that increased digital access to foreign judgments has served to facilitate and promote comparative analysis. Not only has the use of foreign case law generally increased over time, the factors that may influence the selection of case law have also evolved, with judges gradually casting their research beyond the most influential and well-known foreign authorities. Notwithstanding that judgments emanating from the United Kingdom (chiefly from the courts of England and Wales) constitute the most frequently consulted body of jurisprudence, we find evidence that domestic courts favor citing the case law of countries that are geographically proximal.

Keywords: caselaw, complex network, law, graph algorithms, complex adaptive systems (CAS), foreign laws

OPEN ACCESS

Edited by:
J. B. Ruhl,
Vanderbilt University, United States

Reviewed by:
Luca Verginer,
ETH Zurich, Switzerland
Pierpaolo Vivo,
King's College London, United Kingdom

****Correspondence:***
D. Hoadley
daniel.hoadley@mishcon.com

Specialty section:
This article was submitted to Interdisciplinary Physics, a section of the journal Frontiers in Physics

Received: *08 February 2021*
Accepted: *26 May 2021*
Published: *06 July 2021*

Citation:
Hoadley D, Bartolo M, Chesterman R, Faus A, Hernandez W, Kultys B, Moore AP, Nemsic E, Roche N, Shangguan J, Steer B, Tylinski K and West N (2021) A Global Community of Courts? Modelling the Use of Persuasive Authority as a Complex Network. Front. Phys. 9:665719. doi: 10.3389/fphy.2021.665719

1 INTRODUCTION

There is a growing discussion in the legal literature of an emerging "global community of courts" composed of a network of increasing judicial dialogue across borders [1–3]. Here we investigate the use of foreign persuasive authority in common law systems by analyzing the network of citations to foreign case law in a corpus of over 1.5 million judgments of the senior courts of twenty-six common law countries. Our corpus of judgments is derived from the vLex Justis database [1]. In this paper we

[1]vLex Justis (www.vlex.com) is one of the largest online collection of legal information, with more than 120 million documents spanning 130 jurisdictions.

aim to quantify the flow of jurisprudence across the countries in our corpus and to explore factors that may influence a judge's selection of foreign case law.

A fundamental feature of common law legal systems is the doctrine of precedent, which places a binding obligation on judges to follow principles established by coordinate and superior courts in earlier similar cases. This binding obligation to follow the decisions of courts of equal and higher standing (*mandatory* authorities) stands in contrast to the concept of persuasive authority (*optional* authority), the most common example of which are decisions of foreign national courts [4].

The concept of persuasive authority is well-known but imprecise [5], yet there has been growing consensus among legal scholars since the late-1990s that "more and more courts, particularly within the common law world, are looking to the judgments of other jurisdictions" [1].

The practice of cross-jurisdictional citation of persuasive authority sits within the broader context of the emergence of what Slaughter describes as a "global community of courts" [2]–the formation of which has been driven by a range of factors, including increasing similarities between the issues facing courts around the world; the international nature of human rights and the proliferation of international courts and tribunals; advances in technology and vastly improved accessibility of foreign jurisprudence; and increased personal contact among judges [1].

Slaughter acknowledges that the phenomenon of "cross-pollination" of judicial thinking via the citation of one nation's jurisprudence by another is not new and is well established in the Commonwealth [2]. However, Claire L'Heureux-Dubé, a former justice of the Canadian Supreme Court, observes that the contemporary "process of international influences has changed from reception to dialogue. Judges no longer simply receive the cases of other jurisdictions and then apply them or modify them for their own jurisdiction."[1] Instead, according to L'Heureux-Dubé, "... cross-pollination and dialogue between jurisdictions is increasingly occurring. As judgments in different countries build on each other, mutual respect and dialogue are fostered among appellate courts. Judges around the world look to each other for persuasive authority, rather than some judges being "givers" of law while others are "receivers". Reception is turning to dialogue [1]."

The cross-pollination of judicial thinking via the citation of optional, yet persuasive, foreign judgments occurs horizontally between nations independently of the doctrine of precedent as opposed to vertically between nations and the decisions of their supranational counterparts by which the national court is either bound[2] or at the very least obligated to take into account.[3]

There is ample support for L'Heureux-Dubé's conception of judicial dialogue between nations to be found in the decisions of senior common law courts. For example, in the United Kingdom House of Lords case of *Fairchild v Glenhaven Funeral Services Ltd.* [4], which concerned the issue of causation of mesothelioma arising from the appellant's exposure to asbestos during different periods of employment in breach of each employer's duty of care, Lord Bingham, having conducted a survey of case law from Australia, South Africa, the United States, France, Germany and Canada, said: "Development of the law in this country cannot of course depend on a head-count of decisions and codes adopted in other countries around the world ... The law must be developed coherently, in accordance with principle, so as to serve, even-handedly, the ends of justice. If, however, a decision is given in this country which offends one's basic sense of justice, and if consideration of international sources suggests that a different and more acceptable decision would be given in most other jurisdictions, whatever their legal tradition, this must prompt anxious review of the decision in question. In a shrinking world ... there must be some virtue in uniformity of outcome whatever the diversity of approach in reaching that outcome [5]."

Similar sentiments have been expressed, extra-judicially, by former justices of the Canadian Supreme Court [6] and the High Court of Australia [7]. However, the phenomenon of participation in cross-jurisdictional dialogue is not universally embraced. For example, in the United States Supreme Court case of *Foster v Florida*, [6] in which a death-row inmate sought a writ of certiorari on the grounds that the lengthy delay between his sentencing and execution constituted a violation of his Eighth Amendment rights against cruel and unusual punishment, Justice Thomas denigrated Justice Breyer's willingness to cite foreign authorities, stating: "While Congress, as a *legislature*, may wish to consider the actions of other nations on any issue it likes, this Court's Eighth Amendment jurisprudence should not impose foreign moods, fads, or fashions on Americans [7]."

The discourse in academic articles and the cases themselves present a mixed picture where the use of foreign case law is concerned. The courts of some countries, notably Australia and Canada, appear to have taken advantage of increased access to case law from around the world to engage in comparative analysis and dialogue. However, other common law jurisdictions, most notably the US, appear to have adopted a more restrictive approach to the citation of foreign jurisprudence, rarely reaching beyond their borders when seeking guidance on legal issues [8].

A lack of access to judgment data the spans multiple common law systems has made it difficult to analyze the use of foreign case law at scale. Most earlier work has therefore concentrated on the analysis of interactions with foreign case law by a specific court or a specific territorial jurisdiction, such as the United States. This paper utilizes the case law citation network derived from a

[2]For example, the courts of member states of the European Community are bound by the decisions of the Court of Justice of the European Union.

[3]For example, under Section 2 (1) of the Human Rights Act 1998, United Kingdom courts are required to take judgments of the European Court of Human Rights into account when determining a question related to any of the rights conferred by the European Convention on Human Rights.

[4][2003] 1 AC 32, HL(E).

[5]*Fairchild v Glenhaven Funeral Services Ltd.*, at [32].

[6]537 US 990 (2002).

[7]*Ibid*, at 991 (emphasis supplied).

substantial corpus of senior and appellate court judgments from twenty-six common law countries to examine how, and to what extent, domestic courts in common law jurisdictions make use of foreign case law. We anchor our analysis in the overarching theme that there is an emerging global community of courts composed of a network of judicial dialogue flowing between national courts via the mechanism of citation to persuasive foreign case law.

We make five core findings. First, the use of foreign case law has followed a consistent upward trend throughout the 20th century to the present, with a pronounced increase in foreign case law utilization from 1990 onward. The timing of this increase appears to correspond to the rise of digital platforms that facilitate more comprehensive and low-cost systems of case law dissemination and retrieval. Second, although all of the countries we examine participate in the citation of foreign case law, the extent to which they do varies, with some countries making more frequent reference to foreign cases than others. Third, while there is evidence of a historical preference among judges to cite foreign cases that are highly influential in their own domestic jurisprudence, this pattern has given way to citation behavior that potentially indicates a shift in attitudes toward the citation of less well-known cases. This shift corresponds to a rapid expansion in the citation networks that is likely attributable to increased online accessibility to case law. Fourth, domestic courts have a general tendency to cite to the jurisprudence of legal systems that are geographically proximal. Finally, in aggregate we find support for the proposition that there is an *emerging* global community of courts held together by an increasingly seamless web of foreign case law citation. However, that community is dominated by a clique of countries–Australia, Canada, the United Kingdom (chiefly, England and Wales), New Zealand and the United States–that exchange dialogue between themselves and attract the majority of inward citations from the rest of the community.

The remainder of this paper is organized as follows. **Section 2** provides an outline of previous work applying network analysis to judgment corpora. In **Section 3** we present the framework for our study and analyze the global properties of the cross-jurisdictional citation network. **Section 4** utilizes the network of citations to explore the extent to which a judge's decision to cite a particular foreign case is guided by how influential the case is; the degree to which is it well-grounded in established precedent and the country from which it emanates.

2 RELATED WORK

There is a growing line of research that seeks to represent judgment corpora, and their underlying citation structures, as complex networks [9]. Several earlier studies that utilize citation network analysis on judgments have focused on American case law. These studies have aimed to measure the influence of US Supreme Court and Federal Courts of Appeals judges [10,11] and analyze citation patterns between state appellate courts [12–14]. Attention has been directed toward understanding the general internal network dynamics of US Supreme Court decisions [15,16], tracing the evolution of legal principle in that court's body of jurisprudence [17], measuring the importance of its precedents [18] and evaluating how strategic interactions between Supreme Court justices during the court's bargaining process affects citations to precedent in the court's final opinion [19].

Network analysis of citations have also been applied to Canadian [20] and Australian [21] case law, in addition to judgments of international courts and tribunals, including the European Court of Human Rights [22], the Court of Justice of the European Union [23], the International Criminal Court [24,25], the Appellate Body of the World Trade Organization [26] and the International Court of Justice [27].

Other work focused on quantitatively analyzing the use of foreign authority by domestic courts is scarce. A comparative analysis of engagement with foreign case law in the decisions of the highest courts of the United States, Canada and Australia suggests that judges in Canada and Australia promote the use of foreign case law as persuasive authority, particularly in the context of criminal cases, using the case law of other countries both to defend arguments and to refute them, and to clarify a position through comparison and contrast with domestic case law [8]. A survey of US federal court case law citation practice between 1945 and 2005 revealed that citation of foreign decisions is a relatively rare phenomenon in the United States that is generally confined to cases where international issues are squarely presented by the facts [28]. In Australia, an analysis of decisions of the High Court of Australia between 2015 and 2016 found that court tended to cite foreign judicial decisions emanating from jurisdictions that reflect values common to the Australian legal system, particularly where the cited case considers statutory language that is similar to that which is in dispute [29]. Most recently, an analysis of United Kingdom Supreme Court decisions found that citations to foreign jurisprudence occurred in just under 30% of that court's decisions [30].

The work outlined above focuses on citation activity that is specific to an individual court (e.g., the United States Supreme Court, the European Court of Human Rights etc.) or to an individual territorial jurisdiction (e.g., Australia, Canada, the United States etc). This paper provides a cross-jurisdictional perspective on judicial citation interactions between *multiple* countries and their respective senior courts.

3 THE COMMON LAW CITATION NETWORK: GLOBAL PROPERTIES

To demonstrate how citations to earlier cases in judicial decisions can be modeled as a network, consider the following example based on a small selection of cases concerning the ability of an accused person to challenge hearsay evidence admitted against them in criminal proceedings. In *Crawford v Washington* [8] the United States Supreme Court held that the Confrontation Clause

[8] 541 US 36 (2004).

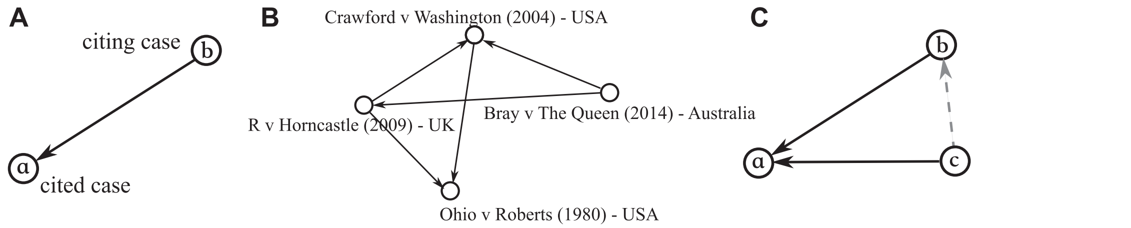


FIGURE 1 | The citation network. **(A)** Nodes *a* and *b* representing cases linked by a directed edge from citing to cited node. **(B)** The graph for four hearsay cases decided in Australia, the United Kingdom and the United States with cross-jurisdictional citation edges. **(C)** Triads between three citation nodes. It is considered a Triangle if the dotted edge *c-b* is included in the graph.

in the Sixth Amendment prohibited the introduction of testimonial hearsay as evidence at trial unless the declarant was unavailable to give evidence in person and the accused had the prior opportunity to cross-examine the declarant. In its reasoning the court cited, and overruled, an earlier decision of the United States Supreme Court addressing the Confrontation Clause–*Ohio v Roberts* [9]. The United States decisions in *Crawford* and *Roberts* were both subsequently considered, as persuasive authority, by the United Kingdom Supreme Court in *R v Horncastle* [10]. Finally, the United States decision in *Crawford* and the United Kingdom decision in *Horncastle* were considered by the Victoria Court of Appeal in Australia in *Bray v The Queen*[11]. The relationships between the cases in this example network are shown in **Figure 1B**.

Our corpus consists of 1,559,807 judgments covering a rich spread of civil and criminal matters given between 1717 and 2020 by the senior and appellate courts of twenty-six common law systems. In addition to setting out the court's reasoning for its determination in a given case, the judgments cite principles settled in earlier cases emanating from their own respective domestic legal systems and, to a lesser extent, from cases decided by courts in foreign jurisdictions. These judgments contain citations to 853,287 unique judgments. Citations in the judgments were identified using a proprietary rules-based engine developed by vLex Justis that allows for ambiguous reporter series abbreviations and the accurate resolution of malformed references. Detected citations are reconciled to unique case entities using a database of parallel citations developed and maintained by vLex Justis.

As outlined in **Section 1**, a fundamental feature of common law systems is the principle that judges are bound by the earlier decisions of judges in superior courts. Notwithstanding that the legal systems of the countries represented in our corpus are not identical, they all broadly conform to a similar hierarchical court structure. *Inferior* or *lower* courts sit at the bottom of that hierarchy. In general, inferior courts are concerned with questions of fact and provide the venue within which the majority of legal disputes are resolved at a local level. For example, in the context of the English and Welsh legal system, the majority of civil disputes are heard in the county court (the lowest civil court) and most criminal matters are heard in the magistrates' court (the lowest criminal court). In contrast, *senior* or *higher* courts, which are located higher up the court hierarchy, are generally concerned with questions of law and exercise supervisory and appellate jurisdiction over the courts below them in the hierarchy. All of the countries represented in our corpus have a "court of last resort" that sits at the apex of the court hierarchy. Apex courts in our corpus include the High Court of Australia, the United Kingdom Supreme Court and the United States Supreme Court. The jurisdiction of apex courts, such as the United Kingdom Supreme Court, is generally reserved for cases of significant public importance and legal complexity. Accordingly, senior jurisdiction is conferred to courts lower down the hierarchical structure. For example, in England and Wales, the United Kingdom Supreme Court sits at apex of the court hierarchy. The Civil and Criminal Divisions of the Court of Appeal are subordinate to the Supreme Court, but are superior to the High Court; and the High Court is subordinate to the Court of Appeal, but is superior to the county court. Collectively, the United Kingdom Supreme Court, the Court of Appeal and the High Court are the *senior* or *higher* courts in the United Kingdom. Earlier work analyzing the use of foreign case law L'Heureux-Dube [1]; Lefler [8]; Tyrrell [30] indicates that the majority of cross-border interactions occur between courts situated at the top or near to the top of the relevant domestic judicial hierarchies. For this reason, we limit the scope of our study to the higher courts of the countries in our corpus.

This paper does not exhaustively cover all of the common law systems available in the vLex Justis collection; judgments of India, Kenya and Sri Lanka, for example, are not included in this study because data on inward and outward citations of cases in judgments emanating from those countries is currently unavailable. Moreover, for the same reason, our corpus does not include data on the number of South African judgments cited by other countries. However, we include South Africa in this study because the data does include foreign cases cited by South African courts. Additionally, our data for the United Kingdom chiefly consists of judgments of the courts of England and Wales, including decisions of the House of Lords and the Supreme Court that consider appeals originating in England and Wales. However, for convenience, we refer to the United Kingdom to

[9] 448 US 56 (1980).
[10] [2009] UKSC 14; [2010] 2 AC 373, HL(E).
[11] [2014] VSCA 276.

TABLE 1 | Composition of the cross-jurisdictional (XJ) network showing the courts included in our analysis. *Citing Cases* is the number of unique cases that have at least one outward edge. For example, cases from the United Kingdom cited to a total of 10,928 foreign cases. *Cited Cases* is the number of unique cases with at least one inward edge. For example, 313,111 cases from the United Kingdom were cited by other countries in the data. The year of the earliest and latest citations are provided for both networks. For the United Kingdom, the year of the earliest case with either an inward or outward edge in the cross-jurisdictional network is 1767 (1713 in the complete network) and the latest is 2020 (also 2020 in the complete network).

Country	Citing cases	Cited cases	Year of earliest citation in XJ network	Year of latest citation in XJ network	Year of earliest citation in complete network	Year of latest citation in complete network	Superior courts included
Anguilla	834	15	1967	2019	1842	2019	High Court, Court of Appeal
Antigua and Barbuda	2,504	23	1959	2019	1808	2018	High Court, Court of Appeal
Australia	75,919	10,387	1903	2020	1,679	2020	Federal Court, High Court, Court of Appeal (Victoria), Supreme Courts (northern territories, New South Wales, Victoria)
Bahamas	11,686	43	1972	2019	1718	2019	Supreme Court, Court of Appeal
Barbados	7,616	326	1950	2019	1721	2017	High Court, Court of Appeal
Belize	3,802	111	1967	2019	1768	2019	High Court, Court of Appeal, Supreme Court
Bermuda	9,584	131	1957	2020	1772	2019	High Court, Court of Appeal, Supreme Court
British Virgin Islands	2,212	12	1967	2019	1828	2016	High Court, Court of Appeal
Canada	17,546	12,156	1938	2020	1722	2020	Supreme Court of Canada, federal Court, federal Court of Appeal, Supreme Court (British Columbia), Court of Appeal (Alberta, British Columbia, Manitoba, newfoundland, New Brunswick, Nova Scotia, nunavaut, Ontario, Yukon territory)
Cayman Islands	1,920	57	1972	2020	1813	2019	Court of Appeal
Grenada	1,675	11	1962	2019	1777	2017	High Court, Court of Appeal
Guyana	7,263	684	1946	2017	1,687	2016	High Court, Court of Appeal
Hong Kong	1,712	0	2019	2020	1838	2020	High Court, Court of Appeal, Court of Final Appeal
Ireland	30,730	2,229	1876	2020	1,682	2020	High Court, Court of Appeal, Court of Criminal Appeal, Supreme Court
Jamaica	27,617	872	1905	2019	1,694	2019	Court of Appeal, Supreme Court
Malaysia	43,295	2,277	1932	2020	1,687	2019	Supreme Court, High Court, federal Court, Court of Appeal
New Zealand	22,579	3,884	1964	2020	1702	2020	High Court, Court of Appeal, Supreme Court
Saint Kitts and Nevis	1,562	1	1967	2019	1774	2019	High Court, Court of Appeal
Saint Lucia	2,109	16	1956	2019	1809	2017	High Court, Court of Appeal
Saint Vincent and the Grenadines	2,136	4	1967	2020	1777	2018	High Court, Court of Appeal
Singapore	19,063	1,542	1967	2018	1725	2017	Court of three judges, High Court
South Africa	21,616	1	1910	2020	1763	2019	Supreme Court of Appeal and other courts[a]
Trinidad and Tobago	24,641	1,319	1948	2019	1707	2019	High Court, Court of Appeal
Turks and Caicos Islands	570	0	1999	2018	1875	2016	Supreme Court, Court of Appeal
United Kingdom	10,928	313,111	1767	2020	1713	2020	Supreme Court, house of lords, privy council, Court of Appeal (england and wales), Court of Appeal (northern Ireland), High Court (england and wales)
United States	6,008	7,910	1968	2020	1,697	2019	United States Supreme Court, State Supreme Courts, United States Courts of Appeals

[a] *Appellate Division, Cape Town - Bloemfontein, Appellate Division, Pietermaritzburg - Cape Town, Appellate Division, Pretoria, Appellate Division, Pretoria - Bloemfontein, Appellate Division, Privy Council, Bhisho High Court, Bophuthatswana Appellate Division, Bophuthatswana High Court, Bophuthatswana Supreme Court, Cape Provincial Division, Ciskei Appellate Division, Ciskei High Court, Ciskei Supreme Court, Constitutional Court, Constitutional Court, Zimbabwe, Durban and Coast Local Division, East London Circuit Local Division, Eastern Cape Division, Eastern Districts Local Division, Free State Division, Bloemfontein, Free State Division, Bloemnfontein, Gauteng Division, Pretoria, Gauteng Local Division, Johannesburg, Griqualand-West Local Division, Hooggeregshof van Venda, KwaZulu-Natal Division, Durban, KwaZulu-Natal Division, Pietermaritzburg, KwaZulu-Natal High Court, Durban, KwaZulu-Natal High Court, Pietermaritzburg, KwaZulu-Natal Local Division, Durban, Lesotho High Court, Maseru, Limpopo Division, Polokwane, Mpumalanga Division (Main Seat), Mpumalanga Division, Nelspruit, Natal Provincial Division, North Gauteng High Court, Pretoria, North West Division, Mahikeng, North West High Court, Mafikeng, North West High Court, Mahikeng, Northern Cape Division, Orange Free State Provincial Division, Privy Council, Rhodesia and Nyasaland Court of Appeal, South Eastern Cape Division, South Eastern Cape Local Division, South Gauteng High Court, Johannesburg, Southern African Development Community Tribunal, Supreme Court of Appeal, Supreme Court of Namibia, Transkei Appellate Division, Transkei Division, Transkei High Court, Transkei Supreme Court, Transvaal Provincial Division, Venda High Court, Venda Supreme Court, Western Cape Division, Cape Town, Western Cape High Court, Cape Town, Witwatersrand Local Division.*

describe this portion of the data. Finally, given its leading role in global affairs and the size of the jurisdiction generally, it would be reasonable to assume that the number of inward and outward citations for the United States would be larger than they appear in our data (6,008 outward citations and 7,910 inward citations). In common with the other countries in our data, we have selected courts that sit at the top of the United States court hierarchy. As discussed elsewhere in this paper, the low citation counts are

TABLE 2 | Properties of the complete and cross-jurisdictional networks.

	Complete network	Cross-jurisdiction network	Percentage
Nodes	1,711,626	169,131	9.88
Edges	12,656,156	355,598	2.81
Average clustering coefficient	0.019	0.006	
Density	4.32×10^{-6}	1.24×10^{-5}	
Transitivity	0.06	0.003	

consistent with earlier work which has found that foreign case law utilization is rare in the United States [28].

We construct the complete directed case-to-case network of domestic and cross-jurisdictional citations in which, as per **Figure 1A**, cases are represented as nodes and citations between them as edges. A directed edge extends from case *b* (the *citing* case) to case *a* (the *cited* case) if case *a* is referred to at least once in the judgment of case *b*. By construction, there are no cycles because a case is only capable of citing *earlier* decisions. The resulting network is a directed acyclic graph that evolves over time.

A majority of the complete network consists of relationships between cases where both the citing case and the cited case emanate from the *same* country (domestic citations), in additional to a smaller proportion of relationships where the citing case and the cited case emanate from *different* countries (foreign citations). Our analysis is principally concerned with relationships falling into the latter category. To construct the second network, the *cross-jurisdictional network*, we exclude all instances of domestic citation from the data (for example, where a United States case cites another United States case) so that the network only consists of cases that have interacted at least once with a case from a different country. A summary of the data in the cross-jurisdictional network is shown in **Table 1**.

3.1 Global Properties of the Networks

We begin by analyzing the global properties of the complete and cross-jurisdictional networks. A summary of the global properties of both networks is shown in **Table 2**.

The first feature of interest emerges from a comparison of the size of both networks. As mentioned above, we construct the cross-jurisdictional network of citation interactions between cases emanating from *different* countries by excluding all cases that do not have at least one interaction (either as a *citing* or a *cited* case) with a case emanating from a different country. We retain 9.88 percent of the nodes in the complete network and 2.81 percent of the edges. This yields the insight that a reasonably large number of the cases in our corpus have engaged, whether actively as a *citing* case or passively as a *cited* case, in the practice of citation of foreign case law.

We also compute the *density*, *average clustering coefficient* and *transitivity* of both networks. The average clustering coefficient is a global measure of the abundance of triangles present in the network. In the context of a case law citation network, a triangle exists where case *b* cites case *a* and both cases *a* and *b* are cited by case *c* (see **Figure 1C**). Given that instances of domestic citation (citations between two cases that emanate from the *same* country) have been removed from the cross-jurisdictional network, for both networks we compute the clustering coefficient for a case, v, by dividing the number of edges between v's neighbours by the number of edges between v's neighbors that *do not emanate* from the same country. In other words, if node j has q_j nearest neighbors with t_j connections between nodes in different jurisdictions, the local clustering coefficient C_j is

$$C_j(q_j) = \frac{t_j}{q_j(q_j - 1)/2} \quad (1)$$

and $\overline{C}$ is the average clustering coefficient of the network

$$\overline{C} = \frac{1}{n}\sum_{i=1}^{n} C_i \quad (2)$$

Transitivity measures the fraction of *possible triangles* by identifying the number of *triads* in the network. A triad of nodes in our networks exists where case *a* is cited by cases *b* and *c*, but no edge exists between cases *b* and *c*.

In common with other real-world complex networks, such as social networks [31]; ecological systems [32]; and patent citation networks [33], the complete and cross-jurisdictional networks present low density (4.32×10^{-6} and 1.24×10^{-5}, respectively). In the complete network, the coefficients of average clustering and transitivity are four orders of magnitude greater than the global density. In the cross-jurisdictional network, the average clustering coefficient and transitivity are two orders of magnitude greater than that network's density. Both networks present global properties that are similar to those found in a similar study that focused on the citation network in a corpus of judgments from the International Criminal Court [25].

The complete network presents clustering behavior that is an order of magnitude higher than its cross-jurisdictional counterpart (see **Table 2**). This is expected, because the cross-jurisdictional network was constructed by pruning all nodes from the complete network that did not have at least one interaction (as the citing or cited case) with a foreign case. The comparatively low degree of clustering in the cross-jurisdictional network provides an indication that instances in which three or more cases become linked through citation are rare. We analyze this further in Section 3.2.

3.2 Degree Distributions

The degree distribution of a network is a fundamental quantity measured in most analyses of complex networks. In this section we analyze the distribution of inward citations (citations *to* a case)

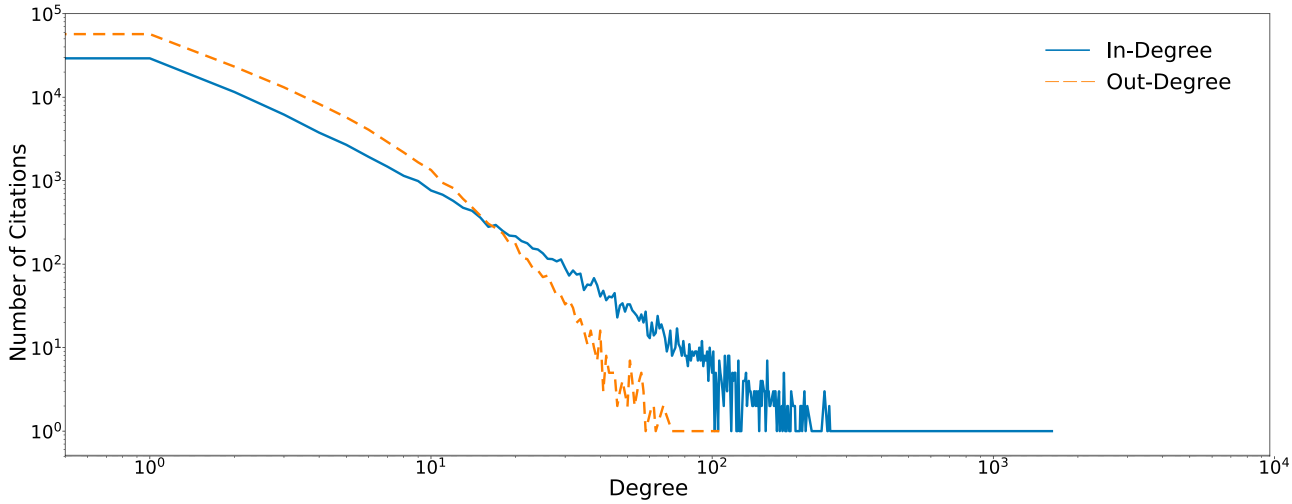


FIGURE 2 | Log-log distributions of inward and outward citations in the cross-jurisdictional network.

and outward citations (citations *from* a case) in the cross-jurisdictional network. The inward and outward degree distributions examined in earlier studies of case law citation networks constructed from the decisions of the US Supreme Court, European Court of Human Rights and the International Criminal Court have each exhibited two common characteristics. First, their inward and outward degree distributions are heterogeneous. Second, both distributions appear to follow a pattern whereby most decisions in the network are cited by relatively few cases, whereas a small number of decisions are cited by many cases. The same pattern applies to the inverse scenario, where most decisions in the network cite relatively few other cases, whereas a minority of decisions cite many cases. Similar patterns have been widely observed in the degree distributions of other large networks, including scientific paper citation networks [34], patent citation networks [35, 36], the structure of the World-Wide Web [37] and social networks [38]. It has been argued that these distributions are the consequence of a process of "preferential attachment" [39, 40] which would indicate, in the context of a judicial citation network, that the more a case has been cited by past cases, the greater the likelihood that it will be cited by future cases. The inward and outward degree distributions in our cross-jurisdictional network, as can be seen in the log-log plots in **Figure 2**, share these properties.

There is evidence of the process of preferential attachment in the cross-jurisdictional network. We observe that 611 cases ranked in the 99th and 100th percentiles by inward citation count constitute a quarter of the total inward citations in the entire cross-jurisdictional network. Examining this small population of 611 cases with high indegrees, which we will refer to as *super authorities* [41, 42], provides initial insights into the dominance of case law emanating from specific countries in the cross-jurisdictional network. We find that the majority (80.4%) of citations in the cross-jurisdictional network were to United Kingdom super authorities, while the second largest group of most cited cases were Canadian super authorities, with 17.3% of the share of inward citations. The dominance of super authorities from the United Kingdom (chiefly, England and Wales) and Canada persists when we limit the pool of cited cases to those decided in or after 2000. However, the proportion of United Kingdom super authorities declines to 67.2%, while the proportion of Canadian cases in the top two percentiles of inward citation count rises to 23.2%.

To examine the extent to which the cases in this group of super authorities possess *landmark* [12] qualities, we calculated the proportion of the top ranking 100 cases emanating from courts in the United Kingdom (which are mainly decisions from the English and Welsh jurisdiction) by indegree that had been reported in England and Wales' leading series of law reports, *The Law Reports*, published by the Incorporated Council of Law Reporting for England and Wales since 1865. For a case to be selected for inclusion in *The Law Reports* it must exhibit, in the view of that series' editors, the potential to have longstanding significance as a precedent [13]. In general, cases selected for inclusion in *The Law Reports* are published in that series within a year to eighteen months from the date of judgment.

We found that 86 of the top ranking 100 United Kingdom cases by inward citations had been reported in that series of law reports and that 66 of these cases were decisions of the United Kingdom *apex* courts: the Supreme Court, the House of Lords and the Judicial Committee of the Privy Council. This suggests that there is a correlation between the indegree of a case and the status of that case as a landmark authority. This

[12]Landmark cases are cases that have a long-term effect on the state of the law.
[13]*The Law Reports* are regarded as the most authoritative series of law reports in England and Wales. See *Practice Direction* (*Citation of Authorities* [2012] 1 WLR 780 at [6]. This Practice Direction is available online at https://www.judiciary.uk/wp-content/uploads/JCO/Documents/Practice+Directions/lcj-pract-dir-citation-authorities-2012.pdf (accessed 21 January 2021).

indication is supported by an inspection of the top three ranking United Kingdom cases, all of which are generally regarded as seminal decisions:

- A*merican Cyanamid Company v Ethicon Ltd.*, [14] the leading authority on applications for interim relief. 1,477 inward citations.
- A*ssociated Provincial Picture Houses Ltd. v Wednesbury Corporation*, [15] a leading case in the sphere of judicial review that established the test of unreasonableness in public body decision making. 860 inward citations.
- *Donoghue v Stevenson*, [16] established the foundations of the tort of negligence. 719 inward citations.

While using *The Law Reports* as a benchmark provides a useful guide for ascertaining the correspondence between indegree and the status of a case as a landmark authority, there are limitations to this approach. Inclusion of a case in *The Law Reports* amounts to a prediction on the part of that series' editors as to the likelihood that the case in question bears the marker of a landmark authority. Publication in that series of law reports therefore has the potential to enhance the visibility and the perceived impact of a given reported case to users of case law, thereby increasing the likelihood that it will be discovered and cited in future cases when compared to a case that was *not* reported in that series. One way to control for this effect, which we leave for future work, may be to examine the rate of citation to a case for a period prior to (generally, a year to eighteen months) and following its publication in *The Law Reports*. Moreover, it has been recognized that citation counts are biased by the age of the cited unit. For example, in the context of academic paper citation networks, the number of citations received by a paper depends on the age of the paper [43]. Older papers have more time to acquire citations than more recent papers–an advantage that is enhanced by the phenomenon of preferential attachment [44, 45]. Both the complete and cross-jurisdictional networks are subject to this bias.

3.3 Average Clustering as a Function of Degree

In their analysis of the citation network constructed from a corpus of International Criminal Court decisions Tarissan and Nollez-Goldbach [25] observed a pattern under which the local clustering coefficient of a decision was inversely proportional to its indegree: decisions with high indegree presented low clustering, whereas decisions with low indegree presented higher local clustering. The authors of that study explain this pattern by noting that small indegree cases in their network tended to deal with esoteric issues pertaining to the court's procedure that raised specific and technical points of law, while larger degree cases addressed substantive issues of broader application. A similar pattern has been observed in other growing directed networks, such as scientific paper author collaboration networks [46].

Translating this trend into the context of our cross-jurisdictional network, it would suggest that large degree cases establish general principles that are applicable to a wide range of factually disparate disputes: case *a* establishes principles or rules of general application that are relevant to the issues to be determined in cases *b* and *c*, but the factual and legal matrices of cases *b* and *c* do not overlap sufficiently for either of those cases to cite the other. The inverse scenario in which low degree cases present higher clustering would suggest that low degree cases have a tendency to address specific factual and legal issues that are relevant to small cliques of cases entering the network.

We compute the average local clustering coefficient (using the approach outlined in Section 3.1) as a function of indegree to explore whether a similar relationship exists in our networks and compare our results to random networks. For both networks, we generate a random network using the degree distribution of the respective real network as the configuration model, removing all self-looping and parallel edges from the generated network. We then randomly assign a country attribute to each node in the generated random network, following the distribution of countries in the respective real network (i.e., the proportion of Canadian nodes in the random cross-jurisdictional network is equal to the proportion of Canadian nodes in the real cross-jurisdictional network). Finally, all edges between nodes from the same country were removed from the random version of the cross-jurisdictional network. Our results are shown in **Figure 3**.

There is some evidence of a correlation between indegree and local clustering in the cross-jurisdictional network (**Figure 3B**), although it is weak. In the random cross-jurisdictional network, shown in orange, the clustering coefficient of cases with low indegree is small before increasing at higher indegrees. There is deviation between both networks at low degree, with the cross-jurisdictional network presenting higher clustering for cases with low inward citations. However, the deviation between the real and the random network is less distinct at high degree. High degree nodes in both networks present low clustering, although the coefficient is larger in the cross-jurisdictional network than in the random cross-jurisdictional. As can be seen, the majority of the cases in the cross-jurisdictional network have high degree. It is therefore possible, even when compared with the random network, that the presence of higher clustering at low degree in the cross-jurisdiction is the product of chance rather than the phenomenon observed by Tarissan and Nollez-Goldbach [25].

The opposite pattern is observed in the complete network (**Figure 3A**), which demonstrates a trend of increased clustering for cases with higher inward citations compared to cases with fewer citations. This trend is also reflected in the random complete network. However, clustering at high indegree in the complete network exhibits far more variability than that observed in the random complete network. This variability may indicate the presence of cases in the complete network that act as frequently cited hubs of jurisprudence on legal issues that are common or prominent across a range of countries in our network. Deeper analysis, which we reserve for future work,

[14][1975] AC 396.
[15][1948] 1 KB 223.
[16][1932] AC 562.

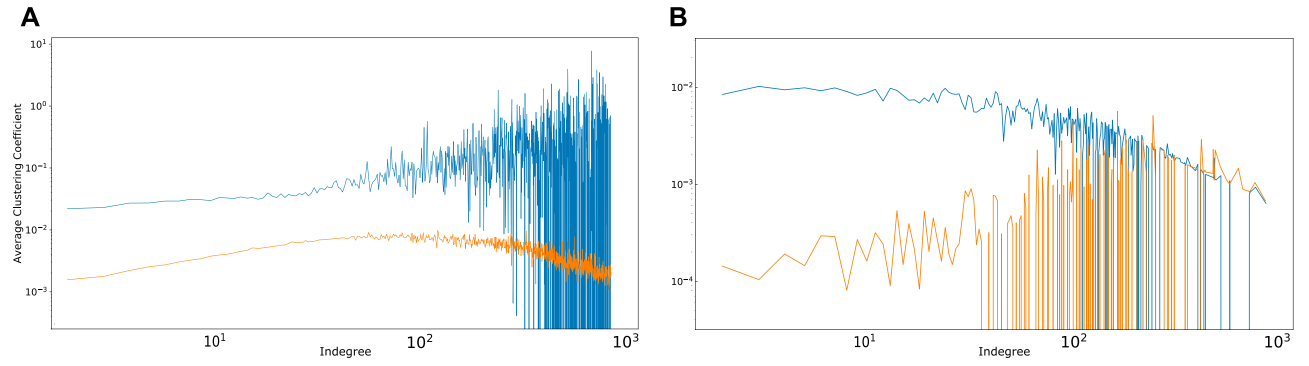


FIGURE 3 | Local average clustering coefficient of cases as a function of indegree. **(A)** The complete network **(B)** The cross-jurisdiction network. Cases with high degrees present lower clustering coefficients. The complete and cross-jurisdictional networks are shown in blue, their randomized counterparts are shown in orange.

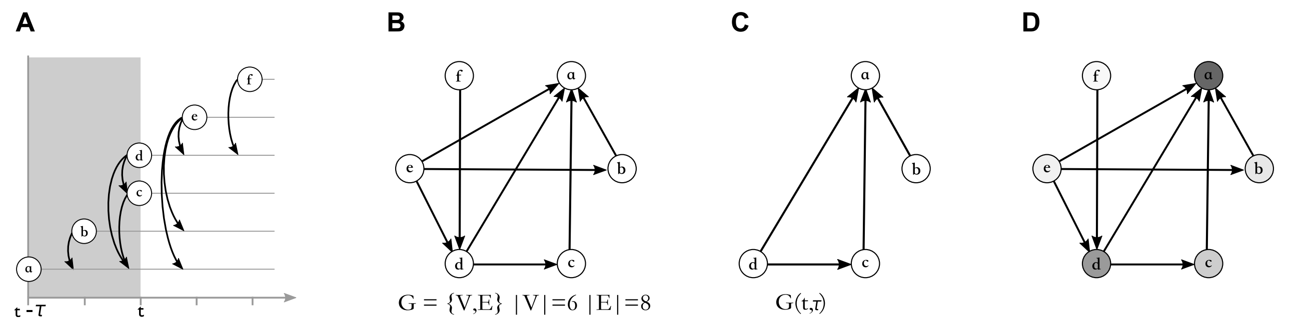


FIGURE 4 | Temporal graph. **(A)** Windowing procedure to construct the graph $G(t, \tau)$ **(B)** Complete graph G with 6 nodes and 8 edges **(C)** Sub-graph $G(t, \tau) \subset G$ **(D)** Authority score. Example where the shade of each node represents its Authority score. Node *a* has high authority.

could be directed to confirm and explore the presence of these hub cases and the role they may play in diffusing key common law concepts across nations.

3.4 Evolution of the Cross-Jurisdictional Network Over Time

Our focus now turns to exploring how the cross-jurisdictional network has evolved over time. To enable this we follow the temporal window approach of Steer et al. [47]. This approach allows us to view the exact state of the networks at any point throughout its lifetime and move through this history at different temporal levels of granularity. By filtering out older entities shorter term patterns may be extracted which would otherwise be obscured by the full data aggregate.

The temporal window approach is illustrated in **Figure 4**. In **Figure 4A** the evolution of a network is plotted over time, showing new vertices joining the network and which existing nodes they are connected to. This can be envisioned within our cross-jurisdictional network context where new judgments are published, citing previously established judgments and, therefore, generating edges. Aggregating all of these vertices and edges together will create the latest version of the graph *G*, seen in the middle of the figure, which is what typical graph analysis will be performed on. Lastly, in **Figure 4C**, we can see a windowed view of the graph $G(t, \tau)$. This consists of the graph as it would have existed at time *t* and with a set window size of τ.

3.4.1 Growth of the Cross-Jurisdictional Network

We begin our temporal analysis of the cross-jurisdictional network by investigating the growth of the network, based on the number of nodes and edges present, between 1940 and 2020. This temporal period is chosen because as **Table 1** shows, the majority of citation interactions in the corpus selected for this study begin in the mid-twentieth century. The results of this analysis are shown in **Figure 5**.

The growth of the network was modest in the fifty years between 1940 and 1990, with an approximate increase of 28,000 cases and 60,000 edges present in the network over that period. The size of the network accelerates from 1990 onward, with a three-fold increase from 60,000 cases in 1990 to 181,618 cases in 2020. The increase in edges in the network is even more striking, rising from approximately 70,000 edges in 1990 to 384,336 edges in 2020. The growth in cases entering the network from 1990 may be attributable to the rise in the use of digital systems to author, store and disseminate case law [1, 48, 49] and the fact that as time

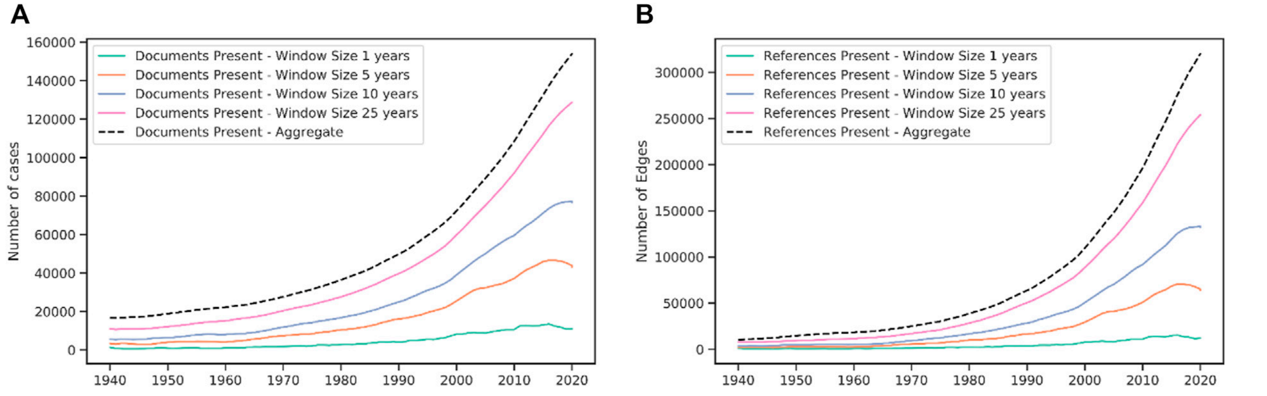


FIGURE 5 | Evolution of the cross-jurisdictional network. **(A)** Nodes and **(B)** edges present over time and four window widths, $\tau = \{1, 5, 10, 25\}$.

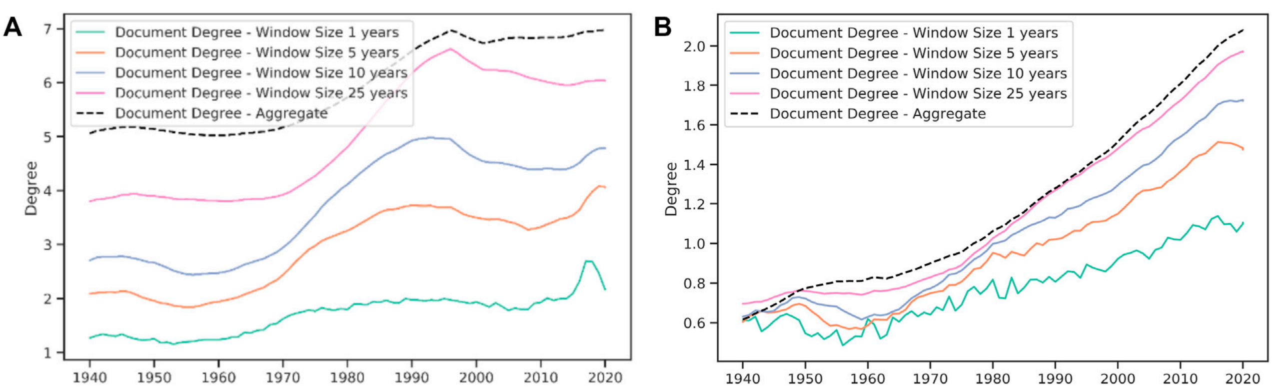


FIGURE 6 | Distribution of case average degree. **(A)** The complete network **(B)** The cross-jurisdiction network. Window widths.$\tau = \{1, 5, 10, 25\}$.

passes the stock of precedent increases–there are more earlier cases available to be cited.

3.4.2 Average Case Degree

The increase in the number edges in the network raises the question as to whether judges have increased the number of cases cited in their judgments. To explore this, we analyzed the mean degree of the cases in the complete network and the cross-jurisdictional network between 1940 and 2020. The results of this analysis are shown in **Figure 6**.

Both networks present an increase in mean degree over time. The mean degree of cases in the complete network, shown in **Figure 6A** rises from five in 1940 to 7 in 2020, with a noticeable increase occurring between 1970 and 2000, at which point the mean degree plateaus. The mean degree of cases in the cross-jurisdictional network, shown in **Figure 6B** are lower, rising from a mean of 0.6 in 1940 to 2 in 2020. This indicates that judges have progressively increased the amount of case law cited in their judgments over time, both in the complete network, which includes domestic citations, and in the cross-jurisdictional network.

Our conclusion that judges have increased the number of cases they cite in their judgments over time leads to the possibility that the network of foreign citations has grown progressively more connected over the same period. This is supported by a temporal analysis of the growth of the largest connected component in the cross-jurisdictional network, the results of which are shown in **Figure 7**. In the context of a case law citation network, a component is a set of cases for which each pair of cases are linked by at least one path through the network (see **Figure 7A**).

The trend presented by our analysis of the growth of the largest connected component in the cross-jurisdiction network in **Figure 7B** shows that the size of that component has more than doubled from where it stood in 1940 at approximately 45% of the network to approximately 95% of the network in 2020. The presence of a single giant component in the network is consistent with the findings of earlier analyses of case law networks [18, 22]. This may be an indication that the body of foreign persuasive authority cited by the countries in the cross-jurisdictional network has achieved a some degree of integration over the passage of time. When citing foreign cases, domestic judges are sampling from, and contributing to, a progressively more seamless web of case law [16] as opposed to fragmented isolated clusters of authority. However, we also observe in **Figure 7B** that the size of the largest component in the cross-jurisdictional network temporarily decreased between 1955 and 1965 when τ is set

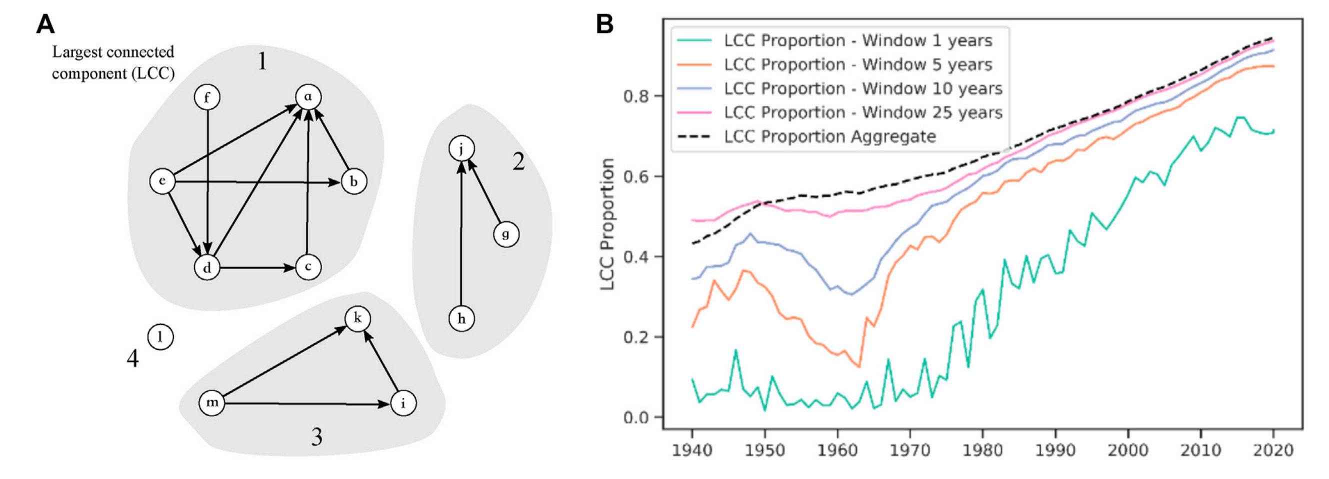


FIGURE 7 | Connected components. **(A)** Connected components on for a graph $G = \{V, E\}$ at time t illustrating the largest connected component (LCC). **(B)** Temporal connected components with five window widths showing the change in the proportion of the LLC in the cross-jurisdictional network.

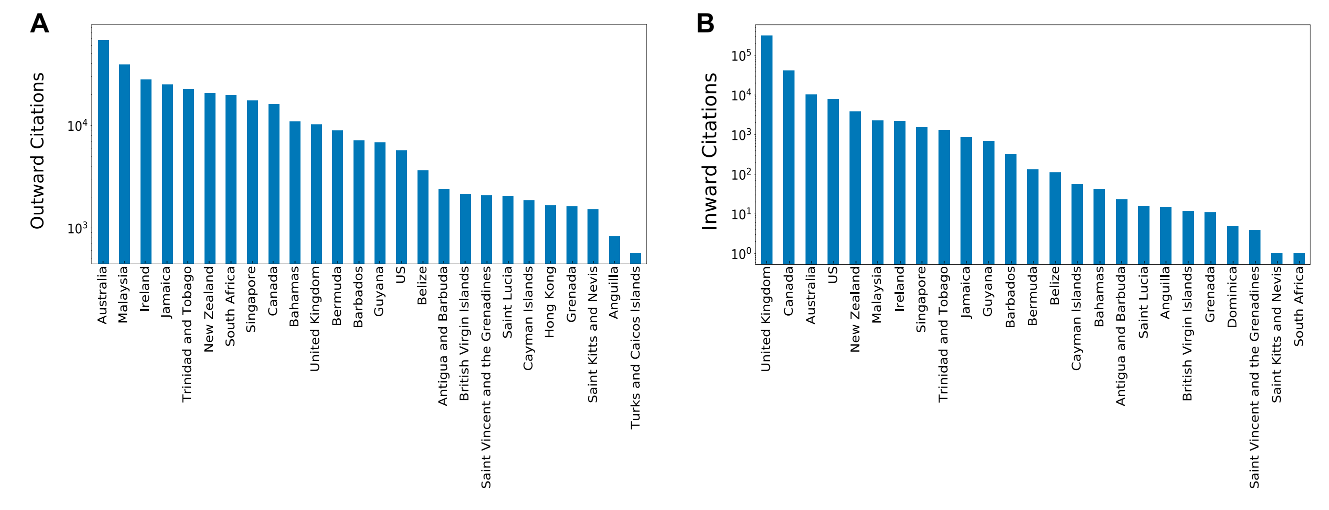


FIGURE 8 | Citation distributions in the cross-jurisdictional network. **(A)** Outward citations for each country **(B)** Inward citations per country. Australia is shown to be the most prolific user of foreign case law by outward citations. Judgments of courts in the United Kingdom, chiefly courts of England and Wales, are the most cited source of case law.

to a window size of 5 and 10 years. This reduction in the size of largest component is likely due to the fact that the majority of non-United Kingdom judgments in our corpus entered the network during this period, which led to the temporary creation of smaller components that were subsequently assimilated into the largest component once they were cited by later cases.

4 A GLOBAL COMMUNITY OF COURTS

4.1 Overview of Foreign Citations by Country

In the last section we presented our analysis of the fundamental global properties of the cross-jurisdictional network, along with an investigation of how the network has evolved over time. Our focus in this section turns to address the extent to which the cross-jurisdictional network reveals evidence of the emergence of a global community of courts formed by the cross-border flow of persuasive authority. The legal scholarship in this area presents a mixed picture in which the extent to which specific domestic legal systems are willing to engage with foreign case law varies considerably. Lefler's [8] analysis of the judgments of the US Supreme Court, the Canadian Supreme Court and the High Court of Australia found that judges in Australia and Canada promote the use of foreign case law, particularly in criminal cases. The United Kingdom Supreme Court also makes regular reference to jurisprudence of other countries [30]. In contrast, earlier work indicates that American judges make relatively little use of foreign case law [8, 28]. **Figure 8** show the log-scale quantities of foreign cases *cited by* and *citing to* each country in the network. Three points of interest emerge. The first is the dominance of cases emanating from the United Kingdom as the most cited legal system in the network. This is likely due to historical factors which are discussed further in Section 4.2. The second concerns Australia, which stands out as the most prolific user, by volume, of foreign authority–this trend is consistent Lefler's [8] findings. The third relates to the United States. Notwithstanding that the use of foreign case law by United States courts has proven to be a contentious issue in the discourse of American legal scholars [42, 50], our data suggests that in contrast to comparable jurisdictions–such as the United Kingdom, Canada and Australia–the instances of United States engagement with foreign case law are rare. This is consistent with earlier work which found that American courts make scarce use of foreign case law [28]. Despite the fact that United States judgments appear to be relatively isolated from the influence of foreign jurisprudence, they form the fourth most consulted body of jurisprudence in the cross-jurisdictional network, after the United Kingdom, Canada and Australia.

4.2 Considerations Relevant to the Citation of Foreign Case Law

Our analysis shows that although the courts of all countries in our corpus engage in the citation of foreign case law, the extent to which they do so varies. In all cases the volume of citations to foreign case law are dwarfed by the volume of citations to domestic law. The variability in interactions with foreign case law across the countries in the cross-jurisdictional network suggests the existence of deeper considerations on the part of judges when reaching beyond their domestic law to persuasive judgments given by the courts of other countries.

We begin from the basic legal and constitutional imperative that in order to maintain their institutional and decision legitimacy, national courts will generally seek to resolve the legal issues before them in accordance with the content of their own *binding* domestic law. From this imperative, we assume that where the resolution of legal issues presented by a case appears to require, and the doctrine of precedent permits, the use of authority by which the court is not bound (*persuasive authority*), the court will prefer domestic persuasive authority over foreign authority. Accordingly, we proceed on the basis that national courts will generally refer to foreign case in limited circumstances as a matter of last resort where 1) the content of domestic law (binding or otherwise) is inconclusive as to the issues presented; 2) the content of domestic law *is* conclusive, but leads to an "unjust" [17] outcome and the court has the latitude to depart from established principle; or 3) the issues of the case squarely present an international dimension [28].

In this section, we test three assumptions to explore the extent to which the selection of foreign case law is guided by the citing court's perception of how influential and well-reasoned a foreign case is and the country from which it emanates.

Assumption 1: **"Importance" of foreign cases**: Our first assumption relates to the perceived "importance" of foreign cases. On the most straightforward view, an "important" case is a case that has subsequently been cited by many important later cases. In view of the evidence that citation of foreign case law is a minority occurrence that arguably deviates from the strict goal of settling domestic legal questions in accordance with domestic law, we assume it is more likely than not that courts will have historically tended to confine themselves to the citation of foreign cases that exhibit high importance and influence in the jurisdictions from which they emanate. However, we assume that the emphasis courts place on the perceived influence of a foreign case as a prerequisite for citation will have declined over time as the practice of foreign citation has grown more commonplace.We use the *authority* score computed by the hyperlink-induced topic search algorithm (HITS) [51] as a proxy for the importance of a case: important cases are those with high authority scores and less important cases are those with low authority scores. We compute HITS over successive temporal partitions of the complete network to analyze how the authority scores of cases cited in the cross-jurisdictional network between 1940 and 2020 change over time. A reduction in the mean authority score of cases cited over time may provide an indication that judges have grown less concerned about limiting themselves to citing foreign cases that are regarded as having particular importance or weight of authority.

Assumption 2: **"Grounding" of foreign cases**: Our second assumption examines the extent to which cited foreign cases are "well-grounded" in their own domestic jurisprudence. On the most straightforward view, a "well-grounded" case is a case that it itself cited many earlier important cases. In common with **Assumption 1**, we assume that judges will have historically preferred to cite foreign cases that were well-grounded in their own domestic case law when they were decided. However, we assume that as the practice of foreign citation has grown more commonplace over time, the emphasis a judge may place on how well-grounded a case is will have reduced.We use the *hub* score computed by HITS as a proxy for how well-grounded a case is: well-grounded cases are those with high hub scores and less well-grounded are those with low hub scores. We compute HITS over successive temporal partitions of the network to analyze how the hub scores of cases cited in the cross-jurisdictional network between 1940 and 2020 change over time. A reduction in the mean hub score of cases cited over time may provide an indication that judges have grown less concerned about limiting themselves to citing foreign cases that are regarded as having been well-grounded in their own domestic law.

Assumption 3: **Geographic proximity**: Homophily, the principle that similarity breeds connection, has been found to influence a range of network settings, including paper citation networks [52] and social networks [53]. A recent study examining an Australian case law citation network observed that judges in that country generally favoured the jurisprudence of countries that reflect their social values and legal traditions [29]. There are any number of points of similarity between national legal systems,

[17]This, for example, was the motivation for recourse to foreign case law in the United Kingdom House of Lords decision in *Fairchild* [2003] 1 AC 32, HL(E).

including socioeconomic factors, language, style of constitution and government. A further point of similarity, which we explore, is geographic proximity. Our third assumption is that judges will tend to favour the citation of foreign case law that emanates from countries that are geographically proximal. We use the geographic distance between the countries in our data as a proxy for similarity, where countries separated by low distance are assumed to be more similar than countries separated by larger distance.

4.2.1 The Importance and Grounding of Cited Foreign Cases

In their analysis of over 30,000 majority opinions of the United States Supreme Court, Fowler and Jeon [17] explore the ways in which the citation network could be harnessed to identify cases that are the most important or influential for establishing precedent. At the most basic level, as we have shown in our analysis of the 100 most frequently cited United Kingdom cases in the cross-jurisdictional network, it is possible to rely on degree centrality–a classic and intuitive measure of importance [54]. However, Fowler and Jeon's study argues that degree centrality fails to fully utilize the information in the case law citation network because all inward citations are treated equally and that, when estimating the importance of a case, we should ideally be able to account for the cases that those cases themselves cite. This is made clearer with an example. Suppose case *a* is cited by a case of considerable importance, case *x*, and that case *b* is cited by a case of low importance, case *y*. This would indicate that case *a* is more important than case *b*, because case *x* is more important than case *y*.

An alternative strategy to the calculation of importance considered by that paper is eigenvector centrality, which computes the importance of nodes in a network based on the centrality of its neighboring nodes [55]. Eigenvector centrality was also discounted by Fowler and Jeon as an appropriate measure of importance for case law because that approach only treats nodes associated by an inward edge as neighbors for the purposes of the calculation. Fowler and Jeon regarded this as problematic because the importance of a case is simultaneously dependent on the importance of the cases *citing it* and the importance of the cases that *it itself cites*. For example, if case *a* is cited by case x, which cited many important cases, and case *b* is cited by case y, which mainly cited cases of low importance, then case *a* may be said to be more important than case *b*, because case *x* is well-grounded in important cases and case *y* is not.

The approach adopted by Fowler and Jeon to assess importance, which we follow in this paper, is the hyperlink-induced topic search algorithm (HITS) [51]. HITS uses two related but distinct scores of importance: the authority score and the hub score. In the context of a case law citation network, an *authority* is a case that is extensively *cited by* other cases in the network and a *hub* is a case that itself *cites* many other cases in the network. The relationship between authorities and hubs is mutually reinforcing–a good authority is generally a case that cites many good hubs; a good hub is a case that cites many good authorities. The authority score for a given case depends on 1) the number of cases it has been cited by; and 2) the hub scores of the cases it cites. This is illustrated in **Figure 4D**. Meanwhile, the hub score for a given case depends on 1) the number of cases it cites; and 2) the authority scores of the cases it cites [18].

We follow Fowler and Jeon [17] and use the *authority* score as a proxy for the importance of cases in the network to test **Assumption 1**, which assumes that the emphasis courts place on the perceived importance of a foreign case when considering whether to cite it will have been historically greater than it is in present times. The framework underlying **Assumption 1** draws on two strands. First, the evidence in the literature and in the cases indicating the phenomenon of "cross-pollination" of judicial wisdom [1] suggests that judges are beginning to cast a wider net when it comes to the citation of foreign case law. There is an increase in cases cited per judgment over time in the cross-jurisdictional network in **Figure 6B**. The motivations driving wider citation of foreign jurisprudence, particularly in cases involving human rights, appear to stem from the recognition among an increasing community of judges that they are fellow professionals engaged in a common endeavor that transcends national borders [2] and improved accessibility of foreign case law online [1, 48, 49]. We therefore assume that the more widely judges are prepared to cite foreign authority, the less emphasis they will place on the "importance" of those authorities.

The second strand we draw on is the "strategic legitimation" model explored by Lupu and Voeten (2011) in their analysis of the European Court of Human Rights' (ECtHR) use of its own body of jurisprudence. In that study Lupu and Voeten found that the selection of authority by ECtHR judges is guided by a strategic concern to persuade the domestic parties, particularly national governments, to comply with the court's decisions by demonstrating impartial and careful decision-making [56]. Strategic legitimation was analyzed through the prism of the extent to which the ECtHR grounds its decisions in its own precedent, using the hub score calculated by HITS as the proxy for how "well-grounded" a given decision is in the court's body of case law. The overarching hypothesis running through that analysis, which the authors confirmed, was that the ECtHR seeks to promote its institutional legitimacy by taking care to ground its decisions, particularly those bearing on controversial issues, in a thorough survey of its own case law.

Returning to our core assumption that domestic courts are subject to the legal and constitutional imperative to decide the cases before them in accordance with their own domestic law, the model of strategic legitimation operates as a counterbalancing check on the freedom of judges to sample and incorporate foreign jurisprudence in their judgments. In the context of this cross-jurisdictional analysis, therefore, there is a tension between the advantages of participation in cross-border judicial dialogue and the countervailing concern not to undermine institutional and decisional legitimacy by over liberally citing foreign case law.

[18]A useful account of how HITS is implemented is provided in A. Langville and C. Meyer, "A survey of eigenvector methods of web information retrieval." SIAM Rev 47(1), 135–161: 137.

To examine how the emphasis placed by judges on the perceived importance of foreign cases changes over time, we compute the authority and hub scores for all of the cases in the *complete network* across nine temporal partitions. We use the complete network for this computation because the importance of a case is a product of *all* of its citation activity, including citations to and from other cases emanating from the *same country*. The first partition contains all cases and their relationships between 1767 (the year of the oldest case in the network) and 1940; the second partition from 1767 to 1950; and so on, where the terminal year is incremented by ten years in each partition until it reaches 2020, at which point authority scores are computed for all of the cases in the complete network. This temporal partition approach was deployed by Fowler and Jeon [17] in order to analyze the rise and fall of a case's authority score over time and allows us to overcome the fact that authority and hub scores would be frozen in a static network. To enable us to compare the evolution of the authority scores with a null network for each partition of the complete network, we generate a random equivalent using the degree distributions from the corresponding partition of the complete network and compute HITS on that random partition.

Given that the majority of the cases in the cross-jurisdictional network date from the mid-twentieth century, we start our analysis of the authority scores from 1940 stepping forward in time toward 2020 in ten year windows. For the first window, we induce the sub-graph of all cases cited between 1940 and 1950 from the cross-jurisdictional network and calculate the mean of the authority scores for that set of cases from the first partition of the complete network (1767–1950). For the next step, we slide the window forward in time to 1950 to 1960 and repeat the procedure against the second partition (1767–1960) of the authority score computed in the complete network, and so on. We repeat this procedure on the complete network to enable a comparison between purely cross-jurisdictional citation interactions one the one hand, and a complete representation of all citation activity, including domestic citation activity, on the other. We replicate this procedure on the random networks.

The evolution of authority scores of cases cited in the complete and cross-jurisdictional networks between 1940 and 2020, along with the authority scores computed in the random networks, are shown in **Figures 9A,B**, respectively.

Our temporal analysis of authority scores in the complete network (**Figure 9A**), shows that the authority scores of cases cited steadily decreases over time. The same trend can also be observed in the evolution of the authority scores in the randomized version of the complete network, shown in orange, however the average score in the random network is consistently lower across the examined time period. The trend in the authority scores of cases cited in the cross-jurisdictional network (**Figure 9B**) is less uniform. We observe, in contrast to the complete network, that the authority scores of cited cases start off low between 1940 and 1970, before sharply increasing between 1970 and 1990, where they peak. This peak in scores is likely due to the fact that in our data most of the citation activity starts in the mid-twentieth century. The peak in scores over this period indicates that extensive citation was being made to a select group of cases of heavily cited cases. From 1990, the average authority score proceeds to steadily decline over the following 30 years toward 2020.

The analysis of the authority scores in the cross-jurisdictional network and the complete network both demonstrate a steady decline in the authority score of cases cited over time, although the points at which the decline commences differ. This is consistent with **Assumption 1**, which proposes that the preference of judges to cite high importance cases will have reduced over time as the practice of foreign citation grows more widespread. However, the fact that the trend of decreasing authority scores is mirrored in the random networks provides an indication that the fall in importance, as measured by the authority score, may in fact be a consequence of the growth in the size of the network rather than a conscious decision on the part of judges to cite lower impact cases, as envisaged by **Assumption 1**.

The decline in authority scores in the cross-jurisdictional network begins in 1990. From this point in time we observe two patterns that are common to both networks: a steep increase in the number of cases entering the networks (**Figure 5A**; Section 3.4) and an increase in the number of cases cited per case (**Figure 5B**). The period from 1990 onward is marked by increased digitization both of the judgments themselves and the platforms used to facilitate their dissemination and retrieval. Technological advances in the online legal research domain in particular have significantly increased the ease with which judges and lawyers are able to access foreign jurisprudence and engage in comparative analysis [48,49]. Prior to the widespread digitization of legal sources, discovery of case law,

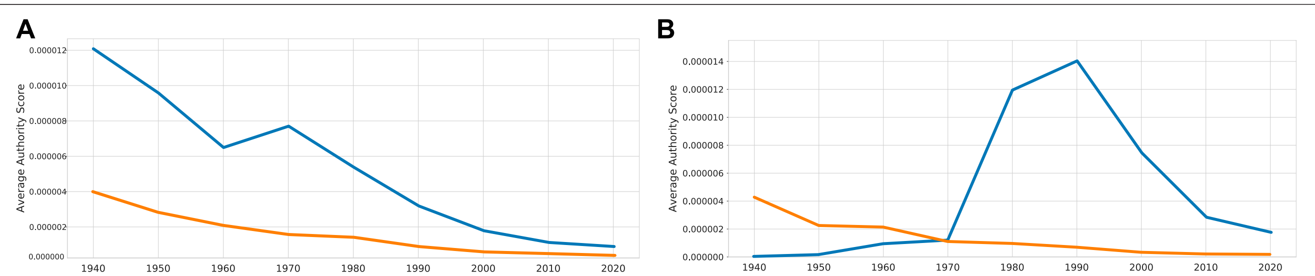


FIGURE 9 | Authority scores of **(A)** The complete network and **(B)** the cross-jurisdictional network between 1940 and 2020. For the real networks the authority scores are shown in blue, for the random networks the scores are shown in orange.

both foreign and domestic, largely depended on access to printed collections of law reports or to textbooks, both of which tended to be limited to the treatment of well-known cases. The ability to access foreign case law in volume online is likely to have resulted in a broadening of the horizons of judges, enabling them to simultaneously cast their nets *wider* across a growing number of bodies of foreign jurisprudence and *deeper* into case law pertaining to esoteric issues.

Accordingly, it is difficult to decouple the emphasis placed by judges on the importance of a foreign case when considering whether to cite it from the effects of dramatically improved accessibility of foreign case law over time and its apparent effect on the growth of the networks. However, the overall trend is that over time, and certainly since the advent of digital access to case law, citations to foreign authority have increased and the importance of the cases that are being cited, as measured by their authority scores, has decreased. This, in turn, indicates an increased willingness among judges to make use of foreign case law and a reduced tendency to confine reference to landmark cases.

Assumption 2 proposes that the preference of judges to cite cases that are well-grounded in existing case law will have reduced over time as the practice of foreign citation grows more widespread. We use the hub score computed by HITS as a proxy for how well-grounded a case is in earlier authority.

The trends in the evolution of the average hub scores closely resemble those of the authority scores. In the complete network, shown in **Figure 10A** the hub scores start high and steadily decline over time. The same steady decline is reflected in the movement of the hub score in the random network, shown in orange. The evolution of the hub score in the cross-jurisdictional network and the random model generated from its degree distribution, shown in **Figure 10B**, is less straightforward. The hub scores in the cross-jurisdictional model are generally low throughout the observed period. The scores are seen to peak in 1990 and then to continually decline from that point onward. As with our analysis of the evolution of the authority scores, this trend is consistent with **Assumption 2**. However, the movement of the hub score in the random cross-jurisdictional network provides strong evidence that the decline in hub scores from 1990 onward is a consequence of the increasing size of the network rather than a conscious change in approach on the part of judges.

As with our analysis of the evolution of the authority scores, our analysis of the hub scores do not provide sufficient evidence to enable us to state that the reduction in hub scores is a *direct consequence* of judges lowering their attention to how well-grounded a particular foreign case is in the domestic authority of the country from which it emanates. Rather, the rise of widespread digital access to case law from 1990 onward is likely to be playing the prominent role in this respect.

Our analysis of the authority score raises a general question as to its effectiveness of as a measure of importance in large evolving case law citation networks. In their study, Fowler and Jeon [17] compared the top ten ranking cases by authority score in their network with three sets of expert rankings of the most influential United States Supreme Court decisions and found that all but one of those cases appeared in at least two of the three sets of rankings. Their analysis therefore revealed a strong association between the authority score and importance. In order to evaluate the validity of the authority when applied to the data in our network, we compare the relationship between case indegree and authority score (the independent variables) with whether a case has been reported in the leading series of law reports for England and Wales, *The Law Reports*. We use the inclusion of the case in *The Law Reports* as a proxy for importance and this serves as our dependant variable. For this comparison, we induce the sub-graph of citations internal to the United Kingdom from the complete network and compute the authority scores in that network. Then for each measure (authority score and indegree), we sample to the top 100 cases, the bottom 60 cases and 60 additional cases around the middle yielding 280 cases for each measure. We then manually check whether the cases for each measure were reported in *The Law Reports*. The results of a logistic regression, along with Pearson correlation scores, on the relationship between these measures and importance, shown in **Table 3**, indicate that there is a significant relationship between indegree and importance using *The Law Reports* as a benchmark. The authority score, on the other hand, has a weaker association with importance in our data.

4.2.2 The Role of Geographic Proximity

In **Assumption 3**, we draw on Spottiswood's [29] study of the High Court of Australia and assume that in making choices as to which foreign country to cite from, judges will generally favor the jurisprudence of countries that share similar values. We use geographic proximity as the proxy for similarity and assume that countries that are close in distance are more likely to share values than countries that are separated by larger distances.

United Kingdom We plot the matrix of citation interactions between the countries in the cross-jurisdictional network in **Figure 11**. Virtually all countries, with the exception of the United States, direct the bulk of their outward citations toward the United Kingdom. This is likely due to historic factors: all of the countries in the network were at some point subject to British imperial rule and, as a consequence, have systems of law modeled on that used in England. Moreover, all countries, with the exception of the United States and the United Kingdom itself, at some point had the Judicial Committee of the Privy Council as their final court of appeal (this remains the case for twelve of the countries in the network).

North America 58% of United States outward citations were to Canadian cases. In this respect, our findings are broadly consistent with those of Zaring [28]. 64% of outward Canadian citations were to decisions emanating from courts in the United Kingdom. This is not surprising given Canada's close historic ties with the United Kingdom. However, the second country most cited by Canada is the United States, the decisions of which amount to 31.7% of Canada's outward citations.

Asia-Pacific When references to decisions of courts in the United Kingdom are removed, Malaysian courts are seen to prefer the case law of Singapore (33.5% of Malaysian outward citations) and Australia (48.2%), and to a lesser extent, New Zealand (7.2%). New Zealand directs most of its non-United Kingdom outward

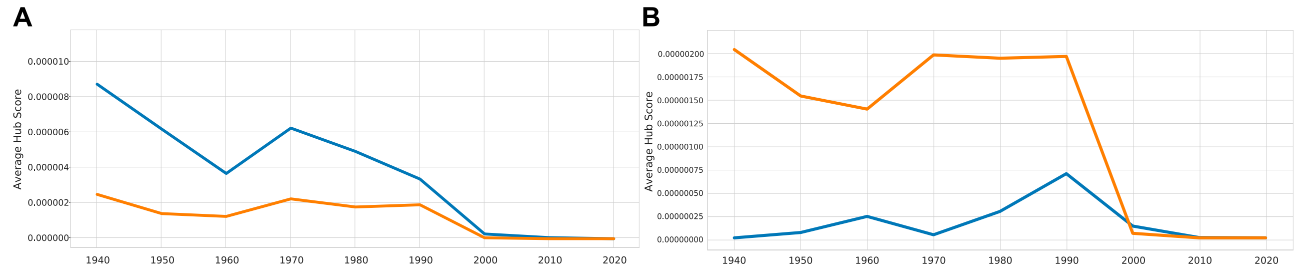


FIGURE 10 | Hub scores of cited cases in **(A)** the complete network and **(B)** the cross-jurisdictional network between 1940 and 2020. For the real networks the hub scores are shown in blue, for the random networks the scores are shown in orange.

TABLE 3 | Relationship between importance and indegree/authority in United Kingdom cases. *N = 280*. Coefficients and standard errors calculated using logit. The dependant variable is whether the case was reported in *The Law Reports*.

Independent variable	Coef	Standard error	p-value	Pearson correlation
Authority score	92.8268	91.998	0.313	0.06
Indegree	0.0081	0.001	0.0	0.55

TABLE 4 | Correlations between log geographic distance and number of citations. Geographic distance is measured in kilometres between national capital cities. “Overall,” shown in bold, denotes correlation calculated on all country pairs (excluding the United Kingdom). Correlations for individual countries are calculated using only the subset of country pairs each country belongs to. Larger negative values indicate stronger negative correlations between distance and citation count.

Country	Pearson correlation coef (w/Caribbean)	Pearson correlation coef (w/o caribbean)
Overall	**−0.055**	**−0.565**
United States	−0.409	−0.727
Canada	−0.391	−0.649
Australia	−0.776	−0.759
New Zealand	−0.789	−0.773
Singapore	−0.889	−0.888
Malaysia	−0.938	−0.929
South Africa	−0.408	−0.408
Hong Kong	−0.305	−0.305
Ireland	0.268	0.075

citations to its closest neighbor, Australia (63.1% of New Zealand’s non-United Kingdom outward citations). The flow of outward citations from Australia to New Zealand is also high (36.5% of Australia’s non-United Kingdom citations), although Australian courts are shown to have a stronger preference for Canadian jurisprudence (41.9%).

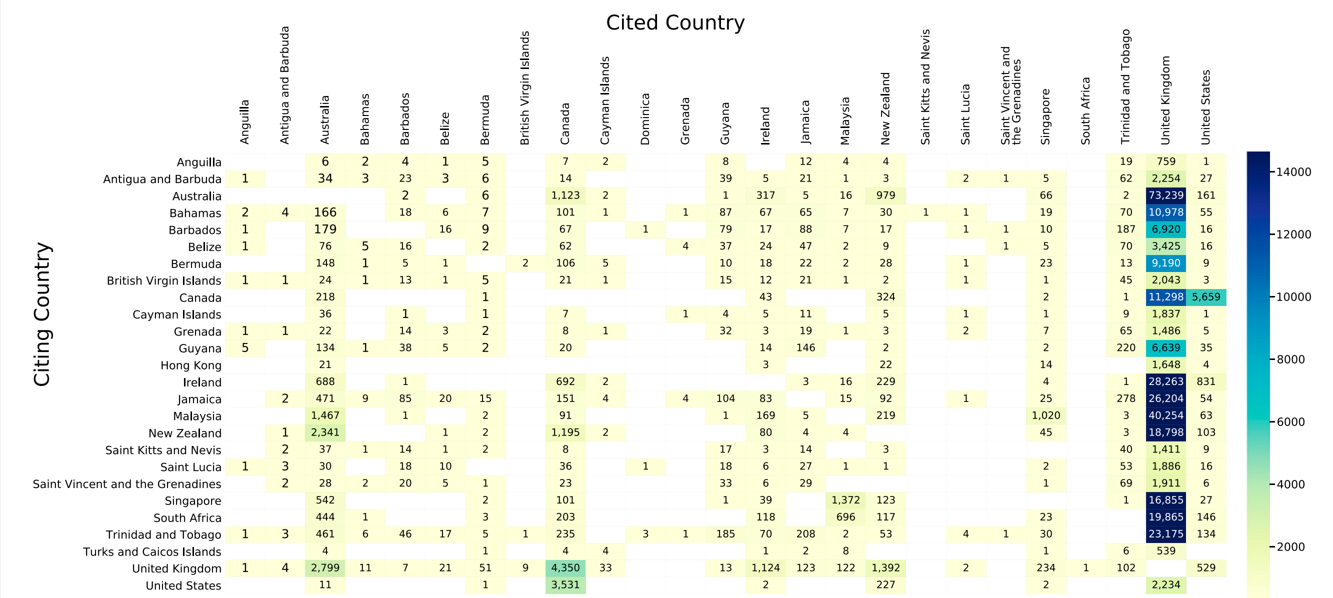
Caribbean Countries in the Caribbean region stand out as the most cosmopolitan users of foreign case law, spreading their outward citations across neighboring jurisdictions in the region. However, outward citations to foreign case law by courts in Caribbean jurisdictions are concentrated on Australia, Canada, United Kingdom and the United States.

To gain a clearer understanding of the extent to which geographic distance between countries influences the selection of foreign case law, we analyze the extent to which there is a negative correlation between distance and citation count. A strong negative correlation would provide support for **Assumption 3**. We calculate the log distances in kilometres between the capital cities of each pair of countries in our data along with the total inward and outward citations between each pair. Owing to its peculiar role from a historical perspective, we exclude the United Kingdom from our analysis. Further, in view of the fact that countries in the Caribbean region cite to many different countries, we compute the correlations both with and without these countries. Our analysis is shown in **Table 4**.

When the Caribbean data is included in the analysis, the overall correlation is weak (−0.055), although the strength of the correlation is high in countries situated in Asia and Australasia. In contrast, the negative correlation becomes stronger (−0.565) when we exclude Caribbean citation activity, with a marked increase in the observed correlation for Canada and the United States. The correlation for Ireland is weak in both scenarios because as shown in **Figure 11** the majority of Ireland’s citations are to the United Kingdom, which was excluded from the analysis.

In general, this analysis supports **Assumption 3**–when citing foreign case law, the courts of the citing country tend to prefer the jurisprudence of their geographic neighbors over the case law of countries that are further afield. However, distance alone is unlikely to provide an optimal surface of similarity between the countries in our data and it would be useful for future work to explore a wider range of factors.

The citation interactions between countries in the network also highlight the existence of a clique of countries–the United Kingdom (chiefly, England and Wales), Canada, Australia, New Zealand and the United States–that serve as *pillars* in the network, acting as the main hubs of case law to



Citing Country	Anguilla	Antigua and Barbuda	Australia	Bahamas	Barbados	Belize	Bermuda	British Virgin Islands	Canada	Cayman Islands	Dominica	Grenada	Guyana	Ireland	Jamaica	Malaysia	New Zealand	Saint Kitts and Nevis	Saint Lucia	Saint Vincent and the Grenadines	Singapore	South Africa	Trinidad and Tobago	United Kingdom	United States
Anguilla			6	2	4	1	5		7	2			8		12	4	4						19	759	1
Antigua and Barbuda	1		34	3	23	3	6		14				39	5	21	1	3		2	1	5		62	2,254	27
Australia					2		6		1,123	2			1	317	5	16	979				66		2	73,239	161
Bahamas	2	4	166		18	6	7		101	1		1	87	67	65	7	30	1	1		19		70	10,978	55
Barbados	1		179			16	9		67		1		79	17	88	7	17		1	1	10		187	6,920	16
Belize	1		76	5	16		2		62			4	37	24	47	2	9			1	5		70	3,425	16
Bermuda			148	1	5	1		2	106	5			10	18	22	2	28		1		23		13	9,190	9
British Virgin Islands	1	1	24	1	13	1	5		21	1			15	12	21	1	2		1		1		45	2,043	3
Canada			218				1							43			324				2		1	11,298	5,659
Cayman Islands			36		1		1		7			1	4	5	11		5		1		1		9	1,837	1
Grenada	1	1	22		14	3	2		8	1			32	3	19	1	3		2		7		65	1,486	5
Guyana	5		134	1	38	5	2		20					14	146		2				2		220	6,639	35
Hong Kong			21											3			22				14			1,648	4
Ireland			688		1				692	2					3	16	229				4		1	28,263	831
Jamaica		2	471	9	85	20	15		151	4		4	104	83		15	92		1		25		278	26,204	54
Malaysia			1,467		1		2		91				1	169	5		219				1,020		3	40,254	63
New Zealand		1	2,341			1	2		1,195	2				80	4	4					45		3	18,798	103
Saint Kitts and Nevis		2	37	1	14	1	2		8				17	3	14		3						40	1,411	9
Saint Lucia	1	3	30		18	10			36		1		18	6	27	1	1				2		53	1,886	16
Saint Vincent and the Grenadines		2	28	2	20	5	1		23				33	6	29						1		69	1,911	6
Singapore			542				2		101				1	39		1,372	123						1	16,855	27
South Africa			444	1			3		203					118		696	117				23			19,865	146
Trinidad and Tobago	1	3	461	6	46	17	5	1	235		3	1	185	70	208	2	53		4	1	30			23,175	134
Turks and Caicos Islands			4				1		4	4				1	2	8					1		6	539	
United Kingdom	1	4	2,799	11	7	21	51	9	4,350	33			13	1,124	123	122	1,392		2		234	1	102		529
United States			11				1		3,531					2			227				2			2,234	

FIGURE 11 | Heatmap of citations between countries in the cross-jurisdictional network. Citing countries are show on the left axis and citing countries are shown on the top axis.

which the other countries in the network reach. Not only do the members of this clique attract the most outward citation from the rest of the network, there is a considerable rate of exchange of authority between them. This suggests, based on our analysis of the data in the cross-jurisdictional network, that to the extent that there is an *emerging* "global community of courts", that community orbits, and is driven by, the jurisprudence of these five countries.

4.3 Evidence of Dialogue Between Domestic Courts Across Borders

L'Heureux-Dubé's [1] and Slaughter's [2] conceptions of a global community of courts place their emphasis on the notion of increased *dialogue* between the domestic courts of the countries in that community. L'Heureux-Dubé distinguishes *reception of law*, under which *country a*, the receiver, extensively cites the jurisprudence of *country b*, the donor, from a model of dialogue, under which *country a* and *country b* cite each other's jurisprudence on a more balanced footing. Reception is marked by a one-way exchange of judicial-wisdom; whereas by definition dialogue requires a two-way exchange of citation.

The clearest support for evidence of dialogue in the cross-jurisdictional network is seen among the jurisdictions in the Caribbean, where our analysis of the role of geographic proximity indicates a fairly even distribution of two-way citation between the courts of the nations in that region. We also observe a reasonable degree of exchange between the cross-jurisdictional network's "pillar" jurisdictions: Australia, Canada, New Zealand, the United Kingdom and the United States). There are indications in the matrix of citation interactions between the countries in the cross-jurisdictions that some of the dominant jurisdictions, notably the United Kingdom, Australia and New Zealand, are beginning to reach outwards beyond their clique toward the jurisprudence of the domestic courts of nations in the Caribbean and Asia-Pacific regions. To this limited extent, on the basis of the data included in this study, we conclude that the early signs of a transition from *reception* to *dialogue* are present in the network.

5 CONCLUSION AND FUTURE WORK

We constructed the network of the cross-border flow of citations to foreign case law in a substantial corpus of judgments given by the senior courts of twenty-six common law countries. In common with other complex networks, the cross-jurisdictional network presented low density and low global clustering. The degree distribution of the cross-jurisdictional network follows a pattern whereby most decisions in the network are cited by relatively few cases, while a small number of decisions are cited very frequently. This revealed a collection of 611 predominately English *super authorities* that attracted most of the citation activity in the network. The number of cases entering the network, the number of connections between those cases and the number of cases cited per judgment, increase from 1990 onward. We attribute the growth in the network to advances in technology during this period that served to improve online access to foreign case law [1,48,49].

The extent to which each country in our corpus engages with foreign case law varies. The courts of Australia and Canada were shown to be the most active users of foreign jurisprudence, while the United States is more restrictive in their use of decisions of other countries. In this respect, our findings correspond with those of Lefler [8] and Zaring [28]. The decisions of courts in the United Kingdom (chiefly, cases decided in the courts of England and Wales) are consulted the most. We attribute this to the fact that the legal systems of all of the countries in our data were based on the English common law model and were, or remain, members of the Commonwealth.

We tested three assumptions to gain a better understanding of the considerations judges may apply when citing foreign case law. Our first assumption was that judges would have historically used foreign cases cautiously and limited citation to cases of high importance, but that as the practice of foreign citation became more widespread, judges would lower their focus on the most influential cases. We used the authority scores calculated by the HITS algorithm [51] as a proxy for importance and found that the authority scores for cases cited declined over time. The same trend was observed in the analysis of our second assumption, which assumed that judges would have historically limited their reference to cases that were well-grounded in their domestic jurisprudence. Following the work of Lupu and Voeten (2011), we used the hub score calculated by HITS as a proxy for the extent to which a judgment was well-grounded in established jurisprudence and found that the average scores we historically high but declined with the passage of time. However, comparisons of these measures against random networks indicates that the reduction in the authority and hub scores over time is more likely to be a consequence of the rapid growth of the network from 1990 onward. Finally, we tested the assumption that judges tend to prefer to cite case law that emanates from countries that are geographically nearby over countries that are less proximal. We found this to be the case in our data, with localized dialogue between domestic courts in North America, the Caribbean and Asia-Pacific.

Notwithstanding evidence that national courts have a tendency to prefer case law from nearby jurisdictions, our analysis provides quantitative support for the overarching theme that there is an *emerging* global community of courts [1, 2]. While there are signs that reception of law is gradually giving way to increased dialogue, we found that the most of this dialogue is concentrated among a clique of countries–the United Kingdom, Canada, Australia, the United States and New Zealand. This work is far from complete, but our analysis lays the groundwork for future studies that can cast yet more light on the use of foreign case law by national courts. While our study utilizes a large corpus of senior court judgments from a broader range of common law countries, future work would ideally extend across a *wider* selection of countries, including systems based on the civil law model, and a *deeper* selection of courts, including courts lower down in the judicial hierarchy.

Finally, our finding that there has been a general increase in engagement with foreign case law over time raises the question as to whether this phenomenon points to a pattern of increasing convergence between national systems, and if so, the degree to which that convergence is distributed across substantive and procedural points of law.

DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because the dataset is proprietary. Requests to access the datasets should be directed to daniel.hoadley@mishcon.com.

AUTHOR CONTRIBUTIONS

DH, RC, AF, KT and AM conceived of the presented paper. RC and AF constructed the original data set. WH, NR and EN preprocessed the data set. BS, NR and WH conducted the temporal analysis experiments. DH and EN designed and conducted experiments with HITS. DH, AM and EN directed the research and interpretation of data. DH, AM and KT drafted the work. MB, JS, NW, BK, RC and AF critically revised the work.

REFERENCES

1. L'Heureux-Dube C. The Importance of Dialogue: Globalization and the International Impact of the Rehnquist Court. *Tulsa L Rev* (1998) 34:15–40.
2. Slaughter A-M. A Global Community of Courts. *Harv Int L J* (2003) 44: 191–219.
3. McLaughlin Mitchell S, and Powell EJ. *Domestic Law Goes Global*. Cambridge University Press (2011). doi:10.1017/CBO9780511783036
4. Schauer F. Authority and Authorities. *Va L Rev* (2008) 94:1931–61.
5. Glenn HP. Persuasive Authority. *McGill L J* (1987) 32:261–4.
6. La Forest GV. The Use of American Precedents in Canadian Courts. *Maine L Rev* (1994) 46:211.
7. Kirby MD. The New Biology and International Sharing: Lessons from the Life and Work of George P Smith Ii. *Ind J Glob Leg Stud* (2000) 7: 425–45.
8. Lefler R. A Comparision of Comparision: Use of Foreign Case Law as Persuasive Authority by the United States Supreme Court, the Supreme Court of canada, and the High Court of australia. *South Calif Interdiscip L J* (2001) 11:165–91.
9. Whalen R. Legal Networks: the Promises and Challenges of Legal Network Analysis. *Mich State L Rev* (2016) 2016:539.
10. Kosma MN. Measuring the Influence of Supreme Court Justices. *J Leg Stud* (1998) 27:333–72. doi:10.1086/468023
11. Landes WM, Lessig L, and Solimine ME. Judicial Influence: A Citation Analysis of Federal Courts of Appeals Judges. *J Leg Stud* (1998) 27: 271–332. doi:10.1086/468022
12. Caldeira GA. The Transmission of Legal Precedent: A Study of State Supreme Courts. *Am Polit Sci Rev* (1985) 79:178–94. doi:10.2307/1956126
13. Harris P. Ecology and Culture in the Communication of Precedent Among State Supreme Courts, 1870-1970. *L Soc Rev* (1985) 19:449–86. doi:10.2307/3053573
14. Hinkle RK, and Nelson MJ. The Transmission of Legal Precedent Among State Supreme Courts in the Twenty-First century. *State Polit Pol Q.* (2016) 16: 391–410. doi:10.1177/1532440015625328
15. Chandler SJ. The Network Structure of Supreme Court Jurisprudence. *Mathematica* (2005) 10 (3): 501–26. doi:10.3888/tmj.10.3-5
16. Bommarito MJ, Katz D, and Zelner J. Law as a Seamless Web? In: Proceedings of the 12th International Conference on Artificial Intelligence and Law (ICAIL 2009) (2009). p. 234–5. doi:10.1145/1568234.1568270
17. Fowler JH, and Jeon S. The Authority of Supreme Court Precedent. *Social networks* (2008) 30:16–30. doi:10.1016/j.socnet.2007.05.001
18. Fowler JH, Johnson TR, Spriggs JF, Jeon S, and Wahlbeck PJ. Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Polit Anal* (2007) 15:324–46. doi:10.1093/pan/mpm011

19. Lupu Y, and Fowler JH. Strategic Citations to Precedent on the U.S. Supreme Court. *J Leg Stud* (2013) 42:151–86. doi:10.1086/669125
20. Neale T. Citation Analysis of Canadian Case Law. *J Open Access L* (2013) 1:1.
21. Smyth R, and Mishra V. The Transmission of Legal Precedent across the Australian State Supreme Courts over the Twentieth century. *L Soc Rev* (2011) 45:139–70. doi:10.1111/j.1540-5893.2011.00430.x
22. Lupu Y, and Voeten E. Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights. *Br J. Polit. Sci.* (2012) 42:413–39. doi:10.1017/s0007123411000433
23. Lindholm J, Derlén M, Rosvall M, and Mirshahvalad A. Coherence Out of Chaos: Mapping european union Law by Running Randomly through the Maze of Cjeu Case Law. *Europarättslig tidskrift* (2012) 517–35.
24. Manley S. Referencing Patterns at the International Criminal Court. *Ejilaw* (2016) 27:191–214. doi:10.1093/ejil/chw002
25. Tarissan F, and Nollez-Goldbach R. Analysing the First Case of the International Criminal Court from a Network-Science Perspective. *J Complex Networks* (2016) 4:616–34. doi:10.1093/comnet/cnw002
26. Pauwelyn J. Minority Rules: Precedent and Participation before the Wto Appellate Body. In: J Jemielniak, L Nielsen, and HP Olsen, editors. *Forthcoming in Judicial Authority in International Economic Law*. Cambridge University Press (2015). p. 141–72. doi:10.1017/CBO9781316544860.006
27. Alscher W, and Charlotin D. The Growing Complexity of the International Court of justice's Self-Citation Network. *Eur J Int L* (2018) 29:83–112. doi:10.1093/ejil/chy002
28. Zaring D. The Use of Foreign Decisions by Federal Courts: An Empirical Analysis. *J Empirical Leg Stud* (2006) 3:297–331. doi:10.1111/j.1740-1461.2006.00071.x
29. Spottiswood S. The Use of Foreign Law by the High Court of australia. *Fed L Rev* (2018) 46:161–91. doi:10.1177/0067205x1804600201
30. Tyrrell H. Foreign Jurisprudence in the UK Supreme Court. In: *Human Rights in the UK and the Influence of Foreign Jurisprudence*. Oxford: Hart Publishing (2018): 74–121. doi:10.5040/9781509904976.ch-005
31. Ugander J, Karrer B, Backstrom L, and Marlow C. The Anatomy of the Facebook Social Graph. (2011). arXiv: abs/1111.4503
32. Busiello D, Suweis S, Hidalgo J, and Maritan A. Explorability and the Origin of Network Sparsity in Living Systems. *Sci Rep* (2017) 7:12323. doi:10.1038/s41598-017-12521-1
33. Liu W, Tao Y, Yang Z, and Bi K. Exploring and Visualizing the Patent Collaboration Network: A Case Study of Smart Grid Field in china. *Sustainability* (2019) 11:465. doi:10.3390/su11020465
34. Albert R, and Barabási A-L. Statistical Mechanics of Complex Networks. *Rev Mod Phys* (2002) 74:47. doi:10.1103/revmodphys.74.47
35. Hung S-W, and Wang A-P. Examining the Small World Phenomenon in the Patent Citation Network: a Case Study of the Radio Frequency Identification (RFID) Network. *Scientometrics* (2010) 82:121–34. doi:10.1007/s11192-009-0032-z
36. Chakraborty M, Byshkin M, and Crestani F. Patent Citation Network Analysis: A Perspective from Descriptive Statistics and Ergms. *PloS one* (2020) 15(12): e0241797. doi:10.1371/journal.pone.0241797
37. Albert R, Jeong H, and Barabási A-L. Diameter of the World-wide Web. *Nature* (1999) 401:130–1. doi:10.1038/43601
38. Ebel H, Mielsch L-I, and Bornholdt S. Scale-free Topology of E-Mail Networks. *Phys Rev E* (2002) 66:035103. doi:10.1103/physreve.66.035103
39. Price Dd. S. A General Theory of Bibliometric and Other Cumulative Advantage Processes. *J Am Soc Inf Sci* (1976) 27(5):292–306. doi:10.1002/asi.4630270505
40. Barabási A-L, and Albert R. Emergence of Scaling in Random Networks. *science* (1999) 286:509–12. doi:10.1126/science.286.5439.509
41. Posner R, and Landes W. Legal Precedent: A Theoretical and Empirical Analysis. *J L Econ* (1976) 19:249–307. doi:10.1086/466868
42. Flanders CW. Toward a Theory of Persuasive Authority. *Okla L Rev* (2009) 62: 55–88.
43. Vinkler P. Evaluation of Some Methods for the Relative Assessment of Scientific Publications. *Scientometrics* (1986) 10(3-4):157–77. doi:10.1007/bf02026039
44. Newman MEJ. The First-Mover Advantage in Scientific Publication. *Europhys Lett* (2009) 86:68001. doi:10.1209/0295-5075/86/68001
45. Vaccario G, Medo M, Wider N, and Mariani MS. Quantifying and Suppressing Ranking Bias in a Large Citation Network. *J Informetrics* (2017) 11:766–82. doi:10.1016/j.joi.2017.05.014
46. Qiu B, Ivanova K, Yen J, and Liu P. Behavior Evolution and Event-Driven Growth Dynamics in Social Networks, In 2010 IEEE Second International Conference on Social Computing, Minneapolis, MN, USA. IEEE (2010). p. 217–24. doi:10.1109/SocialCom.2010.38
47. Steer B, Cuadrado F, and Clegg R. Raphtory: Streaming Analysis of Distributed Temporal Graphs. *Future Generation Comp Syst* (2020) 102:453–64. doi:10.1016/j.future.2019.08.022
48. Abrahamson S, and Fisher MJ. All the World's a Court: Judging in the New Millennium. *Hofstra L Rev* (1997) 26:273–92.
49. Ginsburg RB. Looking Beyond Our Borders: The Value of a Comparative Perspective in Constitutional Adjudication. *Yale L Pol Rev* (2004) 22:329–37.
50. Parrish AL. Storm in a Teacup: The U.S. Supreme Court's Use of Foreign Law. *Ill L Rev* (2007) 2007:637–80. doi:10.2139/ssrn.891269
51. Kleinberg J. Authoritative Sources in a Hyperlinked Environment. In: *SODA '98: Proceedings of the Ninth Annual ACM-SIAM Symposium on Discrete Algorithms* (1998). p. 668–77.
52. Ciotti V, Bonaventura M, Nicosia V, Panzarasa P, and Latora V. Homophily and Missing Links in Citation Networks. *EPJ Data Sci* (2016) 5:7. doi:10.1140/epjds/s13688-016-0068-2
53. Kossinets G, and Watts DJ. Origins of Homophily in an Evolving Social Network. *Am J Sociol* (2009) 115:405–50. doi:10.1086/599247
54. Freeman L. Centrality in Social Networks: Conceptual Clarification. *Social Networkds* (1979) 1:215–39. doi:10.1016/0378-8733(78)90021-7
55. Bonacich P. Power and Centrality: A Family of Measures. *Am J Sociol* (1986) 95:1170–82. doi:10.1086/228631
56. Stone Sweet A. On the Constitutionalisation of the Convention: The European Court of Human Rights as a Constitutional Court. *Revue Trimestrielle des Droits de l'Homme* (2010) 80:923–44.

Conflict of Interest: Authors DH, WH, BK, AM, EN, NR, JS, KT and NW were employed by company Mishcon de Reya LLP. Authors RC and AF were employed by company vLex Justis Limited.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Physics

ORIGINAL RESEARCH
published: 16 July 2021
doi: 10.3389/fphy.2021.665563

Check for updates

Simulating Subject Communities in Case Law Citation Networks

*Jerrold Soh Tsin Howe**

Yong Pung How School of Law, Singapore Management University, Singapore, Singapore

We propose and evaluate generative models for case law citation networks that account for legal authority, subject relevance, and time decay. Since Common Law systems rely heavily on citations to precedent, case law citation networks present a special type of citation graph which existing models do not adequately reproduce. We describe a general framework for simulating node and edge generation processes in such networks, including a procedure for simulating case subjects, and experiment with four methods of modelling subject relevance: using subject similarity as linear features, as fitness coefficients, constraining the citable graph by subject, and computing subject-sensitive PageRank scores. Model properties are studied by simulation and compared against existing baselines. Promising approaches are then benchmarked against empirical networks from the United States and Singapore Supreme Courts. Our models better approximate the structural properties of both benchmarks, particularly in terms of subject structure. We show that differences in the approach for modelling subject relevance, as well as for normalizing attachment probabilities, produce significantly different network structures. Overall, using subject similarities as fitness coefficients in a sum-normalized attachment model provides the best approximation to both benchmarks. Our results shed light on the mechanics of legal citations as well as the community structure of case law citation networks. Researchers may use our models to simulate case law networks for other inquiries in legal network science.

Keywords: case law citation networks, legal network science, physics of law, network modelling, community detection

OPEN ACCESS

Edited by:
J.B. Ruhl,
Vanderbilt University, United States

Reviewed by:
Animesh Mukherjee,
Indian Institute of Technology Kharagpur, India
Diego R Amancio,
University of São Paulo, Brazil
Arthur Dyevre,
KU Leuven, Belgium

***Correspondence:**
Jerrold Soh Tsin Howe
jerroldsoh@smu.edu.sg

Specialty section:
This article was submitted to Interdisciplinary Physics, a section of the journal Frontiers in Physics

Received: *08 February 2021*
Accepted: *02 July 2021*
Published: *16 July 2021*

Citation:
Soh Tsin Howe J (2021) Simulating Subject Communities in Case Law Citation Networks. Front. Phys. 9:665563. doi: 10.3389/fphy.2021.665563

1 INTRODUCTION

Citations between cases form the bedrock of Common Law reasoning, organizing the law into directed graphs ripe for network analysis. A growing number of complexity theorists and legal scholars have sought to exploit legal networks to uncover insights about complex systems in general and legal systems in particular. Clough et al. [1] show that transitive reduction produces different effects on a citation network of judgments from the United States Supreme Court ("USSC") as compared to academic paper and patent networks. Fowler et al. [2, 3] pioneered using centrality analysis to quantify the authority of USSC precedent. This inquiry has been since been extended and applied to other courts [4] such as the Court of Justice of the European Union [5], the European Court of Human Rights [6], and the Singapore Court of Appeal ("SGCA") [7]. Examining case law citation networks ("CLCN"s) from the Supreme Courts of the United States, Canada, and India, Whalen et al. [8] find that cases whose citations have low average ages, but high variance within those ages are significantly more likely to later become highly influential. Beyond case law networks, Bommarito, Katz, and colleagues [9–11] have exploited the network structure of US and German legislation to study the growth of legal systems as well as the law's influence on society.

The community structure of CLCNs has received significantly less attention. This, however, is a rich area of research that retraces to seminal works in network science [12, 13]. Understanding communities broadly as connected subgraphs with denser within-set connectivity than without [14] allows us to automatically uncover network communities by iteratively removing links between otherwise dense subgraphs [13] or stochastically modelling link probabilities [15, 16]. A wide range of community detection techniques [17–21] as well as measures for evaluating community quality [22–24] have been studied. To our knowledge, two prior works have examined community structures in case law. Bommarito et al. [25] develop a distance measure for citation networks which they exploit to uncover communities in USSC judgments. Mirshahvalad and colleagues [26] use a network of European Court of Justice judgments to empirically benchmark a proposed method for identifying the significance of detected communities through random link perturbation.

Studying community structures in CLCNs can reveal deeper insights for both legal studies and network science. For legal studies, how far communities in CLCNs mirror legal doctrinal areas (e.g., torts and contracts) is telling of judicial (citation) practices. A judge who cites solely on doctrinal considerations should generate likewise doctrinal communities; one who cites for other (legal or political) reasons would transmit noisier signals. Community detection algorithms could also further the task of legal topic classification. Thus far, this has primarily been studied from a text-classification approach [27, 28].

For network science, CLCNs present a special case of the citation networks that have been studied extensively by the field. Studies mapping scientific papers as complex networks have demonstrated that they exhibit classic scale-free degree distributions [29] (but *cf* [30]). This has been attributed to preferential attachment, in that papers which have been cited more will be cited more. Other factors shaping paper citations include age [31] and text similarity [32]. These variables' interacting influences on citation formation yield rich structural dynamics in citation networks. For instance, over time, some papers come to be entrenched as central graph nodes while others fade into obsolescence, showing that age alone does not determine centrality [33]. Numerous generative models, discussed further in Part 2.2, have thus been proposed for citation networks, including for web hyperlinks [33–35].

As [1]'s findings suggest, however, the structure of CLCNs may differ from those of these traditional citation networks. In law, judges must consider the authority and relevance of precedent, amongst other things, when citing cases in their judgments. The doctrine of precedent further requires them to prefer certain citations to others. It is thus worth studying how CLCNs relate to traditional citation networks.

To this end, we examine how far generative models proposed for traditional citation networks can successfully replicate CLCNs. After a brief review of existing models (**Section 2.2**), we propose and evaluate a CLCN-tailored model that attempts to account for the unique mechanics of legal citations. The model is premised on an attachment function that attempts to capture aspects of legal authority, subject relevance, and time decay (**Section 2.3**). As measures for legal authority and decay are well-established, we focus on how subject relevance may be modelled. We devise a method for simulating node-level subjects and experiment with alternative attachment functions that incorporate these vectors in four different ways: using subject cosine similarity as a standalone linear feature in the attachment model; using the same as fitness coefficients [36]; constraining nodes to citing within subject-conditional "local worlds" [37]; using subjects to generate subject-sensitive PageRank scores [38] (**Section 2.3.3**). We then study by simulation the topological and community properties of networks produced by these alternative models (**Section 3.1**) and benchmark promising models (and baselines) against two empirical CLCNs: early decisions of the United States Supreme Court and of the Singapore Court of Appeal (**Section 3.2**).[1]

We find that using subject similarity scores as fitness coefficients within a sum-normalized probability function best approximates these actual networks. However, key differences remain between the simulated and actual networks, suggesting that other factors influencing legal citations are remain unaccounted for. Nonetheless, our work represents a first step towards better capturing and studying the mechanics of case law citation networks.

2 MATERIALS AND METHODS

Section 2.1 sets the legal theory and context behind case law citation formation. **Section 2.2** explores how far these are captured by existing models. **Section 2.3** explains our proposed models. **Section 2.4** describes the simulation protocol. **Section 2.5** explains the graph metrics used to evaluate the simulations. **Section 2.6** details the benchmark datasets.

2.1 Legal Context

We define a CLCN as a graph $G(N, E)$ where *all* nodes $n \in N$ are legal case judgments and all edges $e \in E$ citations between them.[2] Let k, k^- and k^+ respectively denote the degree, in-degree, and out-degree distributions of G. Nodes may have attributes such as the authoring judge, decision date, legal subject, and the text of the judgment. Edges may be weighted (e.g., if a judgment cites another more than once), and have attributes such as whether the citation affirms or overrules the cited case.

Like all citation networks, CLCNs are time-directed and acyclic.[3] CLCNs are unique, however, because the probability that a new node n_t cites an earlier n_{t-1} (denoted $P((n_t, n_{t-1}))$ and the entire distribution P) is shaped by legal theory and practice.

[1]All models and simulation methods are implemented in Python and will subsequently be made available in a GitHub repository.

[2]Though judgments often cite to other forms of law such as statutes. We leave these such hybrid networks aside for now.

[3]Judgments published close in time *could* cite one another. This occurs, albeit rarely, in the Singapore dataset below.

Posner [39] identifies five overarching reasons for legal *paper* citations, namely to:

1. Acknowledge priority or influence of prior art
2. Provide bibliographic or substantive information
3. Focus disagreements
4. Appeal to authority
5. Reinforce the prestige of one's own or another's work

Reason (4) is particularly pertinent to *case* citations in Common Law systems characterized by the doctrine of binding precedent. The doctrine, in brief, means propositions of law central to a court's essential holding are taken as binding law for future purposes. Lower courts are bound to follow these holdings. While courts at the same level of hierarchy are not technically bound the same way, great deference is generally accorded to past cases nonetheless.

Recent studies have thus sought to measure legal authority with network centrality measures calibrated for the legal domain [2, 3]. Beyond citation counts, a judgment's authority is further shaped by its subject areas and time context [40]. Lawyers do not think of judgments as simply authoritative in the abstract, but within a given doctrinal subject area (i.e., torts or contracts) and at a given time. Precedential value waxes and wanes as a judgment gets entrenched by subsequent citations, ages into obsolescence over time, and as other complementary or substitute judgments emerge [7, 39]. Relevance, authority, and age are thus key, interconnected drivers of CLCN link formation [8].

2.2 Existing Models

How far are these legal mechanics captured by existing network models? In this section, we review existing citation network generative models and consider how they may be used to simulate CLCNs. Note that this paper is not a comprehensive review and will only highlight illustrative examples.

2.2.1 Degree-based Models

Classic Barabasi-Albert ("BA") [41] preferential attachment sets $P_{BA} = \frac{k_i}{\sum_j k_j}$. This model famously recovers scale-free degree distributions observed in empirical networks. However, in this model earlier nodes acquire a significant and permanent advantage over later ones, particularly if the former are cited early on. Thus, a known limitation [42] of using BA model for simulating citation networks is that, since $P_{BA} = 0$ whenever $k = 0$, new nodes (which necessarily have $k^- = 0$) are very unlikely to gain citations. Thus, the final graph may be such that most subsequent nodes cite the root node. This, of course, does not occur in empirical CLCNs (see also **Section 3.2**).

The "copying" model [34] offers a partial workaround. Links are determined by first randomly choosing one node from N as a "prototype", denoted n_p. Destinations are then either selected randomly from N by a coin toss with manually-specified probability α or copied from n_p otherwise. While the model does not explicitly include k in its attachment process, notice that nodes with zero in-degree may only be cited under the former branch, while nodes with high in-degree are more likely to be cited under the latter branch. The copying model can therefore be broadly understood as a mixture between the Erdos-Renyi and BA models with mixture intensity controlled by α. Setting $\alpha = 1$ recovers Erdos-Renyi completely, though the model with $\alpha = 0$ is not completely equivalent to BA. This allows the copying model to produce scale-free degree distributions while leaving open the possibility for $k = 0$ nodes to be cited. However, these nodes are still less likely (depending on α) to be cited, as they cannot be cited under the copying branch. Moreover, the random process used for selecting prototypes and deciding whether to copy does not accord with legal intuitions. We do not expect new judgments (or papers) to randomly choose older judgments to either cite or copy citations from.

Another alternative proposed by Bommarito et al. ("BEA") [43] is a generalizable attachment function which considers in- and out-degree separately. More precisely,

$$P_{BEA}((n_t, N_{t-1})) = \frac{e^{\alpha k^-_{N_{t-1}} + \beta k^+_{N_{t-1}}}}{\sum_{i=1}^{|N_{t-1}|} e^{\alpha k^-_i + \beta k^+_i}} \tag{1}$$

where k^-_i and k^+_i are node i's in- and out-degree respectively, and $\{\alpha, \beta\} \in \mathbb{R}$ are parameters for tuning their influences on P. Denoting $\{k^-_i, k^+_i\}$ as a single feature vector X_i and $\{\alpha, \beta\}$ as weight vector B, (1) may be rewritten as

$$P_{BEA}((n_t, N_{t-1})) = \frac{e^{BX_{N_{t-1}}}}{\sum_{i=1}^{|N_{t-1}|} e^{BX_i}} = softmax(BX) \tag{2}$$

Because the softmax has a vector smoothing effect, using it over conventional sum normalization ensures that non-zero citation probabilities are assigned to all nodes, even for nodes where $k = 0$. Seen this way, BEA provides a readily-extensible framework for modelling citation networks. BX is capable of encompassing an arbitrary range of weights and features.[4]

The BA and BEA models may be seen as instances of what Pham et al. [44] call the General Temporal ("GT") model. GT generalizes k into an arbitrary function of node degree $A(k)$, known as the "attachment kernel", such that $P_{GT} \propto A(k)$. The GT framework allows a large class of degree-based attachment models to be specified and estimated by maximum likelihood. For instance [44], simulate networks with $A(k) = 3((log(max(k,1)))^2 + 1$. GT thus offers an attractive framework for modelling legal authority in CLCN link formation. But, despite its name, GT attachment does not explicitly model node age. Yet, age has been identified as a factor driving citation networks, including CLCNs [8].

2.2.2 Aging Models

More generally, degree-based models generally ignore the well-documented influence of node age on citation formation [31, 42, 45]. By contrast, "aging" models [31, 45] propose introducing a decay vector $w(n,t)$ such that $P \propto A(k) \times w(n,t)$. Here, $w(n,t)$ can be any standard decay function which takes maps every n to weights bounded by $[0, 1]$ based on their arrival time t_n. Decay

[4]Notice that this effectively models link probabilities with a multi-class logistic model.

functions are further monotonically non-increasing with item age $a_n = t - t_n$, and assign weight 1 to nodes with $a_n = 0$ [46]. For instance, a simple sliding window assigns all items younger than a cut-off age to weight 1 and all other items to weight 0.

The specific aging model proposed by Wang et al. [31] uses node in-degree and exponential decay such that $P_{aging} \propto k^{-} \times \exp(-\tau a)$. In exponential decay functions, the parameter τ controls decay rate and induces a fixed half-life of $\frac{ln(2)}{\tau}$. Thus, the aging model gives younger nodes a higher chance to be cited than older ones with the same in-degree. However, nodes with $k^{-} = 0$ still have zero probability of being cited, regardless of age.

One extension of the aging model which incorporates intuitions from copying models are Singh et al.'s "relay" models [33]. Like copying models, relay models first choose a prototype n_p and performs a coin toss to determine the next step. But unlike copying models, prototypes are chosen by BA preferential attachment instead of randomly. The first coin's head probability is given by $\exp(-\tau a_n)$ (that is, exponential decay). On heads, n_p itself is cited and the process ends. On tails, a *second* coin toss with manually-specified head probability θ decides if citations are "relayed" (on heads) or if n_p will be cited nonetheless (on tails). In a "relay", a new prototype is selected from within the set of nodes *citing* n_p via a specified distribution D (Singh et al. use either uniform-random or preferential attachment). The process repeats until broken by a coin toss or the maximum specified relay depth is reached (in which case the final prototype is cited).

The exponential decay which parameterizes the first coin toss means aged papers are less likely to be cited themselves than they are to relay citations to younger papers citing them. At the same time, since prototypes are chosen initially by preferential attachment and subsequently re-chosen by D, relay models incorporate aspects of degree-based, scale-free models [33]. show that relay models better fit empirically-observed distributions of paper citation age gaps (i.e. the age difference between citing and cited papers) than the classic aging model. Relay models thus provide a more sophisticated method to account for both degree and age simultaneously. What remains missing, however, is a way to incorporate subject relevance as well. We thus turn to examine "fitness" models.

2.2.3 Fitness Models

Fitness models [47] attempt to account for each node's innate ability to compete for citations. This is generally achieved by introducing a vector of node fitness coefficients η_i. For instance, the Bianconi-Barabasi model [36] introduces a vector of uniform-randomly sampled ηs to classic preferential attachment such that $P_{BB} = \frac{\eta_i k_i}{\sum_j \eta_j k_j}$. Introducing η weakens the monopoly k holds over citation probabilities in P_{BA}. A fit node has a good chance of being cited even if its degree is low (though not if its degree is zero) [48]. further propose introducing a time-decay vector w, such that the final attachment function becomes $P = \frac{\eta_i A(k_i) w}{\sum_j \eta_j A(k_j) w}$. Notice that fitness in this regard may represent any arbitrary attribute other than degree which is believed to influence citation probabilities. For instance [42], use the ratio between (a) the theoretical number of citations a node should receive under BA and (b) the actual number received to measure the "relevance" of a node.

Likewise, if we conceptualize η, $A(k)$, and w as capturing legal relevance, authority, and time effects respectively, this three-variable model appears ideal for modelling legal citations. Here, degree-based centrality scores (an instance of $A(k)$) have been shown to capture legal authority well (see **Section 2.3.1** below). Modelling time effects with w is also relatively standard. The crux, therefore, is devising *eta*s that capture subject relevance. This turns on the distribution it is sampled from. Drawing fitness uniformly from [0, 1], as in [36], yields in expectation an evenly distributed node ranking inconsistent with the intuition that nodes sharing subjects with the citing node should be fitter (that is, more relevant) than others.

One workaround is to calculate *eta* empirically from the *text content* of actual papers [35]. use the cosine similarity between (stopped and lemmatized) word frequency distributions of two papers' texts to their content similarity. They then propose a "three-feature model" which places content similarity scores alongside in-degree preferential attachment and power-law time decay as competing node attachment distributions. The model randomly chooses one of the three (with probability α, β, γ respectively, $\alpha + \beta + \gamma = 1$) as the final attachment function. This creates a probabilistic mixture between the content similarity, degree, and aging models, although only one model is ultimate used to generate any given edge.

Using text similarity measures to capture content overlap is intuitively logical and allows us to exploit the growing literature on text embeddings [49] (which find increasing representation in legal studies as well [7, 50]). The main drawback is that because text is difficult to simulate, we are limited to simulating *edges* between actual, existing nodes. To illustrate, for CLCNs, we can compute text similarity between empirically-observed case judgments and simulate citations between them. This would, of course, reveal important insights about CLCNs. However, generating the case judgments themselves would be difficult. We would not be able to deviate from empirically-observed node attributes.

To summarise, existing literature provides a wealth of citation network generation models. Each have their own strengths and weaknesses when theoretically applied to CLCNs. At the same time, we are not aware of any study that attempts to do this. Building on this literature, the next Section proposes a model tailored for CLCNs.

2.3 Modelling Case Law Citation Networks

Following [43], we start at time $t = 0$ with G_0 comprising $|N_0| = 1$ node and $|E_0| = 0$ edges. For each t till a specified stop T, $|N_t| \sim Po(\lambda)$ new nodes are added. Each n_t cites $|E_{n_t}| \sim Po(\mu)$ prior nodes (re-drawn independently per node). $\{\lambda, \mu\} \in \mathbb{R}$ thus control network growth rates. The main innovation of our model lies in the attachment function. In the abstract, we use a probability function $P = f(\Lambda, \rho, w)$ where Λ generalizes $A(k)$ above to encompass any function, including functions not based on k alone, capable of measuring legal authority, ρ measures subject relevance, and w is a time-decay function.

The goal is to calibrate these variables in a legally-contextualized manner. Below we expand on each variable in turn.

2.3.1 Authority

In place of k^- and k^+ above, we use Kleinberg's [51] authority and hub scores. These have been shown to accurately recover legally-significant cases from CLCNs [2, 3, 5, 40, 52]. We denote authority and hub scores as $A(k^-)$ and $A(k^+)$ respectively. While we might intuitively expect in-degree to be more representative of authority than out-degree, legal scholars have found that out-degree-based scores can be a better predictor of future case influence [40]. Cases which discuss and synthesize a large number of authorities tend to represent important disputes into which significant legal and financial resources are poured. For similar reasons, they also tend to become important legal checkpoints themselves. We therefore include both score types in the model. In any event, α or β could be set to zero to remove either score.

2.3.2 Time

Following the aging model, we use a standard exponential decay where $w(n,t) = \exp(-\tau a_n)$. This suits the legal context because citations to centuries-old judgments are not uncommon. Thus, discrete decay functions like the sliding window that apply a standard discount to all papers above a certain age would not model this observation well. Given the law's respect for old authority, we assume throughout this paper that $\tau = 0.01$, resulting in a precedent half-life of about 70 periods. Of course, future work could explore how τ may be empirically estimated (see [31]) and how it may vary over time, jurisdiction, subject, or even judge.

2.3.3 Subject Relevance

To derive relevance, we first need to simulate subjects for each node. Drawing inspiration from Latent Dirichlet Allocation ("LDA") [53], we assign each node a vector of m subjects $\phi_i \sim Dir(\psi, m)$. ψ is a m-sized vector that controls subject skew. If we want some subjects to occur more than others, possibly following a power-law, a similarly skewed ψ may be used. As a null model, however, we may set $\psi = \frac{1}{\mathbf{m}}$.

One drawback of the Dirichlet is that non-zero probabilities may occur across many subjects. This is inconsistent with how legal cases generally discuss only a few subjects. Thus, we set a minimum cut-off of 0.1 below which subject values are floored to 0. The vector is then normalized to sum back to 1. Should this cutoff result in an entirely zero vector, one randomly-chosen subject is assigned weight 1. Because LDA treats documents as finite mixtures over m latent overlapping 'topics' that are in turn multinomial distributions across words, such a cut-off is intuitively similar to assuming that any subject generating less than 10% of the words in a judgment is not a subject that should be ascribed to it.

These subject vectors are analogous to overlapping community belonging coefficients [24], though it is always possible to partition nodes into discrete subjects by taking, for instance, $max(\phi_i)$). Here, we interpret ϕ as non-fuzzy subject *proportions* rather than probabilities. That is, a case with $\psi_i = \{0.51, 0.25, 0.24\}$ has subjects 1, 2, and 3, each with probability 1. But it is *primarily about* subject 1, in that 51% of its content is expected to come from the same.

Given ϕ, subject relevance can be modelled in at least four ways:

As Linear Features: First, we can derive subject similarity scores $\rho^{sim}_{n_t,N_{t-1}} = g(\phi_{n_t}, \phi_{N_{t-1}})$, where g is some vector similarity measure. Many options for g exist, but for now we default to cosine similarity given its established use in document clustering, including for legal documents [54]. The simplest way to model relevance is then to include ρ^{sim} as a standalone linear feature with its own weight γ such that $P_{linear} \propto w \times (\alpha A(k^-) + \beta A(k^+) + \gamma\rho^{sim})$.

As Fitness Coefficients: Including ρ^{sim} linearly is attractively simple, but may fail to account for interaction effects between authority, relevance, and time. Thus, a second approach is to model ρ^{sim} as fitness coefficients, so $P_{fitness} \propto \rho_{sim} \times w \times (\alpha A(k^-) + \beta A(k^+))$. This is broadly similar to the model proposed in [48], except that fitness values are computed from simulated subjects. Notice that, unlike with the linear features approach, using subject similarities as coefficients ensures that prior nodes with zero subject overlap with the citing node will be assigned $P = 0$ as well.

As Locality Constraints: Another more direct to enforce this is to limited nodes to citing within subject-conditional "local words" [37]. Within each locality, nodes can be selected by any subject-conditional probability distribution [37].'s original paper used the uniform distribution. To account for legal authority and time effects, we continue to choose nodes using HITS scores and exponential decay. To approximate the idea of nodes being authoritative within subjects, HITS scores are re-computed within the subgraph of nodes sharing at least one subject with the citer. More precisely then, $P_{local} \propto w \times (\alpha A_{local}(k^-) + \beta A_{local}(k^+))$, with the subscript 'local' denoting that the vector is computed on a subject-local subgraph.

For Subject-Sensitive PageRank: Another way to interact subject overlap with degree-based authority is to use ϕ to compute so-called topic-sensitive PageRank ("TSPR") scores [38] that may be used in place of HITS scores. While conventional PageRank [55] produces one global ranking that disregards node topic, TSPR first calculates m (the total number of topics) different rankings by setting non-uniform personalization vectors for each topic given by

$$v(n_i) = \begin{cases} \dfrac{1}{|N_j|} & \text{if } n_i \in N_j \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where N_m is the set of nodes with subject j. Given a query node q with topic weights ϕ_q, TSPR then returns $TSPR_q = \sum \phi_q \cdot PR_j$, with PR_j being personalized PageRank scores for topic' j. While TSPR has not been studied in legal networks literature, it offers a promising way to simultaneously account for authority and subject relevance using just one centrality measure. Thus, $P_{TSPR} \propto w \times TSPR_q$.

The subject models above are, of course, based on the literature reviewed in **Section 2.2**. The best approach for modelling subject relevance is not obvious. Neither are the approaches mutually exclusive. For instance, after constraining the citable node set by subject, we may still include ρ as a linear feature while also using TSPR scores to model authority. Other combinations are also theoretically possible. But doing so may lead to contradictions. For example, calculating TSPR within a subject-constrained subgraph will return the simple PageRank score of that subgraph. It may also overplay the importance of subject relevance. For now, we study the properties of the networks produced by each approach independently.

To summarise, the proposed subject models begin with one root node and, at each time step t, adds $|N_t| \sim Po(\lambda)$ nodes with $|E_{n_t}| \sim Po(\mu)$ edges per node. Attachment probabilities are specified generally by $P = f(A(k), w(t), \phi)$, where $A(k)$ is some degree-based centrality measure including simple in/out-degree, HITS scores, and TSPR, $w(t)$ is an exponential decay function, and $\phi \sim Dir(\psi, m)$ a vector of simulated node subjects which may be incorporated into P in four different ways proposed here (though we do not rule out alternatives).

2.4 Model Simulations

To study the properties of our proposed models, we simulate 50 iterations of $T = 500$ steps for each subject model. Building on [43], we experimented with softmax as well as sum normalization for each model. For ease of reference, below we refer to the four proposed subject models as Linear, Fitness, Locality, and TSPR respectively. We use brackets to identify the normalization scheme. To illustrate, TSPR (sum) refers to a sum-normalized attachment model based on exponentially-decayed subject-sensitive PageRank scores.

For baseline comparison, we also simulated the BA, BEA, aging, copying, and relay models. We ran BA with *degree*-based preferential attachment (not in-degree), following the original model. Likewise, BEA was simulated with α and β both equal to 1. The aging baseline follows [31]'s specification, using only in-degree and an exponential decay with $tau = 0.01$. A softmax-normalized alternative was tested as well. The copying model was run with copying probability $\alpha = 0.5$ (not to be confused with the in-degree weight α in our models). Finally, we used preferential attachment relay and set relay depth at 1, $\tau = 0.3$, and $\theta = 0.9$. This follows optimal parameters found by Singh et al. for approximating the scientific paper networks they studied.

Including 5 baselines, subject models, and alternative normalizations, a total of 16 different models are run for 50 iterations each.[5] To promote comparisons across models, we fix a few key parameters in our simulations. First, $|N_t|$ is fixed universally at 1. Thus, exactly one node is added at every step for every simulation. Second, the number of subjects is fixed at 30. Third, within each iteration we first draw all $|E_{n_t}|$s from $Po(5)$ and all ϕ_{n_t}s from $Dir\left(\frac{1}{30}, 30\right)$ and use the same inputs across all models/approaches. This means the out-degree distribution of all models in the same iteration will be similar. Further, because the same subject vectors are used across all parameterizations *within* the same iteration, only 50×501 individual node subject vectors are generated.[6] Fourth, all weights $\{\alpha, \beta, \gamma\}$ are set at 1 whenever relevant (though recall that γ is only used by the linear feature subject model). Finally, an exponential decay with $\tau = 0.01$ is used for all models (except the relay model).

A few implementation details are worth noting. First, because $|E_{n_t}|$ is randomly-drawn, it can exceed the total number of nodes in the existing, citable graph. Further, some attachment models result in zero citation probabilities for certain nodes, further limiting the citable node set. Thus, whenever $|N_{t-1}| < |E_{n_t}|$, we draw only $|N_{t-1}$ destinations (while still using P). As a result, a node's realized out-degree can be lower than its initially-drawn out-degree. This is more likely to occur in the Locality models since nodes may only cite within subject-local subgraphs. This accounts for minor differences in total edge counts across model simulations. Second, nodes and edges are added in batches *after* attachment probabilities and edge destinations for every new node at a given t is determined. All computations are based solely on G_{t-1}, so nodes and edges added at the same t do not influence computations for each other. Third, after P is calculated, destination nodes are selected *without* replacement, so $|E_{n_t}|$ *unique* destination nodes are always drawn. This follows prior literature which (implicitly) samples without replacement [43].

Finally, we use NetworkX's [56] Python package to compute HITS scores. Since convergence is not guaranteed, we allow the algorithm to run for a maximum of 300 power iterations, three times the package default. We modify the package slightly to return prevailing scores if convergence is not achieved by then. To facilitate convergence (and save computational resources), we exploit the intuition that HITS scores for step $t + 1$ should not differ too much from those of step t and provide HITS scores from previous steps as warm starts. Note that this cannot be done for Locality. Because the citable node set in that model varies from node to node, relevant prior HITS scores vary.

2.5 Model Properties

An important preliminary question is whether the subject models yield scale-free degree distributions even as they seek to model time and relevance effects. As our simulation protocol fixes out-degree distributions, so comparing out-degree or total degree distribution is less meaningful. Thus, we begin by examining each model's realised in-degree distributions. To compute the average distribution across 50 iterations of the same model, we stack distributions on each other to produce a 50×501 matrix of in-degree counts. We then take the column-wise *mode* of this matrix[7] as the average distribution and compute the frequency-rank distribution of the same.

To further examine how the baseline and subject models differ in subject structure, we also derive subject signatures for each network. These are broadly inspired by [33]'s temporal bucket signatures. More precisely, denote the subject edge histogram of a

[5]Other parameterizations for the copying ($\alpha \in \{0.1, 0.9\}$) and relay models(uniform relay) were run but do not change our results.

[6]This includes the root node plus 500 node-time steps.

[7]Mean and median are not suitable here since they yield decimal values that would render frequency counts problematic. Where more than one mode exists we take the smallest value.

graph G whose nodes fall within m subjects as a size $m \times m$ matrix $H(i, j)$ where each entry is the total number of times nodes with subject j have cited nodes with subject i. Because one node may have many subjects, a single edge can add to many entries.[8] Thus, $|H| \geq |E|$.

The global-sum-normalized matrix $H_{norm} = \frac{H}{|H|}$ then yields the unconditional probability distribution for the cited and citing subject pairings of an arbitrary edge. Meanwhile, normalizing H column-wise so that $H_{col} = \frac{H}{|H(j)|}$ yields in each column the probability of subject i being cited conditional on the citing edge having subject j. In this way, H, H_{norm} and H_{col} offer different insights on the subject signature of a single network. Subject signatures for each *model* may then be computed by averaging these matrices across model iterations.

Finally, we compute a range of network density and community quality metrics for selected models. These include intra/inter-community edge ratios [57] and link modularity, being [18]'s modularity scores extended to the case of overlapping, directed communities [22]. We compute these metrics against (1) the simulated ϕs themselves (as ground truth subject labels) and (2) communities recovered by k-clique percolation. k-Clique percolation is a useful baseline because it is an established method for *overlapping* community detection [14]. Briefly, it recovers communities by percolating k-sized cliques to adjacent k-cliques (i.e., those sharing k-1 nodes). Its time-efficiency also means running the algorithm on all simulated networks is more practical than other equally established but less efficient algorithms, such as Girvan-Newman edge-betweenness [13]. We fix k at 3, the smallest meaningful input, to allow more, smaller communities to be returned. Though a directed k-clique algorithm exists [58], we could not find any open-source implementation. As an accessible baseline was desired, we relied on NetworkX's undirected implementation instead. Code for k-clique percolation and all community quality metrics are from CDLIB [59].

2.6 Empirical Benchmarks

Studying the structural properties and subject signatures of the simulations identifies certain more promising approaches for modelling CLCNs. As a final step, we benchmark selected approaches against two empirical CLCNs. The first is the internal network of USSC judgments that is well-studied in legal network science. To obtain legal subjects, we join Fowler et al.'s [3] edgelist with metadata from the Spaeth database [60], particularly the "issue areas" identified for each case. The second is an internal network of SGCA judgments that has also been studied in prior work [7]. The dataset covers citations between all reported decisions of the SGCA from 1965 (the year Singapore gained independence) to 2017. Judgments are assigned to subjects using catchwords provided by the Singapore Law Reports, the authoritative reporter of SGCA judgments. Following [28], we map these subject labels to 31 unique subject areas (including the "Others" category). Note that in both datasets subjects are overlapping in that the same case may belong to more than one subject.

These networks represent different CLCN archetypes. Although both the United States and Singapore inherited English law, the legal, social, and time contexts in which each system originated and developed is vastly different. Further, while the USSC primarily (and selectively) *reviews* cases of federal and constitutional significance, the SGCA, as its name suggests, routinely considers *appeals* on matters of substantive (including private) law. The datasets are also practically usefully because both provide human-labelled legal subjects. To be sure, this also implies that comparisons between the two networks must be made with caution. On top of their different legal contexts, the legal subjects provided by each database also differ. The Spaeth database uses broad issue areas such as "Civil Rights" and "Due Process". The Singapore dataset uses specific doctrinal areas such as torts and contracts. Below we refrain from drawing comparisons between the two networks except on broad properties such as degree distributions.

As we are primarily interested in network generation, we focus on the first 2001 nodes of the USSC network and the first 1001 of the SGCA network (making allowance for one root node). More USSC nodes were used because the early USSC graph was sparser. The resultant USSC and SGCA benchmark graphs had 777 and 779 edges each. The USSC benchmark spans from the year 1791 (the first node) to 1852 (the 2001th node), while the SGCA benchmark spans from 1970 to 1999. More detailed properties of both graphs are discussed in **Section 3.2**.

To set up the comparison, we first calibrated the model with the empirical properties of each network. Specifically, USSC simulations used an average edge rate per case of $\frac{777}{2001} = 0.388$. Subject vectors were drawn across 14 issue areas from $Dir(\psi_{us}, 14)$, with ψ_{us} being the normalized issue frequency distribution of the benchmark graph. Likewise, SGCA simulations were run using edge rates of $\frac{779}{1001} = 0.778$ with subjects drawn from $Dir(\psi_{sg}, 31)$. As with the initial simulations, we fix $|N_t| = 1$ for all t and pre-draw all $|E_t|$s and ϕs per iteration.

All benchmark simulations are run and assessed using the same protocols and metrics described in **Section 2.4**. The implementation details noted there apply. Additionally, we exploit the subject signatures to compute vector distances between the simulated and empirical graphs. In particular, we take the distance between H_{norm} as an indicator of the aggregate differences between the subject structure of two graphs. We also measure distances between the main diagonals/off-diagonals alone for insight into differences between intra/inter-subject structure. Column-wise distances between H_{col}s can be taken as a measure of per-subject differences in structure.

Amongst the wide range of graph distance measures available (see [61] for a review) we use the L1 distance because of its simple, intuitive interpretation as the sum of absolute differences.[9] More precisely, $L1(H_1, H_2) = \sum_j \sum_i |H_1(i, j) - H_2(i, j)|$. At the same time,

[8] To illustrate, if a node with subjects 1 and 2 cites a node with subjects 3 and 4, entries(3,1), (4,1), (3,2) and (4,2) are all incremented.

[9] We also experimented with Kullbeck-Leibler divergence and obtained similar results, though that measure often returns infinity.

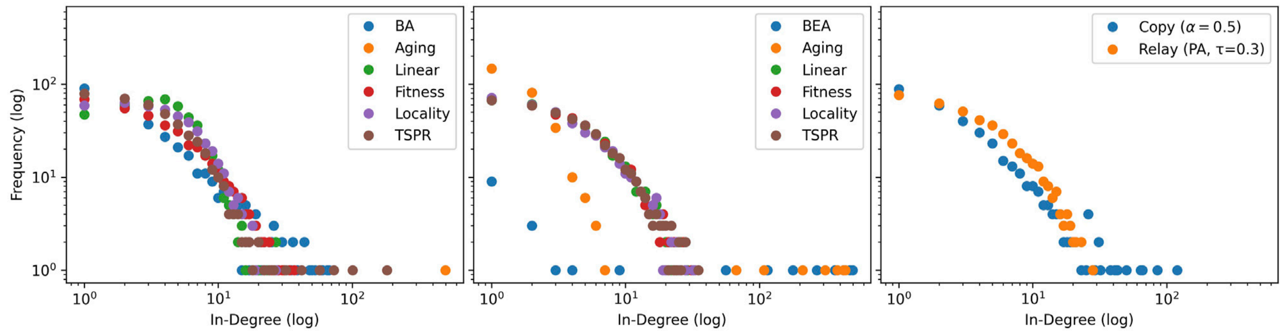


FIGURE 1 | Mode-averaged in-degree distributions for simulated sum-normalized **(left)**, softmax-normalized(centre), and other baseline models **(right)**. 50 iterations of 500 node steps are run per model. Models in the same iteration number use identical, pre-drawn out-degree counts from $|E_n| \sim Po(5)$ and subjects from $\phi \sim Dir\left(\frac{1}{30}, 30\right)$.

classic measures such as the Hamming and Jaccard distances are meant primarily for binary (adjacency) matrices. Future work could explore more tailored distance measures. In particular, since H and H_{norm} may also be interpreted as adjacency matrices for a weighted and directed meta-graph between the *subjects*, a distance measure specialized for such graphs (e.g. [62, 63]) may be useful.

Importantly, note that subject indexes must be aligned before computing most vector distances, including the L1, as entry order affects results. In our context, this is akin to ensuring that subject i of H_1 is comparable to subject i of H_2. As our simulated subjects are arbitrary, we cannot order them by substantial content: say, to place torts at index 0 and contracts at index 1. Instead, we place the most frequent subject at index 0 and the least frequent at $m - 1$. Since the most common subjects are often also more likely to be cited (simply as a function of frequency), frequency indexing desirably concentrates probabilities towards the top-left quarter of the matrix, presenting a visually-readable signature (see **Section 3.2**). When computing signature distances, therefore, we are comparing citation patterns across subjects as ranked by frequency. If the most common subjects in graphs G_1 and G_2 both primarily cite other common subjects, signature distances would be relatively small. But if common subjects in G2 primarily cite its least common subjects, signature distances would be larger.

3 RESULTS

3.1 Model Simulations

As shown in **Figure 1**, most subject models successfully generate scale-free in-degree distributions similar to baselines. We also observe that Aging (sum) produces an average in-degree distribution with one node monopolizing most edges. Because aging models consider only in-degree, sum-normalization leads exactly to the problem, discussed in **Section 2.2**, where new nodes are never cited. This is partially addressed in by softmax normalization, but the Aging (softmax) model still manifests a visibly more imbalanced in-degree distribution than others. The same is also true for the BEA model. This is because even though the softmax function generally assigns non-zero probabilities, small-valued input elements are quickly assigned vanishingly small values. Consider for example that $softmax(5, 5, 1) = 0.495, 0.495, 0.009$ at three decimal places. This affects models like BEA and Aging, which use simple degree counts as inputs, more directly because differences in feature values are larger.

The models are further differentiated by subject signatures. **Figure 2** shows that the baseline models do not reproduce intra-subject citations. Instead, citation densities are evenly spread across all subjects. As expected, this also applies to the softmax-normalized subject models (except Locality). The softmax function's smoothing effect is particularly significant here because we use features — such as authority score and cosine similarities — that range within $[0, 1]$. Much of the potentially differentiating information captured by these variables are expressed in terms of small decimal differences which are easily smoothened away. Therefore, except when used in the context of the Locality model (which would impose strict subject constraints to begin with), softmax normalization appears unsuited for our purposes.

The remaining subject models have largely similar subject signatures. To further distinguish them, we examine specific graph properties presented in **Tables 1** and **2**. While our subject models are similar to baselines in some respects like connectivity, they have clear structural differences, particularly in terms of subject communities.

General Properties: Clustering coefficients were generally low (save in the BEA and Aging (softmax) models). All models produced giant components encompassing most (around 99.9%) of the graph. However, in-degree distributions in baselines are generally more imbalanced than in the subject models. Gini coefficients for the baselines ranged between 0.52 (relay) to 0.998 (Aging (sum)) whereas those for the subject models ranged between 0.417 (Linear (sum)) to 0.598 (TSPR (sum)).

Subject Structure: Relative to all others, sum-normalized subject models yielded more intra-subject edges, lower expansion and conductance scores for communities defined by

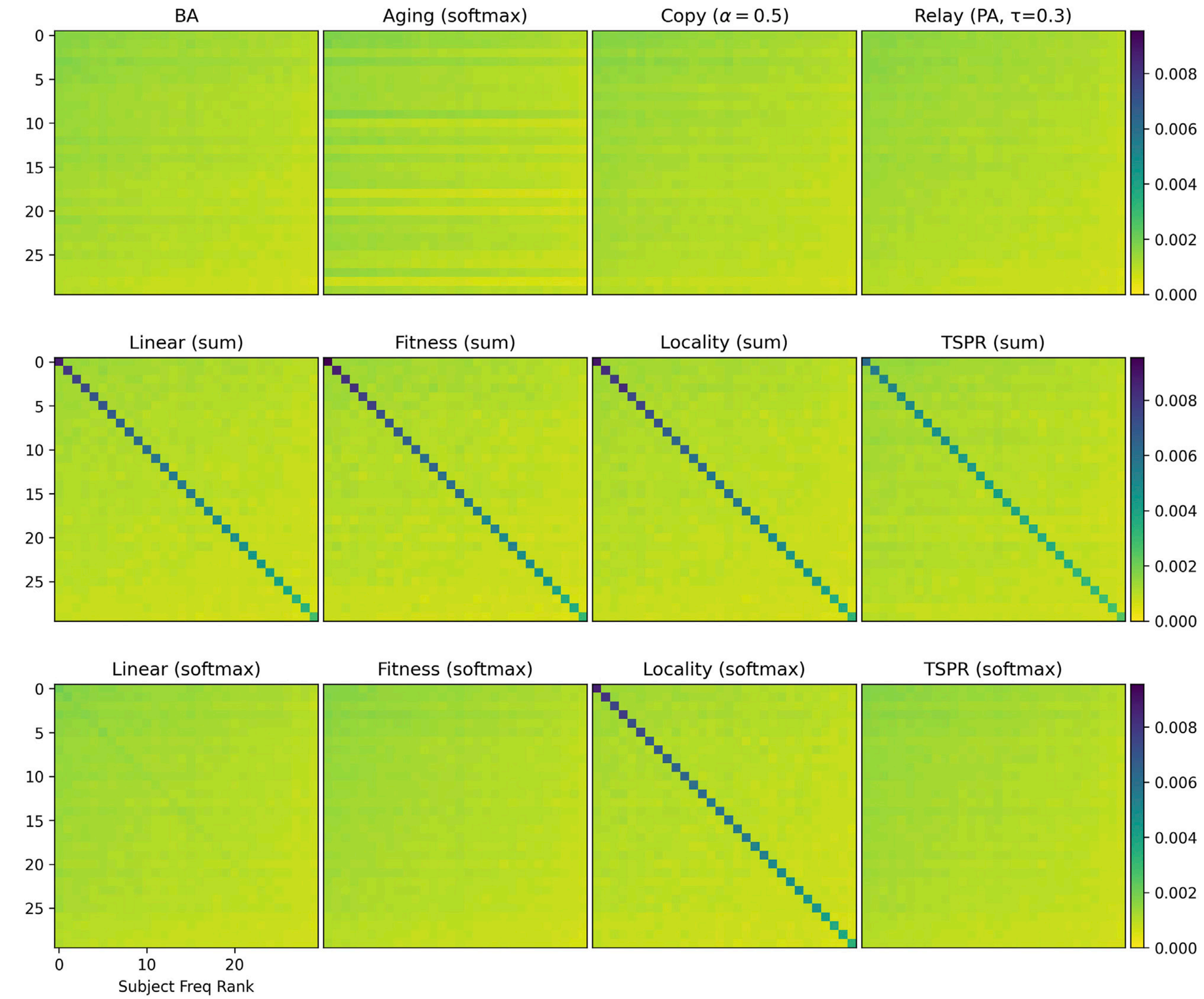


FIGURE 2 | H_{norm} subject signatures for simulated baselines **(top row)**, sum-normalized subject models **(middle row)**, and softmax-normalized subject models **(bottom row)**. Subjects are indexed by descending frequency from 0 to 30.

the gold labels, as well as higher link modularities. These networks therefore exhibit stronger within-subject clustering (an attribute which, to recall, should in theory characterize CLCNs). Conversely, the softmaxed subject models differed only slightly from the baselines. To be sure, stronger conformity with subject labels is not *necessarily* better, since legal citations are influenced by more than subject relevance alone. Depending on the extent of subject clustering desired, therefore, TSPR approaches may offer a middle-ground.

k-Clique Structure: The k-clique metrics paint a less coherent picture. There is no clear correlation between how well the models retrace gold label subjects and the average number, size, and quality of the communities recovered by k-clique percolation. The average number (size) of communities uncovered amongst sum-normalized subject models varies from around 28 to 82 (7–18) whereas the softmaxed subject models tend to yield around 70 k-clique communities of 5–8 nodes. Insofar as our models approximate the legal citation process, these results suggest that k-clique percolation may be less useful for clustering (and classifying) case law by legal subjects. Communities recovered by the algorithm do not appear to reflect actual legal areas, suggesting that legal subject clustering does not follow the assumptions of k-clique percolation. Nonetheless, the observed k-clique communities may be the result of other clustering mechanics inherent in legal citation networks. To this extent, they offer an independent basis for assessing the structural similarity of different simulated models. For instance, it is clear from **Table 2** that the BEA model, which always results in exactly 1 large k-clique community that encapsulates the whole network, is structurally distinct from the rest. The sum-normalized Fitness and TSPR models

TABLE 1 | Simulated properties for sum-normalized and relay models.

	BA	Aging	Relay	Linear	Fitness	Locality	TSPR
General Structure							
Avg Clust Coef	0.043 (0.0)	0.0 (0.0)	0.018 (0.0)	0.059 (0.0)	0.086 (0.01)	0.058 (0.0)	0.07 (0.0)
Giant Comp %	0.994 (0.0)	0.994 (0.0)	0.994 (0.0)	1.0 (0.0)	0.994 (0.0)	0.999 (0.0)	0.999 (0.0)
Gini (In-Deg)	0.739 (0.01)	0.998 (0.0)	0.52 (0.01)	0.417 (0.01)	0.597 (0.01)	0.456 (0.01)	0.598 (0.01)
Subject Structure							
#Intra Edges	420.08 (29.29)	83.84 (33.3)	405.86 (25.59)	2228.54 (52.4)	2507.18 (60.4)	2449.08 (59.06)	1740.94 (46.04)
Expansion	4.559 (0.11)	0.917 (0.03)	4.435 (0.11)	2.981 (0.08)	2.594 (0.07)	2.626 (0.08)	3.388 (0.09)
Conductance	0.865 (0.01)	0.923 (0.03)	0.866 (0.01)	0.437 (0.01)	0.38 (0.01)	0.387 (0.01)	0.528 (0.01)
Link Modularity	0.005 (0.0)	0.005 (0.0)	0.005 (0.0)	0.031 (0.0)	0.035 (0.0)	0.034 (0.0)	0.023 (0.0)
k-Clique Structure							
#Comms Recov'd	32.66 (7.64)	0.0 (0.0)	84.74 (10.69)	72.36 (9.53)	28.14 (5.45)	81.92 (10.15)	52.48 (6.4)
Avg Comm Size	13.485 (2.92)	–	5.205 (0.38)	8.929 (0.98)	18.058 (3.09)	7.661 (0.81)	10.02 (1.04)
#Intra Edges	1819.5 (94.97)	–	983.42 (85.87)	1915.8 (124.49)	2107.3 (89.04)	1740.84 (116.24)	1890.7 (103.13)
Expansion	4.563 (0.23)	–	4.911 (0.19)	3.864 (0.14)	4.097 (0.27)	4.01 (0.15)	4.078 (0.2)
Conductance	0.663 (0.01)	–	0.68 (0.01)	0.598 (0.01)	0.62 (0.02)	0.617 (0.01)	0.628 (0.01)
Link Modularity	0.044 (0.0)	–	0.019 (0.0)	0.042 (0.0)	0.056 (0.0)	0.038 (0.0)	0.047 (0.0)

Notes: Values represent the mean (standard deviation) of the relevant metric over 50 iterations. Relay was simulated with $\theta = 0.9$, $\tau = -0.3$, relay depth 1, and preferential relay. Relay is not *a sum-normalized model but because the relay mechanism relies on sum-normalized preferential attachment, it is expedient to present its properties here instead of in a separate table.*

TABLE 2 | Simulated properties for softmax-normalized and copying models.

	BEA	Aging	Copy	Linear	Fitness	Locality	TSPR
General Structure							
Avg Clust Coef	0.469 (0.01)	0.341 (0.02)	0.044 (0.0)	0.015 (0.0)	0.015 (0.0)	0.051 (0.0)	0.016 (0.0)
Giant Comp %	0.994 (0.0)	0.997 (0.0)	0.999 (0.0)	0.999 (0.0)	0.999 (0.0)	0.999 (0.0)	0.999 (0.0)
Gini (In-Deg)	0.988 (0.0)	0.905 (0.01)	0.719 (0.01)	0.544 (0.01)	0.548 (0.01)	0.564 (0.01)	0.549 (0.01)
Subject Structure							
#Intra Edges	417.82 (64.92)	411.02 (48.24)	423.82 (24.52)	476.0 (25.78)	420.42 (25.01)	2448.34 (58.85)	424.32 (26.85)
Expansion	4.561 (0.14)	4.565 (0.12)	4.557 (0.11)	4.511 (0.1)	4.559 (0.1)	2.642 (0.08)	4.555 (0.1)
Conductance	0.878 (0.02)	0.874 (0.01)	0.863 (0.01)	0.847 (0.01)	0.864 (0.01)	0.385 (0.01)	0.862 (0.01)
Link Modularity	0.005 (0.0)	0.005 (0.0)	0.005 (0.0)	0.006 (0.0)	0.005 (0.0)	0.034 (0.0)	0.005 (0.0)
k-Clique Structure							
#Comms Recov'd	1.0 (0.0)	1.32 (0.87)	27.94 (5.58)	77.34 (8.28)	76.52 (8.07)	62.96 (7.16)	77.46 (8.91)
Avg Comm Size	480.68 (4.38)	419.163 (109.05)	14.826 (2.86)	5.245 (0.27)	5.293 (0.31)	8.565 (0.89)	5.257 (0.35)
#Intra Edges	2452.9 (56.38)	2388.24 (61.4)	1680.34 (95.53)	986.3 (81.47)	1000.64 (84.98)	1742.48 (110.98)	997.5 (83.89)
Expansion	0.0 (0.0)	0.353 (0.74)	4.035 (0.27)	4.502 (0.18)	4.476 (0.16)	4.027 (0.2)	4.469 (0.15)
Conductance	0.0 (0.0)	0.064 (0.13)	0.629 (0.02)	0.663 (0.01)	0.662 (0.01)	0.629 (0.01)	0.661 (0.01)
Link Modularity	0.072 (0.0)	0.07 (0.0)	0.042 (0.0)	0.019 (0.0)	0.02 (0.0)	0.04 (0.0)	0.02 (0.0)

Notes: Values represent the mean (standard deviation) of the relevant metric over 50 iterations. Copy was simulated with $\alpha = 0.5$. Note that Copy is not a softmax-normalized model but is tabulated here for brevity.

also stand out from the other sum-normalized subject models as they tend to produce fewer but larger k-clique communities (see **Table 1**).

In sum, preliminary simulations demonstrate that the specific approach used to incorporate subject relevance induces significantly different network structures. While using subject cosine similarities as fitness values in a sum-normalized model may be *theoretically* similar to setting subject constraints on citable localities, the resulting networks differ in key aspects such as in-degree distribution and k-clique structure. The normalization function used also makes a key difference. This in turn underscores the importance of carefully selecting the subject model. Here we do not declare any one model as correct or best. To the extent that an ideal model exists, its parameterization likely turns on the specific type of CLCN we are trying to simulate. The legal and institutional context underlying the network would be relevant. Citation practices in one court at a given time may fall closer to the Fitness model, whereas another court may cite more in line with the TSPR model.

That said, it is also clear that certain subject models, especially those using softmax normalization, are unlikely to successfully capture the nuances of legal citations for the reasons given above. For benchmarking purposes, therefore, we focus on approaches that appear more promising, being Linear (sum), Fitness (sum), Locality (softmax), and TSPR (sum). For brevity, we only present results for more promising baselines as well (being BA, Copying, and Relay).[10]

[10]Results for the other baselines and varying paramaterizations of Copying and Relay(on file) do not affect our conclusions.

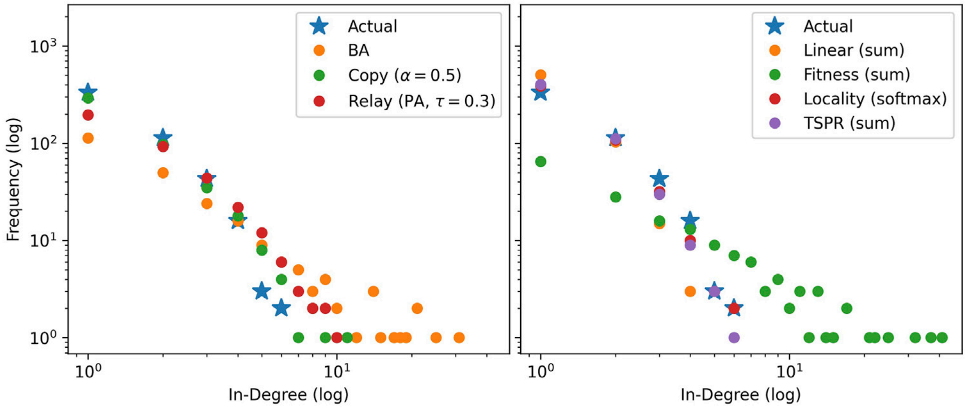


FIGURE 3 | Mode-averaged in-degree distributions for USSC-calibrated baselines **(left)** and subject models **(right)** relative to the actual. Mirroring the actual network, 50 iterations of 2000 node steps are run per model. Models in the same iteration use identical, pre-drawn out-degree counts from $|E_n| \sim Po(0.388)$ and subjects from $\phi \sim Dir(\psi_{us}, 14)$, mirroring the USSC network's actual properties.

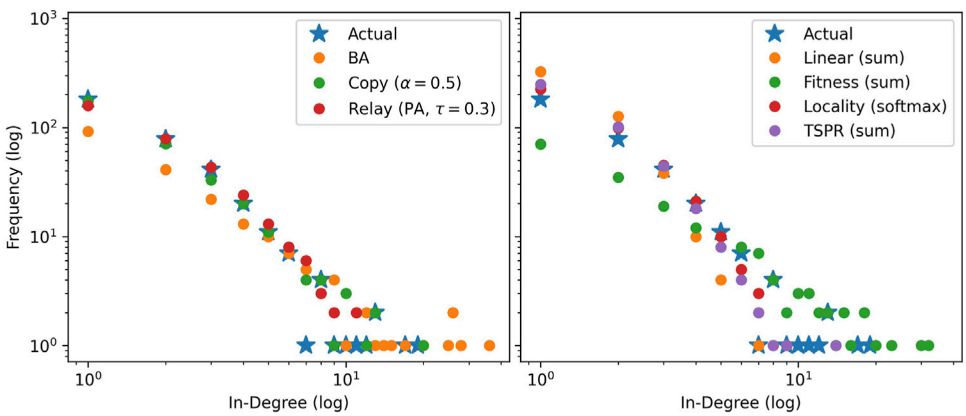


FIGURE 4 | Mode-averaged in-degree distributions for SGCA-calibrated baselines **(left)** and subject models **(right)** relative to the actual. Mirroring the actual network, 50 iterations of 1000 node steps are run per model. Models in the same iteration use identical, pre-drawn out-degree counts from $|E_n| \sim Po(0.778)$ and subjects from $\phi \sim Dir(\psi_{sg}, 31)$.

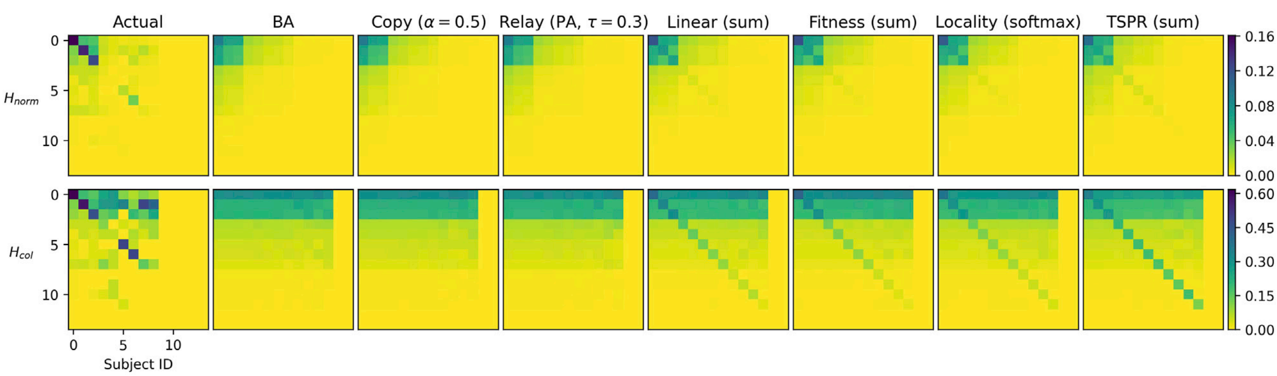


FIGURE 5 | H_{norm} **(top row)** and H_{col} **(bottom row)** subject signatures for USSC-calibrated baselines and selected subject models. Subjects are indexed by descending frequency from 0 to 14. Note that color scales differ across rows.

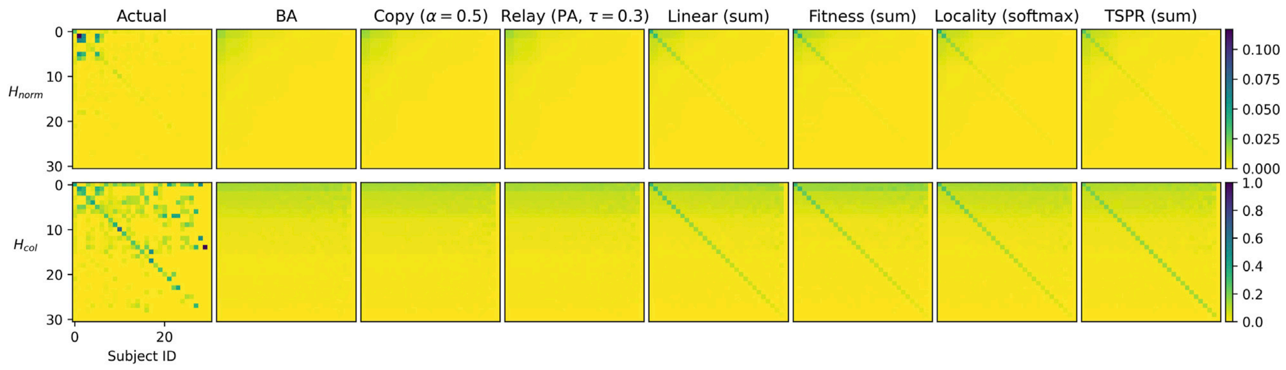


FIGURE 6 | H_{norm} **(top row)** and H_{col} **(bottom row)** subject signatures for SGCA-calibrated baselines and selected subject models. Subjects are indexed by descending frequency from 0 to 31. Note that color scales differ across rows.

3.2 Empirical Benchmarks

Figures 3 and **4** compare in-degree distributions produced by the calibrated baseline and subject models against actual distributions from the USSC and SGCA benchmarks respectively. Both benchmark distributions are broadly reproduced by all models (including baselines), though BA and Fitness (sum) produce slightly gentler-sloping distributions in both cases. At the same time, **Figures 5** and **6** show that only the subject models successfully reproduce empirical subject signatures. To see this, first notice that since subject distributions for both the USSC and SGCA are imbalanced, citation densities are naturally concentrated within the top left quarter (which indexes the most frequent subjects). Second, observe that main diagonals for both benchmarks are also denser than their off-diagonals, showing that subject overlap materially influences actual legal citations. Although the general top-left concentration is reproduced by all models (recall that simulated subjects are drawn from a Dirichlet parameterized by actual subject frequencies), only subject models exhibit the benchmarks' distinctive signature.

These observations apply to both H_{norm} and H_{col} signatures. As expected, H_{col} densities for both the actual and simulated graphs are concentrated upwards across all rows, reflecting how nodes with the most frequent subjects are relatively likely to be cited, regardless of citing subject. This may be explained by subject frequency imbalance. To illustrate, most nodes, by definition, will have subject 0. Assuming node i has subjects $\{0, 1\}$ and node j has $\{0, 30\}$, j may cite i under the subject models due to the shared subject 0. The edge (j, i) would count not only towards $H(0, 0)$ but $H(30, 0)$ and $H(30, 1)$ as well.

Downward density gradients for the benchmarks' signatures are noticeably less smooth than for the simulations. This is clearest for the SGCA benchmark, where (within the top 15 subjects) some less frequent subjects are more likely to be cited than the most frequent. This is suggestive of specific correlations between legal subjects. For example, subject 15 may be legally very relevant to and therefore often cited by subject 5 even though the former rarely occurs. Our models do not currently account for such subject covariance and future work could explore this further.

Figure 7 tabulates L1 distances between the benchmark and simulated networks and provides further confirmation that the subject models offer a closer approximation of CLCNs than baselines. Total, diagonal, off-diagonal, and column-wise distances are consistently smaller for the subject models than for the baselines. Note that results are similar if we include other (previously discussed) baselines, including alternative paramaterizations of Copying and Relay, and also if we use Kullbeck-Leibler divergence instead of L1 distance.[11] Surprisingly, none of the subject models are clearly superior to the others. All clock in comparable numbers for every metric.

Therefore, to determine which models produce better empirical fit, we look into detailed network properties for each benchmark as presented in **Tables 3** and **4**. Results here are consistent with the preliminary simulations. Clustering coefficients and giant component percentages for all models are broadly similar, but in-degree distributions vary across models.

As expected, subject models yield networks with stronger subject communities than all baselines. The subject models fit the SGCA benchmark well, each producing around 800 intra-subject edges compared to the benchmark's 889. They also yield expansion, conductance, and link modularity scores indicating better subject community quality. However, the community quality metrics for the benchmark's network are *even* better, suggesting that our models can place even more weight on subject overlap.

Fit for the USSC benchmark is less clear. The USSC network has relatively few intra-subject edges. Resultantly, the baselines fall closer to the benchmark on this metric. However, community quality for the subject models are significantly closer to the actual. Re-creating the USSC network may therefore require model paramaterizations which create smaller but even more distinctive communities.

In terms of k-clique structure, the model closest to both benchmarks is Fitness (sum). The model produces on average 13.62 communities of 3.85 nodes (compared to 24 and 4.042) and

[11] Though Kullbeck-Leibler returns infinity on some graphs.

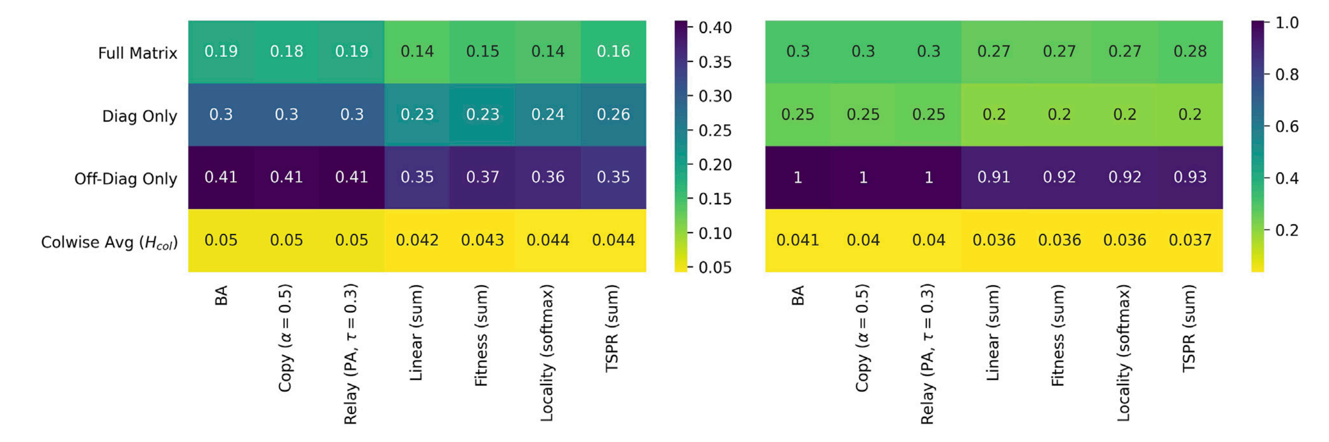


FIGURE 7 | L1 distances for USSC-calibrated **(left)** and SGCA-calibrated models **(right)** relative to their respective benchmarks. Distances in the top three rows are computed from H_{norm} while the column-wise average L1 is computed from H_{col}(though for L1 distances computing this from H_{norm} yields the same result).

TABLE 3 | Actual versus simulated properties for the USSC network.

	Actual	BA	Copy	Relay	Linear (sum)	Fitness (sum)	Locality (softmax)	TSPR (sum)
General Structure								
Avg Clust Coef	0.009	0.001 (0.0)	0.0 (0.0)	0.0(0.0)	0.0(0.0)	**0.002(0.0)**	0.0(0.0)	0.001(0.0)
Giant Comp %	0.259	0.32(0.01)	0.041(0.02)	0.32(0.01)	0.025(0.01)	**0.319(0.01)**	0.101(0.04)	0.037(0.02)
Gini(In-Deg)	0.811	0.94(0.0)	0.841(0.01)	0.876(0.01)	0.736(0.01)	0.962(0.0)	**0.792(0.01)**	0.785(0.01)
Subject Structure								
#Intra Edges	384	623.76(42.38)	**618.4(31.41)**	618.54(35.13)	956.0(42.03)	964.0(46.54)	909.98(38.74)	986.66(43.5)
Expansion	0.206	0.325(0.03)	0.325(0.03)	0.323(0.03)	0.263(0.03)	0.266(0.02)	0.27(0.03)	**0.206(0.03)**
Conductance	0.356	0.764(0.02)	0.759(0.02)	0.758(0.02)	0.562(0.04)	0.585(0.04)	0.585(0.04)	**0.382(0.04)**
Link Modularity	0.036	**0.028(0.0)**	**0.028(0.0)**	**0.028(0.0)**	0.046(0.0)	0.045(0.0)	0.043(0.0)	0.044(0.0)
k-Clique Structure								
#Comms Recov'd	24	6.12(2.15)	0.14(0.5)	2.66(1.62)	1.0(0.9)	**13.62(2.93)**	0.52(0.68)	1.4(1.07)
Avg Comm Size	4.042	3.429(0.47)	3.0(0.0)	3.204(0.5)	3.095(0.26)	**3.852(0.3)**	3.091(0.29)	3.009(0.05)
#Intra Edges	124	23.26(8.63)	4.2(2.68)	9.082(5.34)	4.6(2.51)	**63.8(13.67)**	3.727(1.55)	5.73(2.42)
Expansion	1.925	**0.596(0.15)**	0.267(0.28)	0.591(0.22)	0.252(0.3)	0.468(0.13)	0.321(0.35)	0.296(0.24)
Conductance	0.378	0.207(0.04)	0.107(0.11)	**0.213(0.06)**	0.095(0.1)	0.162(0.04)	0.12(0.12)	0.114(0.08)
Link Modularity	0.011	0.002(0.0)	0.0(0.0)	0.001(0.0)	0.0(0.0)	**0.004(0.0)**	0.0(0.0)	0.001(0.0)

Notes: Except in the "actual" column, values are the mean(standard deviation) of the relevant statistic over 50 simulations per model. For the actual network properties the case issue areas assigned by the Spaeth database are used as subject labels. k-Clique expansion, conductance, and link modularity scores are only computed for and averaged within simulations which return at least 1 k-clique community and should be interpreted in this light. 45, 15, 28, and 13 iterations of the Copy, Linear, Locality, and TSPR models respectively returned 0 k-clique communities. All other models tabulated always returned at least 1. In each row, results numerically closest to the benchmark are bolded.

20.52 communities of size 4.452 (compared to 45 and 4.978) for the actual SGCA network). All other models yield significantly fewer k-clique communities. Indeed, 15 iterations of Linear (sum), 28 iterations of Locality (softmax), and 13 iterations of TSPR produced 0 k-clique communities. Conversely, Fitness (sum) always returns at least one k-clique community (including when calibrated with the SGCA network).

4 DISCUSSION

The mechanics of case law citations involve complex interactions between the legal authority of a case, its relevance to the citer's subjects, as well as the case's age. We may represent the case law citation function in abstract as a probability distribution $P = f(\Lambda, \rho, w)$, where each input variable captures each of these attributes respectively. Determining the exact functional form to be used, however, is difficult. Numerous reasonable approaches can be imagined. For subject relevance alone, we experimented with using subject similarities as linear features, fitness values, to constrain citable horizons by subject, and to use subject-sensitive centrality scores. These all draw on existing literature on citation networks, but other approaches may be possible.

In this light, this paper studied by simulation the expected properties of networks generated by the four approaches above and compared them against established network models. We then compared more promising models against two actual CLCNs

TABLE 4 | Actual versus simulated properties for the SGCA network.

	Actual	BA	Copy	Relay	Linear(sum)	Fitness(sum)	Locality(softmax)	TSPR(sum)
General Structure								
Avg Clust Coef	0.044	**0.005(0.0)**	0.001(0.0)	0.002(0.0)	0.003(0.0)	0.009(0.0)	0.002(0.0)	**0.005(0.0)**
Giant Comp %	0.347	0.535(0.02)	0.557(0.03)	0.535(0.02)	0.572(0.04)	**0.534(0.02)**	0.584(0.03)	0.589(0.03)
Gini(In-Deg)	0.789	0.903(0.0)	0.804(0.01)	**0.789(0.01)**	0.624(0.01)	0.911(0.01)	0.721(0.01)	0.71(0.01)
Subject Structure								
#Intra Edges	889	259.96(24.28)	260.7(19.54)	257.22(21.85)	815.82(32.9)	**831.18(41.47)**	813.04(31.06)	768.5(27.58)
Expansion	0.332	0.712(0.03)	0.713(0.03)	0.709(0.03)	0.534(0.03)	0.531(0.03)	0.533(0.03)	**0.479(0.03)**
Conductance	0.449	0.87(0.01)	0.869(0.02)	0.872(0.02)	0.564(0.02)	0.577(0.03)	0.563(0.03)	**0.462(0.03)**
Link Modularity	0.047	0.011(0.0)	0.011(0.0)	0.011(0.0)	**0.037(0.0)**	**0.037(0.0)**	0.036(0.0)	0.033(0.0)
k-Clique Structure								
#Comms Recov'd	45	11.98(3.03)	2.06(1.46)	7.08(2.66)	6.14(2.15)	**20.52(4.19)**	4.74(2.08)	9.86(2.85)
Avg Comm Size	4.978	4.482(0.93)	3.659(1.5)	3.233(0.25)	3.448(0.42)	**4.457(0.64)**	3.172(0.22)	3.167(0.21)
#Intra Edges	386	69.44(15.58)	9.545(6.04)	24.56(9.53)	23.64(8.11)	**122.0(24.07)**	15.88(7.24)	33.12(10.63)
Expansion	1.139	0.922(0.15)	0.391(0.24)	**1.079(0.18)**	0.63(0.19)	0.78(0.15)	0.61(0.24)	0.519(0.14)
Conductance	0.273	**0.283(0.03)**	0.14(0.08)	0.323(0.04)	0.21(0.05)	0.242(0.04)	0.207(0.07)	0.183(0.04)
Link Modularity	0.031	0.005(0.0)	0.001(0.0)	0.002(0.0)	0.002(0.0)	**0.009(0.0)**	0.001(0.0)	0.003(0.0)

Notes: Except in the "actual" column, values are the mean(standard deviation) of the relevant statistic over 50 simulations per model. For the actual network properties, case catchwords assigned by the Singapore Law Reports are used as subject labels. k-Clique expansion, conductance, and link modularity scores are only computed for and averaged within simulations which return at least 1 k-clique community and should be interpreted in this light. 6 iterations of the Copy model returned 0 k-clique communities. All other models tabulated always returned at least 1. In each row, results numerically closest to the benchmark are bolded.

from the USSC and SGCA respectively. Our findings underscore the importance of a legally-informed model of link generation process. The properties of all proposed subject models were substantially different from those of the baseline models, (provided that sum-normalization was used), and closer to empirical benchmarks in terms of both graph properties and subject signature.

Amongst the range of approaches studied, we found softmax normalization generally unsuited for the task because it smoothens away distinguishing differences in case attributes. We also see that model properties vary substantially when alternative means of modelling subject relevance were tested. All subject models performed comparably in terms of subject signature distance from both benchmarks. However, while three of the four subject models studied yielded average in-degree distributions very close to both benchmarks, the Fitness (sum) model yielded a noticeably gentler degree-rank slope. Nonetheless, Fitness (sum) best fits the k-clique structure of both benchmarks.

The emergent picture is that all subject models provide plausible (and superior to baseline) approaches for modelling CLCNs. For our two benchmarks, however, Fitness (sum) slightly edges out the other approaches as the most promising method for modelling CLCNs. To recall, this model takes the cosine similarity between the subject belonging vectors of two cases as a fitness coefficient that modifies the weight computed from each case's authority and age. Notice that this model is similar to the General Temporal model in many respects, and may be seen as an extension of that model tailored to the legal citation context. An important caveat here is that although Fitness (sum) is *relatively* better when compared to alternative models, it is not in *absolute* terms a perfect approximation of either benchmark. Key differences remain between the actual and fitness-simulated networks. Further, alternative models may be better suited for CLCNs from other courts.

Our findings hint at possible universality in terms of the way courts think about subject relevance when deciding which cases to cite, though we hasten to add that the two benchmarks ran do not offer sufficient evidence. Universal or otherwise, the alternative models proposed may be useful for examining how courts and possibly individual judges differ when selecting cases to cite. If court/judge A's citation network is better approximated by model X while court/judge B's network is better approximated by model Y. Our models are also helpful for generating better simulated data that may be used to research other questions in legal network science. For instance, researchers could benchmark centrality algorithms against networks simulated with these models to study how far these algorithms recover legally-significant nodes. This, to recall **Section 1**, is a rich area of legal networks research. Since the data is simulated, it becomes possible to specifically dictate which nodes are legally-significant.

5 LIMITATIONS AND FUTURE WORK

This paper is the first to study how the unique mechanics of case law citations may be simulated and studied using network models. As a first step, it necessarily leaves a number of important questions unexplored.

First, although we have identified promising approaches for modelling CLCNs, alternative models for legal authority, subject relevance, and time decay remain to be studied. We have focused on comparing different methods for modelling subject relevance because this is the least explored question. Nonetheless, the effect of varying authority and time-decay models are worth studying further. Varying time-decay models in particular may yield insights on how quickly the value of precedent depreciates (a question often raised by legal scholars [8, 39]), as well as how much deference different courts accord to antiquated precedent.

Second, future work can explore a larger space of model parameterizations. An immediate extension would be using time-variant network growth rates. That is, both λ and μ may vary across time. Further, the feature weights α, β, and γ may be adjusted to generate and also reflect different judicial attitudes towards assessing authority and relevance. It may be possible to learn these weights from empirical data, whether using exponential random graph models or machine learning techniques. This would provide a means of quantitatively measuring which factors most influence legal citation decisions, providing a common metric for comparing how these differ (or remain the same) across judges, time, and space.

Third, this study was limited by data availability. Despite growing literature in case law citations analysis, few publicly available edgelists can be linked to case-level (subject) metadata. For now, we have benchmarked our models against two empirical networks produced by apex Common Law courts. Future work can consider how closely these models approximate the citation mechanics of courts in other jurisdictions, particularly those of Civil Law jurisdictions where the doctrine of precedent theoretically holds less sway.

Fourth, while we have focused primarily on network structure, the microscopic properties of our proposed networks remain largely unexplored. Future work could use node-level metrics such as centrality and accessibility to study what kinds of cases and subjects are most likely to become entrenched in the core of such networks (e.g. [64]). Additionally, the task of replicating and studying CLCNs would also benefit from also exploiting the textual content of case judgments, as was done in [35]. Methods that exploit both network structure and node attributes for community detection (e.g. [65]) could be explored. This would connect our work to the growing literature on legal language processing (e.g. [66, 67]).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

JS was in charge of all coding, data analysis, and writing.

ACKNOWLEDGMENTS

We thank Daniel Katz, Michael Bommarito, Ryan Whalen, and Kenny Chng for insightful discussions on the topic, participants of the Physics of Law conference for valuable input, the editors of the special collection on the Physics of Law for inspiring this research, and three reviewers for constructive input. We thank also our loved ones for supporting our research unconditionally.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fphy.2021.665563/full#supplementary-material

REFERENCES

1. Clough JR, Gollings J, Loach TV, and Evans TS. Transitive Reduction of Citation Networks. *J Complex Networks* (2015) 3:189–203. doi:10.1093/comnet/cnu039
2. Fowler JH, Johnson TR, Spriggs JF, Jeon S, and Wahlbeck PJ. Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Polit Anal* (2007) 15:324–46. doi:10.1093/pan/mpm011
3. Fowler JH, and Jeon S. The Authority of Supreme Court Precedent. *Soc. Networks* (2008) 30:16–30. doi:10.1016/j.socnet.2007.05.001
4. Leibon G, Livermore M, Harder R, Riddell A, and Rockmore D. Bending the Law: Geometric Tools for Quantifying Influence in the Multinetwork of Legal Opinions. *Artif Intell L* (2018) 26:145–67. doi:10.1007/s10506-018-9224-2
5. Derlén M, and Lindholm J. Characteristics of Precedent: The Case Law of the European Court of Justice in Three Dimensions. *Ger Law J* (2015) 16:1073–98. doi:10.1017/S2071832200021040
6. Lupu Y, and Voeten E. Precedent in International Courts: A Network Analysis of Case Citations by the European Court of Human Rights. *Br J. Polit. Sci.* (2012) 42:413–39. doi:10.1017/S0007123411000433
7. Soh J. A Network Analysis of the Singapore Court of Appeal's Citations to Precedent. *SSRN J* (2019) 31:246–84. doi:10.2139/ssrn.3346422
8. Whalen R, Uzzi B, and Mukherjee S. Common Law Evolution and Judicial Impact in the Age of Information. *Elon L Rev* (2017) 9:115–70.
9. Bommarito MJ, II, and Katz DM. *Properties of the United States Code Citation Network* (2010). Available from: https://https://arxiv.org/abs/0911.1751 (Accessed 09 July, 2021).
10. Katz DM, Coupette C, Beckedorf J, and Hartung D. Complex Societies and the Growth of the Law. *Sci Rep* (2020) 10:18737, 2020 . ArXiv: 2005.07646. doi:10.1038/s41598-020-73623-x
11. Coupette C, Beckedorf J, Hartung D, Bommarito M, and Katz DM. Measuring Law over Time: A Network Analytical Framework with an Application to Statutes and Regulations in the United States and Germany. *Front Physics* (2021) 9:269. doi:10.3389/fphy.2021.658463
12. Watts D, and Strogatz S. Collective Dynamics of 'Small-World' Networks. *Nature* (1998) 393:3. doi:10.1038/30918
13. Girvan M, and Newman MEJ. Community Structure in Social and Biological Networks. *Proc Natl Acad Sci* (2002) 99:7821–6. Institution: National Academy of Sciences Publisher: National Academy of Sciences Section: Physical Sciences. doi:10.1073/pnas.122653799
14. Porter MA, Onnela JP, and Mucha PJ. Communities in Networks. *Notices Am Math Soc* (2009) 56:1082–166.
15. Holland PW, Laskey KB, and Leinhardt S. Stochastic Blockmodels: First Steps. *Soc Networks* (1983) 5:109–37. doi:10.1016/0378-8733(83)90021-7
16. Karrer B, and Newman MEJ. Stochastic Blockmodels and Community Structure in Networks. *Phys Rev E* (2011) 83:016107. doi:10.1103/PhysRevE.83.016107
17. Newman MEJ. Fast Algorithm for Detecting Community Structure in Networks. *Phys Rev E* (2004) 69:066133. doi:10.1103/PhysRevE.69.066133
18. Newman MEJ, and Girvan M. Finding and Evaluating Community Structure in Networks. *Phys Rev E* (2004) 69:026113. doi:10.1103/PhysRevE.69.026113
19. Clauset A, Newman MEJ, and Moore C. Finding Community Structure in Very Large Networks. *Phys Rev E* (2004) 70:066111. doi:10.1103/PhysRevE.70.066111
20. Derényi I, Palla G, and Vicsek T. Clique Percolation in Random Networks. *Phys Rev Lett* (2005) 94:160202. doi:10.1103/PhysRevLett.94.160202

21. Newman MEJ. Finding Community Structure in Networks Using the Eigenvectors of Matrices. *Phys Rev E* (2006) 74:036104. doi:10.1103/PhysRevE.74.036104
22. Nicosia V, Mangioni G, Carchiolo V, and Malgeri M. Extending the Definition of Modularity to Directed Graphs with Overlapping Communities. *J Stat Mech* (2009) 2009:P03024. doi:10.1088/1742-5468/2009/03/P03024
23. Mucha PJ, Richardson T, Macon K, Porter MA, and Onnela J-P. Community Structure in Time-dependent, Multiscale, and Multiplex Networks. *Science* (2010) 328:876–8. doi:10.1126/science.1184819
24. Chen M, and Szymanski BK. Fuzzy Overlapping Community Quality Metrics. *Soc Netw Anal Min* (2015) 5:40. doi:10.1007/s13278-015-0279-8
25. Bommarito MJ, Katz DM, Zelner JL, and Fowler JH. Distance Measures for Dynamic Citation Networks. *Physica A: Stat Mech its Appl* (2010) 389:4201–8. doi:10.1016/j.physa.2010.06.003
26. Mirshahvalad A, Lindholm J, Derlén M, and Rosvall M. Significant Communities in Large Sparse Networks. *PLoS One* (2012) 7:e33721. Public Library of Science. doi:10.1371/journal.pone.0033721
27. Sulea OM, Zampieri M, Malmasi S, Vela M, Dinu LP, and van Genabith J. Exploring the Use of Text Classification in the Legal Domain (2017) Minneapolis, Minnesota, 2-7 June 2019 CoRR abs/1710.09306.
28. Soh J, Lim HK, and Chai IE. Legal Area Classification: A Comparative Study of Text Classifiers on Singapore Supreme Court Judgments. In: Proceedings of the Natural Legal Language Processing Workshop 2019. Minneapolis, Minnesota: Association for Computational Linguistics (2019). 67–77. doi:10.18653/v1/W19-2208
29. de Solla Price DJ. Networks of Scientific Papers. *Science* (1965) 149:510–5. Institution: American Association for the Advancement of Science Section: Articles. doi:10.1126/science.149.3683.510
30. Broido AD, and Clauset A. Scale-free Networks Are Rare. *Nat Commun* (2019) 10(1):1017. Institution: Nature Publishing Group. doi:10.1038/s41467-019-08746-5
31. Wang M, Yu G, and Yu D. Effect of the Age of Papers on the Preferential Attachment in Citation Networks. *Physica A: Stat Mech its Appl* (2009) 388: 4273–6. doi:10.1016/j.physa.2009.05.008
32. Menczer F. Evolution of Document Networks. *Proc Natl Acad Sci* (2004) 101: 5261–5. doi:10.1073/pnas.0307554100
33. Singh M, Sarkar R, Goyal P, Mukherjee A, and Chakrabarti S. Relay-Linking Models for Prominence and Obsolescence in Evolving Networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, Nova Scotia, 13-17 August 2017 (2017). 1077–86. ArXiv: 1609.08371. doi:10.1145/3097983.3098146
34. Kumar R, Raghavan P, Rajagopalan S, Sivakumar D, Tomkins A, and Upfal E. Stochastic Models for the Web Graph. In: Proceedings 41st Annual Symposium on Foundations of Computer Science. 12-14 November 2000, Redondo Beach, CA: IEEE Comput. Soc (2000). 57–65. doi:10.1109/SFCS.2000.892065
35. Amancio DR, Oliveira ON, and da Fontoura Costa L. Three-feature Model to Reproduce the Topology of Citation Networks and the Effects from Authors' Visibility on Their H-index. *J Informetrics* (2012) 6:427–34. doi:10.1016/j.joi.2012.02.005
36. Bianconi G, and Barabási A-L. Competition and Multiscaling in Evolving Networks. *Europhys Lett* (2001) 54:436–42. doi:10.1209/epl/i2001-00260-6
37. Li X, and Chen G. A Local-World Evolving Network Model. *Physica A: Stat Mech its Appl* (2003) 328:274–86. doi:10.1016/S0378-4371(03)00604-6
38. Haveliwala TH. Topic-sensitive Pagerank. In: Proceedings of the 11th International Conference on World Wide Web, 7-11 May 2002, New York, NY, USA: Association for Computing Machinery (2002). 517–526. doi:10.1145/511446.511513
39. Posner R. An Economic Analysis of the Use of Citations in the Law. *Am L Econ Rev* (2000) 2:381–406. doi:10.1093/aler/2.2.381
40. Carmichael I, Wudel J, Kim M, and Jushchuk J. Examining the Evolution of Legal Precedent Through Citation Network Analysis. *N. C. Law Rev.* (2017) 96: 44. https://scholarship.law.unc.edu/nclr/vol96/iss1/6/
41. Barabasi AL, and Albert R. Emergence of Scaling in Random Networks. *Science* (1999) 286:4. doi:10.1126/science.286.5439.509
42. Medo M, Cimini G, and Gualdi S. Temporal Effects in the Growth of Networks. *Phys Rev Lett* (2011) 107:238701. doi:10.1103/PhysRevLett.107.238701
43. Ii MJB, Katz DM, and Zelner JL. On the Stability of Community Detection Algorithms on Longitudinal Citation Data. *Proced - Soc Behav Sci* (2010) 4: 26–37. doi:10.1016/j.sbspro.2010.07.480
44. Pham T, Sheridan P, and Shimodaira H. PAFit: A Statistical Method for Measuring Preferential Attachment in Temporal Complex Networks. *PLoS One* (2015) 10:e0137796. doi:10.1371/journal.pone.0137796
45. Hajra KB, and Sen P. Aging in Citation Networks. *Physica A: Stat Mech its Appl* (2005) 346:44–8. doi:10.1016/j.physa.2004.08.048
46. Cormode G, Shkapenyuk V, Srivastava D, and Xu B. Forward Decay: A Practical Time Decay Model for Streaming Systems. In: 2009 IEEE 25th International Conference on Data Engineering, 29 March-2 April 2009, Shanghai, China: IEEE (2009). 138–49. doi:10.1109/ICDE.2009.65
47. Caldarelli G, Capocci A, De Los Rios P, and Muñoz MA. Scale-Free Networks from Varying Vertex Intrinsic Fitness. *Phys Rev Lett* (2002) 89:258702. doi:10.1103/PhysRevLett.89.258702
48. Sun J, Staab S, and Karimi F. Decay of Relevance in Exponentially Growing Networks. In: Proceedings of the 10th ACM Conference on Web Science, 30 June-3 July 2019, Boston, MA (2018) 343–51. doi:10.1145/3201064.3201084
49. Mikolov T, Sutskever I, Chen K, Corrado GS, and Dean J. Distributed Representations of Words and Phrases and Their Compositionality. In: CJC Burges, L Bottou, M Welling, Z Ghahramani, and KQ Weinberger, editors. *Advances in Neural Information Processing Systems 26*. Curran Associates, Inc.) (2013). 3111–9.
50. Chen DL, Ash E, and Livermore MA. Case Vectors: Spatial Representations of the Law Using Document Embeddings. In: DN Rockmore, editor. *Law as Data: Computation, Text, & the Future of Legal Analysis*. Santa Fe: SFI Press (2019). HOLLIS number: 99153793511303941
51. Kleinberg JM. Authoritative Sources in a Hyperlinked Environment. *J ACM* (1999) 46:604–32. doi:10.1145/324133.324140
52. Cross FB, and Spriggs JFI. The Most Important(And Best) Supreme Court Opinions and Justices. *Emory Law J.* (2010) 60:407–502. https://scholarlycommons.law.emory.edu/elj/vol60/iss2/5/
53. Blei DM, Ng AY, and Jordan MI. Latent Dirichlet Allocation. *J Machine Learn Res* (2003) 3:993–1022. doi:10.5555/944919.944937
54. Bhattacharya P, Ghosh K, Pal A, and Ghosh S. *Methods for Computing Legal Document Similarity: A Comparative Study* . Madrid, Spain: 5th Workshop on Legal Data Analysis (2019).
55. Brin S, and Page L. *The PageRank Citation Ranking: Bringing Order to the Web* (1998). Available from: http://ilpubs.stanford.edu:8090/422/1/1999-66.pdf (Accessed 09 July, 2021).
56. Hagberg AA, Schult DA, and Swart PJ. Exploring Network Structure, Dynamics, and Function Using Networkx. In: G Varoquaux, T Vaught, and J Millman, editors. Proceedings of the 7th Python in Science Conference. Pasadena, CA: (2008). 11–5.
57. Radicchi F, Castellano C, Cecconi F, Loreto V, and Parisi D. Defining and Identifying Communities in Networks. *Proc Natl Acad Sci* 101. Institution: National Academy of Sciences Section: Physical Sciences (2004). 2658–63. doi:10.1073/pnas.0400054101
58. Palla G, Farkas IJ, Pollner P, Derényi I, and Vicsek T. Directed Network Modules. *New J Phys* (2007) 9:186. doi:10.1088/1367-2630/9/6/186
59. Rossetti G, Milli L, and Cazabet R. CDLIB: a python Library to Extract, Compare and Evaluate Communities from Complex Networks. *Appl Netw Sci* (2019) 4(1):1–26. Institution: Springer. doi:10.1007/s41109-019-0165-9
60. Spaeth HJ, Epstein L, Martin AD, Segal JA, Ruger TJ, and Benesh SC. *The Supreme Court Database*. Version 2017 Release 1. St Louis, Missouri: Publication Title: The Supreme Court Database (2017).
61. Donnat C, and Holmes S. Tracking Network Dynamics: A Survey Using Graph Distances. *Ann Appl Stat* (2018) 12:971–1012. doi:10.1214/18-AOAS1176
62. Hammond DK, Gur Y, and Johnson CR. Graph Diffusion Distance: A Difference Measure for Weighted Graphs Based on the Graph Laplacian Exponential Kernel. in: 2013 IEEE Global Conference on Signal and Information Processing, 3-5 December 2013, Austin, TX (2013). 419–22. doi:10.1109/GlobalSIP.2013.6736904
63. Xu Y, Salapaka SM, and Beck CL. A Distance Metric between Directed Weighted Graphs. In: 52nd IEEE Conference on Decision and Control, 10-

13 December 2013, Firenze: IEEE (2013). p. 6359–64. doi:10.1109/CDC.2013.6760895

64. Amancio DR, Nunes MGV, Oliveira ON, and Costa Ld. F. Extractive Summarization Using Complex Networks and Syntactic Dependency. *Physica A: Stat Mech its Appl* (2012) 391:1855–64. doi:10.1016/j.physa.2011.10.015
65. Yang J, McAuley J, and Leskovec J. Community Detection in Networks with Node Attributes. In: 2013 IEEE 13th International Conference on Data Mining, 7-10 December 2013, Dallas, TX (2013). 1151–6. doi:10.1109/ICDM.2013.167
66. Aletras N, Tsarapatsanis D, Preoţiuc-Pietro D, and Lampos V. Predicting Judicial Decisions of the European Court of Human Rights: A Natural Language Processing Perspective. *PeerJ Comput Sci* (2016) 2:e93. doi:10.7717/peerj-cs.93
67. Medvedeva M, Vols M, and Wieling M. Using Machine Learning to Predict Decisions of the European Court of Human Rights. *Artif Intell L* (2019) 28: 237–66. doi:10.1007/s10506-019-09255-y

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

ORIGINAL RESEARCH
published: 07 January 2022
doi: 10.3389/fphy.2021.695219

The Structure and Dynamics of Modern United States Federal Case Law

Keerthi Adusumilli[1], Bradford Brown[2], Joey Harrison[3], Matthew Koehler[1]*, Jason Kutarnia[4], Shaun Michel[1], Max Olivier[1], Craig Pfeifer[5], Zoryanna Slater[4], William Thompson[1], Dianna Vetter[6] and Renee Zacharowicz[7]

[1]Model-Based Analytics Departement, The MITRE Corporation, McLean, VA, United States, [2]Government Administration and Judiciary Division, The MITRE Corporation, McLean, VA, United States, [3]Simulation, Experimentation, and Gaming Department, The MITRE Corporation, McLean, VA, United States, [4]Data Science Department, The MITRE Corporation, McLean, VA, United States, [5]Human Language Technology Department, The MITRE Corporation, McLean, VA, United States, [6]IT and Architecture Engineering Department, The MITRE Corporation, McLean, VA, United States, [7]Open Health Services Department, The MITRE Corporation, McLean, VA, United States

OPEN ACCESS

Edited by:
*Pierpaolo Vivo,
King's College London,
United Kingdom*

Reviewed by:
*James Chen,
Michigan State University,
United States
Dirk Hartung,
Bucerius Law School, Germany*

***Correspondence:**
*Matthew Koehler
mkoehler@mitre.org*

Specialty section:
This article was submitted to Social Physics, a section of the journal Frontiers in Physics

Received: 14 April 2021
Accepted: 24 November 2021
Published: 07 January 2022

Citation:
Adusumilli K, Brown B, Harrison J, Koehler M, Kutarnia J, Michel S, Olivier M, Pfeifer C, Slater Z, Thompson W, Vetter D and Zacharowicz R (2022) The Structure and Dynamics of Modern United States Federal Case Law. Front. Phys. 9:695219. doi: 10.3389/fphy.2021.695219

The structure and dynamics of modern United States Federal Case Law are examined here. The analyses utilize large-scale network analysis tools, natural language processing techniques, and information theory to examine all the federal opinions in the Court Listener database, containing approximately 1.3 million judicial opinions and 11.4 million citations. The analyses are focused on modern United States Federal Case Law, as cases in the Court Listener database range from approximately 1926–2020 and include most Federal jurisdictions. We examine the data set from a structural perspective using the citation network, overall and by time and space (jurisdiction). In addition to citation structure, we examine the dataset from a topical and information theoretic perspective, again, overall and by time and space.

Keywords: judicial opinions, citation networks, topic models, information theory, law

INTRODUCTION

This paper examines the structure and emergent dynamics of United States Federal Case Law over time and across jurisdictions. The motivation of this work is twofold. First, we wish to demonstrate analysis of the legal corpus at-scale: to the best of our knowledge this is the largest study by case count to date. Second, we wish to conduct a sociolegal meta-analysis to explain the legal corpus' network properties and behavior as endemic to the legal system.

To justify the need for a big data approach to studying law, take, for example, the analysis of transcripts from the Old Bailey court of London showing an increasing intolerance of violence over time [1]. Unfortunately, like the corpus from the Old Bailey court, this corpus is far too large for any single individual or team to review and analyze. Therefore, we must develop automated, computational methods to explore and understand this set of artifacts from a society's significant institutions. This way we may be able to quantitatively understand how the legal system functions, how it changes, and how it relates to the society within which it is embedded.

As information scientists first, we study the citation network using graphical analysis and information theory to investigate the novelty and attachment dynamics of incoming cases, exposing self-reinforcing citation hubs as well as sparsely-connected loci. In parallel, we categorize and interrogate corpus content using topic modeling and natural language processing techniques. We then overlay the topical space with the citation network using a DeepWalk Graphical Neural Network to demonstrate the correlation between citation and topical dimensions in navigating the legal corpus. Finally, we model law over time, assessing the developing importance of cases and how circuit cases scale with population growth to inform classification of the legal system as a social structure. With the requisite amount of storage and computing power at our disposal, we use this constellation of information theoretical perspectives to shed light on the contours of the legal landscape.

By no means is this composite picture a comprehensive study; this work's intention is to quantify and qualify assumptions of a network generated not randomly but on principles of *stare decisis* and subject to precedential constraints of a hierarchical court structure. We hope these insights signal avenues for innovation in legal technologies and set the stage for broad-ranging data-driven sociolegal commentary as previous examinations of the records created by a nation's judicial system have done.

PRIOR WORK

A large and growing body of literature borrows from graph theory, information theory, and physics to systemically analyze the body of law. The edited volume Law as Data [2] represents a recent compilation of much of this work. The reader is referred to [3]; [4]; [5]; [6]; [7,8] for ground laying work here but specifically to Liebon et al. [9] whose work is closest to this paper in concept as well as Smith [10] and Coupette et al. [11] whose work is closest to this paper in scope. With the exception of Smith and Coupette et al., all aforementioned work hones in on subsets of the data analyzed in the ensuing paper. Smith and Coupette et al. respectively offer structural hypotheses and make inroads toward a generalized analytic framework. Finally, while not an application to publish opinions, Katz et al. [12] proved an example of the use of these techniques to help answer a specific research question: namely, what is driving the growth of the law and its complexity in the United States and Germany.

DATASET SCOPE AND HANDLING

The data set used for this analysis was obtained from Court Listener (www.courtlistener.com) which supports a nearly complete set of federal cases from 1926 and onward and a virtually complete set of United States Supreme Court cases prior to 1926. We use the modifier "modern" for this work, given that the dataset is dense for all federal courts' published opinions as of 1926 (see infra). We leveraged Court Listener's bulk download option to obtain a data set of 1,317,233 federal judicial opinions spanning the Supreme Court, the Appellate Courts, District Courts, Bankruptcy Courts, and most courts of special jurisdiction. Citations to corpora not a part of the Court Listener database, for example state court opinions, statutes, regulations, law reviews, etc., were removed, leaving 11,451,351 citations from federal case to federal case for network construction.

TABLE 1 | Judicial opinion network properties.

Property	Value
Number of Cases	1,317,233
Number of Citations	11,451,351
Number of Isolated Cases	299,973
Graph Density	0.000007
Minimum Case Out Degree	0
Maximum Case Out Degree	411
Average Case Out Degree	8.69
Degree Assortativity	0.099184
Number of Connected Components	1,246,635
Size of Largest Component	63,289 (4.80%)

The dataset was stored both in corpus form as well as a citation network representation with document identifiers. We supported a MongoDB database, optimized for storing documents, and a Neo4J database, optimized for storing graph networks. In-text citations define directed edges which form the citation graph network. Given the size of the network (**Table 1**), we utilized NetworKit [13], a python package which runs distributed across multiple central processing unit (CPU) cores for reasonably fast processing speeds in a high-performance computing (HPC) environment. NetworKit computed betweenness and centrality metrics for all nodes in parallel. These computations measured 164 days of core-wall time but just under 6 days in MITRE's HPC environment.

Prior Assumptions and Preliminary Analysis

Given the hierarchical structure of the judicial system and the importance it places on the concepts of *stare decisis* and precedence, we hypothesize our analyses should show, *inter alia*:

- Sparsity, because some cases have more precedential importance than others; the more-important cases should be preferentially cited at the expense of less-precedentially-important cases.
- Highly skewed (possibly power-law distributed) degree centrality among cases; we expect a Proportional Attachment [14] dynamic is at play here (also known as a Preferential Attachment dynamic [15]). We expect this near-scale-free dynamic to anneal to a power-law distribution of node centrality as law settles over time, though it may truncate the extreme tail of a true power-law distribution.
- Cases from the Supreme Court should be most central, followed by appellate courts, then district courts. Generally, nodes' degree centrality to be proportional to the hierarchy of courts represented.
- Given that cases express specific legal concepts that build upon each other, we should be able to use these concepts to trace the development of legal doctrine through time.
- There is an intuition that jurisdictions tend to specialize over time, if that is the case, then we should find consolidation in the topical distribution associated with a particular jurisdiction.

Finally, law is a product of social interaction (this is especially true in the United States where courts are constrained by the

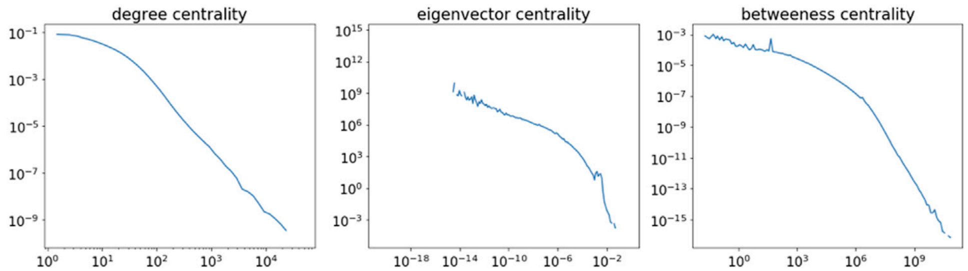


FIGURE 1 | Centrality measures.

"cases and controversies" clause of the United States Constitution; this being the case, we should find legal productivity is correlated with social systems scaling dynamics.

These assumptions serve as the starting point of preliminary empirical analysis to expose network effects of hierarchy and precedence.

Throughout of the present examination, we use the term "Federal Case Law" to mean the ideas and statements contained within the published opinions of the Article III, United States Federal Courts. We make no assertion as to the force of law of federal opinions or whether they constitute a federal common law in the same way as state common law relates to state statutes, only that they are part of a judicial system and express the ideas of the courts across time and space while being subject to a hierarchical structure and the constraints of *stare decisis* and precedence.

In what follows we discuss our initial work analyzing this large corpus and highlight some of our significant findings. Some findings confirm our intuition about the judicial system, while others raise additional questions and point to what analyses should be performed next. First, we discuss the structural features of the citation graph defined among the court opinions. Next, we discuss the topical analyses of the corpus. Third, we explore the relationship between the citation network and the semantic information of topic fits by training a Graphical Neural Network to learn structural embeddings which predict topical embeddings. Fourth, we turn to into an information-theoretic deep dive of the corpus using our topical analyis. Finally, we conclude our analyses with a discussion of how the legal system fits into broader social dynamics.

FEDERAL JUDICIAL OPINIONS CITATION GRAPH

To examine the structure of federal judicial opinions, we first treat the corpus as a citation network, in which in-text citations from opinion-to-opinion form directed links. The corpus analyzed contains 1,317,233 opinions and 11,451,351 citations between them. As discussed, *supra*, given the size of the graph, we utilized the python package NetworKit [16], which runs distributed across multiple central processing unit (CPU) cores, allowing us to perform the analyses of the structural properties of the Court Listener data set in a reasonable period of time. Basic network statistics are given in **Table 1**. Of note is the sparsity of the graph and that there are isolated cases and more than 1.2 million individual "islands" of interconnected cases. These features are consistent with the structure and function of the judicial system, namely, that only cases on point and with precedential value are cited by a given case. It is also interesting to note that there is not a citation path between all pairs of cases. This could be a feature of distinct legal concepts that are not dependent upon each other. This will be discussed further *infra*.

Next, we employed a high-performance computing (HPC) environment to compute centrality measures for each case in the network; the NetworKit python package supports parallel centrality measure calculations. Note that NetworKit has two algorithms that approximate the betweeness centrality of all nodes in a graph, as well as an algorithm that directly computes it. Using the HPC environment we were able to run the direct computation in just under 6 days (using nearly 164 days of core-wall time). The centrality distributions are shown in **Figure 1**.

Node Centrality

In order to validate some of our primary assumptions, we computed Betweeness, Degree Centrality, and Eigenvector Centrality scores for each case-node. Each case received a ranked score between 0 and 1,317,233 for each metric and a composite rank score to measure its overall node centrality. The composite rank score was created by summing the three rankings together to create a single, ordinal scale ranking which, in essence, served as a convenient method to normalize the centrality metrics. As expected, the top central cases are Supreme Court opinions (**Table 2**) with the most highly-cited Supreme Court case, Anderson v. Liberty, exhibiting the highest score for each metric and composite centrality score.

In order to comprehensively assess the hierarchical influence on node centrality, we gathered the top 1,000 nodes and grouped them by jurisdiction. This data is shown in **Figure 2**.

Of note is the striking predominance of Supreme Court case decisions. These empirical data underscore the hierarchical structure of the legal system and the role of *stare decisis* and precedence. As a whole, this initial structural analysis of the citation network among judicial opinions largely conformed with our hypotheses: exhibiting

TABLE 2 | Most central opinions.

Betweeness	Degree	Eigenvector	Date filled	Opinion	Jurisdiction	Score
0	0	0	6/25/1986	anderson-v-liberty-lobby	scotus	0
1	1	1	6/25/1986	celotex-corporation-v-myrtle-nell-catrett-administratrix-of-the-estate-of	scotus	3
11	3	2	3/26/1986	matsushita-electric-industrial-co-ltd-v-zenith-radio-corporation	scotus	16
12	5	3	5/14/1973	mcdonnell-douglas-corp-v-green	scotus	20
6	7	13	11/18/1957	conley-v-gibson	scotus	26
18	9	16	5/21/2007	bell-atlantic-corp-v-twombly	scotus	43
5	18	20	4/25/1938	erie-r-co-v-tompkins	scotus	43
28	14	8	6/6/1978	monell-v-new-york-city-dept-of-social-servs	scotus	50
3	12	41	6/25/1984	chevron-usa-inc-v-natural-resources-defense-council-inc	scotus	56
38	15	7	6/24/1982	harlow-v-fitzgerald	scotus	60

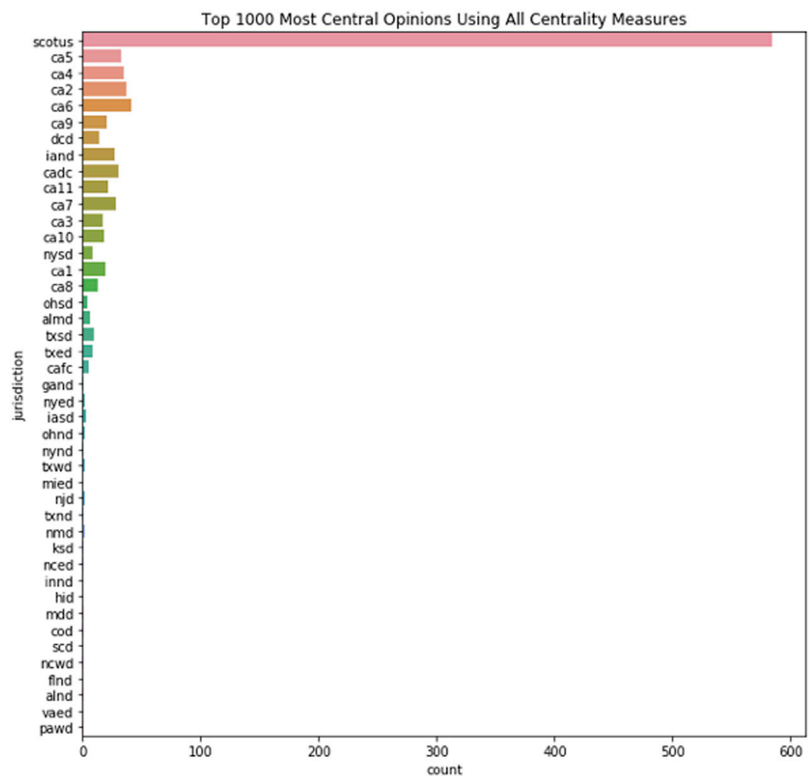


FIGURE 2 | Jurisdiction of most central options.

sparsity, highly skewed degree centrality among cases, and centrality proportional to courts' hierarchy. These points are consistent with the results of prior work and how the judicial system functions. In the next section we discuss our results of an analysis of the text of the opinions by creating a set of topic models of the judicial opinions.

Federal Opinion Topic Modeling

A topic model identifies groups of words that meaningfully co-occur to represent topics latent in a data set of documents. Topics are represented as probability distributions over the data set's vocabulary, while a probability distribution over topics is learned for each document [see [16] for a conceptual introduction]. Topic models learn topics in an unsupervised manner, enabling both inferential analysis and improved utilization of the data set, without the cost and bias of human labeling. Thematic topics that emerge from a distribution over salient words can be explicitly labeled by post-hoc analysis or may persist as unidentified groupings of words for downstream tasks on the data set.

Various implementation approaches are taken to capture latent ideas in a data set's bag-of-words (BoW) representation. Of note for its high popularity is Latent Dirichlet Allocation (LDA) [17], a probabilistic topic model that smooths its distributions by Dirichlet concentration parameters. As a generative model, LDA does not assign but draws topics and word-identities from probability distributions. With distributional sparsity, LDA encourages a constraint of perhaps more than one but not too many topics discussed in a document, and perhaps more than one but not too many uses for a given word.

LDA has been implemented in multiple ways: sampling, optimization, and trained models. Initial implementations used Gibbs sampling as implemented in MALLET [18]. Gibbs sampling is a Markov Chain Monte Carlo sampling technique that allows for sampling a distribution with a large number of variables. Instead of sampling from all variables at once, each variable is sampled in turn, conditioned on the samples of the preceding variables. Gibbs sampling is computationally expensive and non-deterministic; processing larger data sets requires approximation methods formed as an optimization problem.

Online variational Bayes [19] as implemented in sci-kit learn [20] converts the sampling problem into an optimization problem that can be solved *via* stochastic gradient descent. The algorithm alternates between subsampling the data and adjusting the hidden structure based on the subsample [21].

Here we utilize the Topic Modeling Neural Toolkit (TMNT) (https://tmnt.readthedocs.io/, paper forthcoming), an open-source topic modeling toolkit designed to compute topic models on large text collections using a neural network variational autoencoder (VAE) [22]; [23]. It is similar in approach to the Autoencoded Variational Inference for Topic Model (AVITM) as described in [24]. A neural network is trained as an inference network that maps a document directly to a distribution of topics. This is useful because it allows one to utilize a computer's graphical processing unit (GPU) to perform many of the necessary calculations. Efficiency is important here as we need to build the topic model for all 1.3 million documents in our corpus, thus using *all* opinions and jurisdictions contained in the Court Listener database.

Topical Granularity

In order to stress-test the number of useful topics supported by the legal corpus, we trained models on the Court Listener data set constrained to three different counts of latent topics: 20, 40, 80. Upon evaluating the highest performers for each topic-count, we discovered that model variants constrained to a larger number of topics were splitting topics of lower-count models: increasing the number of latent topics did not simply encourage the model to find more distinct topics but encouraged the model to find increasingly granular subtopics. To demonstrate this, we will trace a single legal issue through the 20-, 40-, and 80-topics models to show how it trickles down into topics and subtopics as the number of topics increases. In the following example, we chose intellectual property topics as they seemed particularly clear, however similar dynamics were seen across the discovered topics.

In our 20-topics model, Topic 11 is focused broadly on intellectual property as can be seen from its salient terms which includes property-protection word-forms of "copyright," "patent," and "trademark," while balancing economic legal issues of "infringement" and "monopoly" in the context of "trade" and "competition" (**Figure 3**). While the 20-topics model's Topic 10 also contains salient terms "invention" and "patent" as relevant terms, it contains more terms dealing with environmental topics, specifically "nepa" (national environmental protection act), "eis" (environmental impact statement), "specification," "environmental," and "epa." As such, it is not focused on intellectual property but on environmental policy and processes. Therefore, we comfortably assess one and only one of the model's 20 topics to be dealing with intellectual property.

In our 40-topics model we find two topics semantically-similar to the 20-topics's model Topic 11: Topics 6 and 29 (**Figure 4**). These topics show the single intellectual property topic has been split into two topics, with each further specialization. Topic 6 is focused on trademark and copyright, as evidenced by the words "trademark," "copyright," and "mark," while Topic 29 centers around patents as shown by the words "invention," "patent," and "patented."

Moreover, increasing the number of topics increases the idea granularity. For instance, the 40-topics model becomes increasingly specific in its ideas: "foia" processes, "similarity" and "confusion" issues, and "advertising" and "website" context. In turn, the 80-topics model's[1] topics derivative of "intellectual property" includes for example a product-driven topic with salient terms "features," "device," and "design." Thus, though we see an increase in diversity of concepts as the number of topics increase, these seem to emerge as finer points of their broader topic headers with fewer topics and coarser classification.

Extending the Topic Model to Include Structure

Card et al. [25] incorporate metadata into their VAE-TM implementation to guide an otherwise unsupervised clustering process [25]. The authors differentiate "labels" that were jointly generated with their document (in the generative story) from "covariate" priors that influence the topics latent in their ensuing document. With artful demonstration as to the value and effectiveness of tracing topics along a covariate structure or embedding topics together with their labels, Card et al. (2018, p. 2037–2038) motivate Structural Topic Modeling (STM) for the legal corpus with the expectation that metadata such as a case's year or jurisdiction could influence or bring to focus a case's latent topics.

Setup

Though standard vanilla LDA does not make use of priors, VAE's latent representation medium lends itself to enmeshing metadata priors together with the document's BoW. VAE-based TMNT is built on top of popular machine learning libraries such as AutoGluon [26] for plug-and-play usability as well as parameter grid-search and pyLDAvis [27] for topic model visualization and qualitative analysis (**Figures 3**, **4**, **7**, **8**).

[1]An interesting connection that arose in the 80-topics model was a high relevance of trademark to a topic centered around banking. Banks were initially tightly regulated, highly local institutions with nondescript, similar names. As regulation loosened and the local banks began to increase their geographic footprint, name collisions occurred, and court cases ensued.

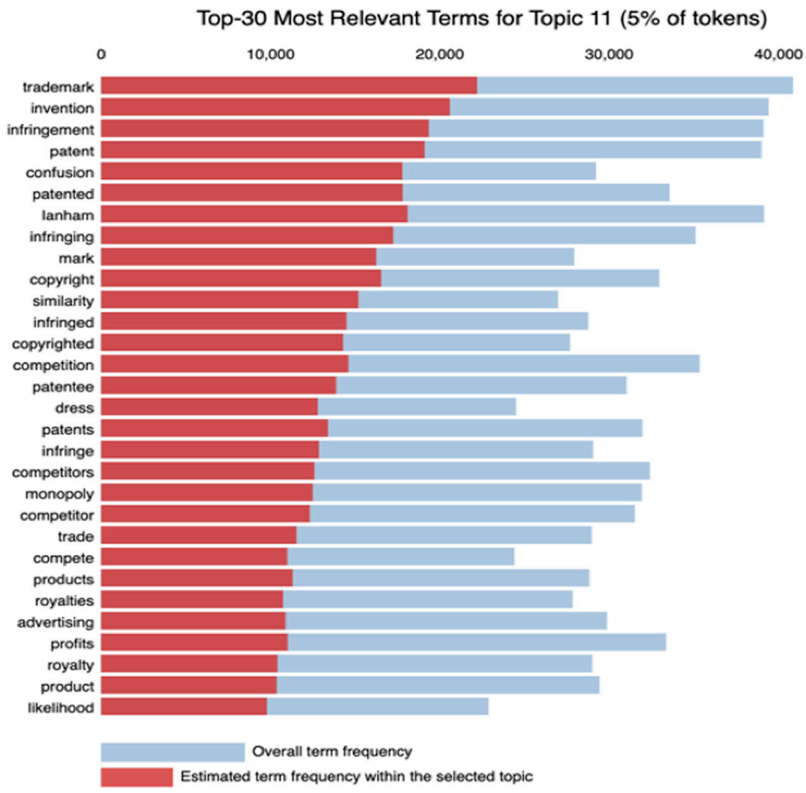


FIGURE 3 | Most relevant terms for Topic 11 of 20-tpoic model.

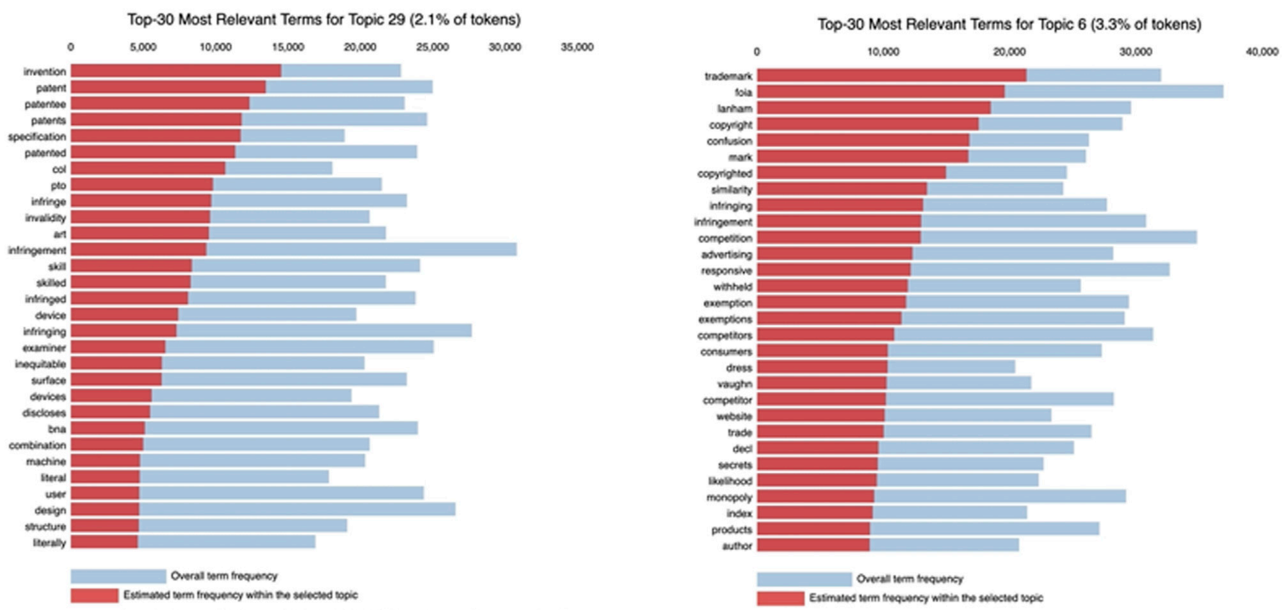


FIGURE 4 | The most relevant terms from Topics 29 and six from the 40-topic model.

TMNT also supports covariance incorporation and is, therefore, a tool of choice for STM efforts.

We start with a subnet of the connected citation hub between Plessy v. Ferguson, 163 United States 537 (1896) and the case that began to chip away at its decision: Brown v. Board of Education, 347 United States 483 (1954). This 1,808 document-sized component incorporates each central node's first-edge citation neighborhood. Considering Brown v. Board of education cites Plessy v. Ferguson directly, each belongs to the other's first-edge citation network, resulting in a tightly connected component with a maximum distance of three citation-edges between any two nodes. Given the historically-progressive nature of this component then, we might expect themes to emerge when conditioned on year, with perhaps further complexities presenting when year interacts with another covariate, jurisdiction, and topics are conditioned on the compounded covariate.

Potential Drawbacks

Though the topically-connected nature of the Plessy-Brown component enables local testing with tight experimental control, such a training set could carry significant extrapolative limitations. Foremost, we might expect loss in topic coherence and perplexity if the fully-connected citation component is separated into more topics than it can support. Furthermore, the training set (1,265 documents with a 70 percent -15 percent—15 percent train-validation-test split) may be too small to meaningfully represent tokens with vectors instantiated by random assignment. Card et al. (2018, p.2037) suggests training using pretrained word embeddings to represent tokens. TMNT supports pretrained word embeddings including word2vec [28], GloVe [29], and fasttext [30] and will be used in future work.

Incorporating Covariate Priors

To mirror aforementioned work on topical granularity, we ran TMNT on a range of topics, with 20 topics as its lower search bound and 80 topics as its upper bound. The vanilla run converged on 22 topics with topics lying along an interaction line between its two principal components. Salient topic terms include ideas of education and (de)segregation as expected from an unsupervised modeling of a Plessy-Brown (BoW) subnet.

With compelling need to pull apart the topics, we conditioned the Plessy-Brown training set on the compound covariate of case jurisdiction and year (decade). With 400 compound covariate options, the network forcefully separated its topics but converged on 50 topics, far too many latent topics than could be supported by our small connected data set. "Too many" was assessed by a dramatic drop in Normalized Pointwise Mutual Information score between training and testing, implying an overfit to statistical artifacts.

With a closer look at the principal components within Plessy-Brown's Topics' distribution, some linear progression from tolerance to criminalization becomes apparent: topics in the second quadrant (**Figure 5**) dispassionately relate to "elementary," "attendance," "zones," "neighborhood," and "desegregation" while a progression along the identity line reaches topics of "criminality" and "prosecution." With a second attempt at structural topic modeling but this time conditioned just on its decade (14 decades in total), the topic model incorporating a "decade" covariate only ever so slightly pulls apart the topics but converges back to 22 topics, mitigating concern of over-imposed conditionality. Noticeably, the Decade variant seems to project the Vanilla model's topics' distribution onto a lower-dimensional space. While there is a retention of progression from unobjected acceptance to prosecutorial offense, from this projection there seem to emerge broader ideas of rights and liberty, including women's (reproductive) rights (**Figure 6**), which historically emerged alongside the civil rights movement[2]. Thorough investigation and analysis are required to meaningfully pull apart the legal corpus while maintaining the integrity of its underlying statistics. As our work advances, we will slowly move outward from focusing on a singular connected component toward topic modeling the full legal corpus for improved structural network analysis and utilization. However, these initial analyses demonstrate the ability to find social change and legal doctrine development within a corpus utilizing NLP and statistical methods.

The topic model, while insightful by itself, also provides us with additional data we can then assign to each opinion in the corpus and make use of in conjunction with the citation network. Similar to Leibon, et al. (2018), we can now analyze the relationship between network structure and the topics contained within each opinion. We first turn to these results and follow that analysis with an information theoretic analysis of the Federal Case law across time and space (jurisdiction).

Linking the Citation Network and Topic Model

Another way to investigate the link between citation network and topic space is to employ graph-based machine learning techniques where the goal is to learn a structural node embedding. This embedding, also known as a feature vector, is then used as input into a statistical model trained to infer topic from embedding. While the 20-topics model identifies major legal areas in the Court Listener dataset useful for broad classification, the 40 and 80 topics models are able to elicit more fine-grained issues. As such, document-vectors produced from the 20-topics model were used downstream as graph neural network embeddings which provide a structural overview of the corpus. The accuracy of this model is proportional to the strength of the link between the citation space and the topic space. It answers the question of whether it is possible to use an opinion's location in the citation network to accurately infer its topic.

[2]See for example, Quanquin, H., 2019 who illustrates the push-and-pull yet side-by-side co-emergence of feminism and Black civil rights [45] and Pierson, M.D., 2005 who revisits the "patriarchal institution" of slavery in the 1850s, cautioning that while "the flow of influence ran in both directions" (p. 387), "too little attention has been paid to the ways in which the early feminist campaign affected antislavery" (p. 398) [46].

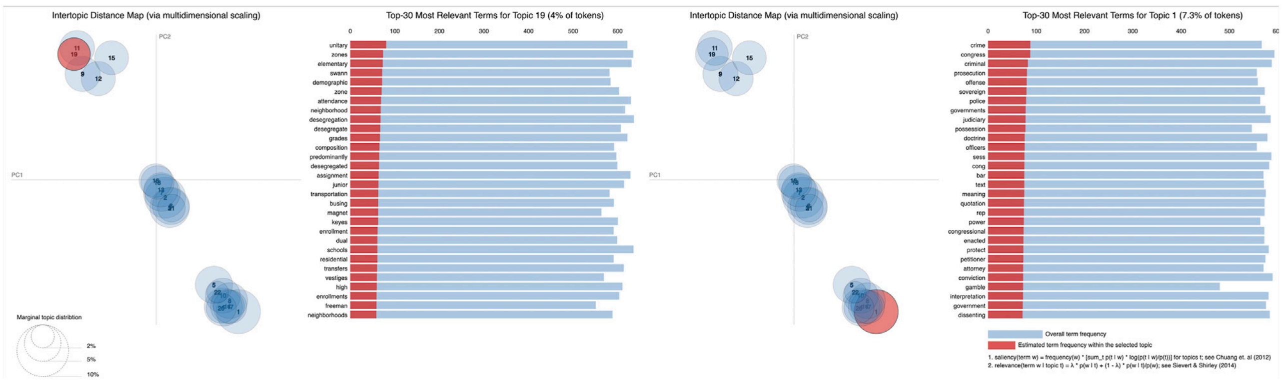


FIGURE 5 | Sentimental progression in topic space of the Plessy-Brown subnet, vanilla topic model.

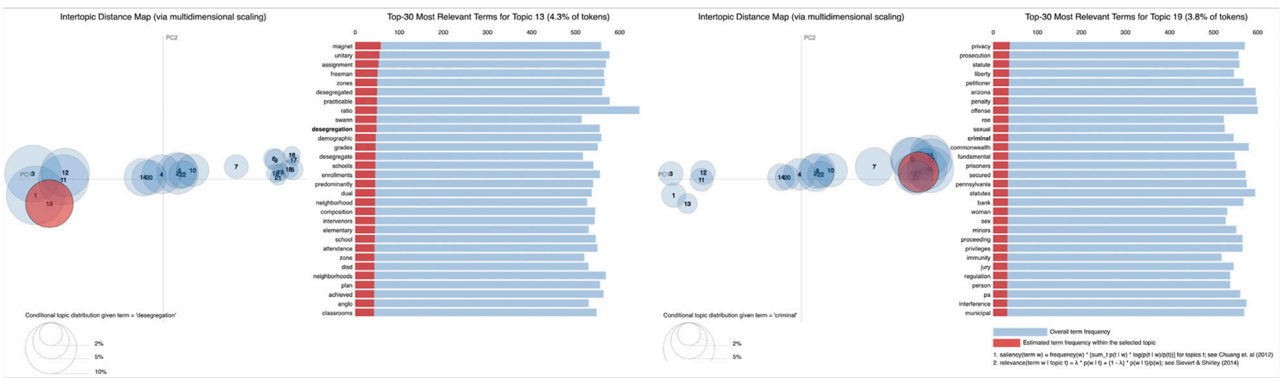


FIGURE 6 | Sentimental progression in topic space of the Plessy-Brown topic model conditioned on case decade.

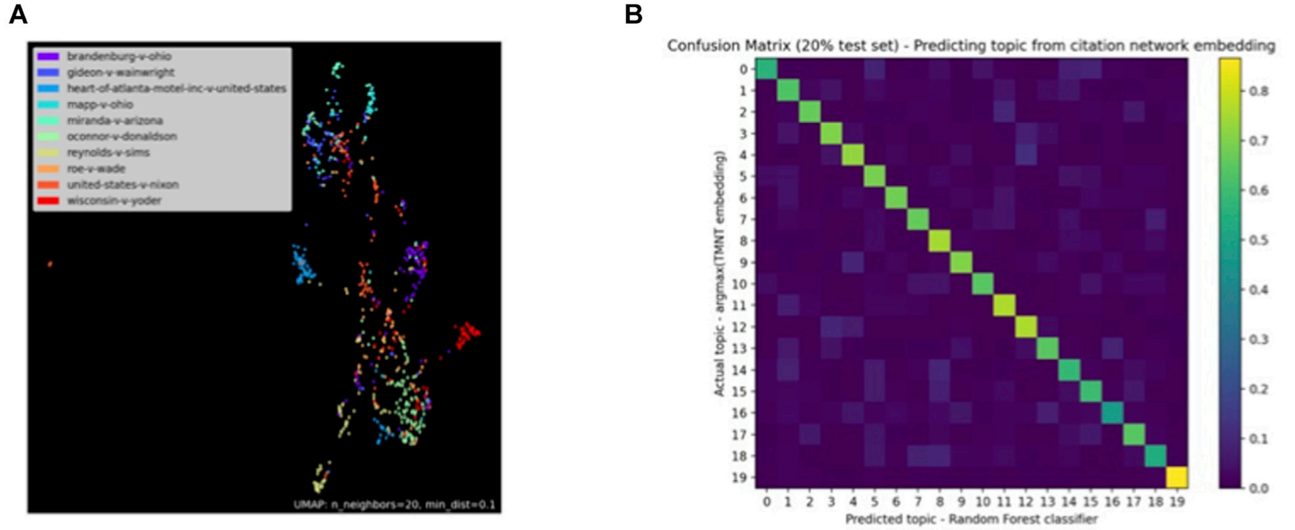
In our experiments we choose the DeepWalk algorithm [31] to learn the structural embedding of the federal opinion citation network, due to its impressive performance with some benchmark multi-label classification problems performed on a graph. The algorithm is also highly scalable, which makes it ideal for a network of this size. DeepWalk is an unsupervised feature learning technique that generalizes proven research on word embeddings originating from the natural language processing (NLP) community. The intuition being that random walks on a graph are analogous to the sentences that feed the NLP methods—we transition from sequences of words to considering short random walks along the graph. This approach enables the features learned by DeepWalk to encode community structure and neighborhood similarity. The node degree distribution of many social networks follows a power law; consequently, the appearance of vertices in a short random walk will also follow a power law. The original paper points out that word frequency in natural language follows a power law as well, and because the techniques from NLP account for this fact, it is intuitive that approaches based on NLP methods would also achieve impressive performance here. The DeepWalk algorithm is based around the following optimization problem:

$$\underset{\Phi}{Minimize} \quad -\log Pr\left(\{v_{i-w}, \cdots, v_{i-1}, v_{i+1}, \cdots, v_{i+w}\} | \Phi(v_i)\right)$$

Where Φ represents the node embedding, v_i represents the vertices, and w represents the window around node v_i. This formulation determines the embedding that maximizes the probability of a node's context, i.e., its neighbors within a certain number of hops.

Figure 7A demonstrates DeepWalk's ability to distinguish between the neighborhoods of 10 landmark Supreme Court cases using only two of its 128 dimensions. Each dot in the figure represents an opinion, and the landmark case cited is identified by the dot's color. Separation is already apparent using just two dimensions, and the full embedding encodes much more-nuanced structural information that was fully leveraged in the machine learning models discussed *supra*.

It turns out it is possible to infer the topic of an opinion (argmax of the TMNT embedding) accurately using just the citation network embedding (output from DeepWalk). It's not unexpected that the link between the structural embedding and the textual embedding would be strong. The classification target for each opinion was the largest contributing topic from the 20-topics model, and the input features were the 128 values from the



The DeepWalk differentiated neighborhoods of 10 landmark Supreme Court cases, color indicates the cited landmark case.

The confusion matrix for the random forest model, the diagonal indicates correct predictions.

FIGURE 7 | DeepWalk prediction results for 10 Landmark Supreme Court cases: **(A)** A reduced dimensional scatter plot of cases showing relative neighborhood (spatial location), and case relationship via color; **(B)** the confusion matrix for the random forest model, correct prediction is shown as the lighter color diagonal.

structural embedding produced using DeepWalk. We trained a multinomial logistic classifier and a random forest, with the random forest doing approximately two percent better on the 20 percent hold-out set. **Figure 7B** shows the confusion matrix from the random-forest model. The diagonal indicates correct predictions, and we see that the link between DeepWalk's structural embedding and TMNT's topical embedding is very strong.

INFORMATION THEORY METHODS

Upon completion of the topic modeling (discussed *supra*), we had additional information about each opinion (node) in the citation graph in the form of a vector of numbers representing how likely each discovered topic was present in the opinion. This, in addition to information about the opinion, such as date published and jurisdiction, provided additional ways to explore the dynamics of the United States federal courts. In this analysis we leveraged techniques from Information Theory to quantify changes among the opinions in time and space. First, we will discuss the methods and then move to results. As discussed *infra*, some discovered dynamics have plausible explanations, while others will require additional, follow-on analyses to adequately understand.

Novelty, Transience, and Resonance

The number of citations and other centrality measures are commonly used to quantify the importance and influence in citation networks. These measures rely on the presence of a citation network, however. We attempt to quantify the influence of an opinion purely semantically, without reference to the citation network.

The structure of the United States courts makes United States Federal Case Law a uniquely interesting corpus in which to study semantic influence. The precedential doctrine of *stare decisis,* along with the hierarchical structure of the court system, imply that a single opinion from a high court can profoundly affect the future opinions of lower courts. Moreover, the influence is asymmetric; a single decision in a lower court is not likely to exert the same influence on higher courts or even courts in other districts. Here an exogenous structure specifies the manner and flow of influence, unlike a corpus on academic papers that have, in theory, a weaker exogenous structure.

To quantify semantic influence, we begin with the information theoretic quantity of Kullback-Leibler divergence (KLD). Broadly speaking, KLD measures the difference between two probability distributions. Results in cognitive science support the interpretation of KLD as a measure of surprise [32]. Others have previously employed KLD, in combination with topic modeling, as a measure of surprise in the context of natural language [33,34]. Here, we use an 80-topic model trained on the entire corpus to infer a topic probability distribution for each document. We perform all subsequent analysis on these document vectors.

To further quantify semantic influence, we employ three measures defined by Barron et al.: novelty, transience, and resonance [35]. These measures rely on the per-document topic probability distributions produced by the aforementioned topic model. *Novelty* measures how surprising a document is, given the topic distributions of the preceding documents. High-novelty documents introduce new topics or combinations of topics into the corpus. Likewise, *transience* measures the surprise of a document, given the subsequent documents. High-transience documents contain topics or combinations of topics that do not persist within the corpus. Finally, *resonance* is the difference between the novelty and the transience. A document with high resonance introduces new topics that persist, while low resonance is indicative of documents that introduce new topics that do not persist.

The formal mathematical specification of our measures follows:

We use a topic model to generate a topic probability distribution for each document in the corpus. This distribution can be thought of as an N-dimensional vector f^i where N is the number of topics specified by the topic model. We use the 80-topic model in our subsequent analyses, so here N = 80. The nth element f_n^i of each document quantifies the relative prevalence of Topic n within document i.

The KLD between two document vectors is defined as follows:

For two opinions i and j, the KLD between their corresponding document vectors f^i and f^j is:

$$D_{KL}(f^i \parallel f^j) \equiv \sum_{n=1}^{N} f_n^i \log\left(\frac{f_n^i}{f_n^j}\right)$$

Note that KL divergence is asymmetric. This implies that:

$$D_{KL}(f^i \parallel f^j) \neq D_{KL}(f^j \parallel f^i)$$

Then we define the set of all cases in the same circuit published within the previous and subsequent w years as $S_{prev}(i)$ and $S_{sub}(i)$, respectively. All the analysis here used a window of 10 years. While many different window sizes were tested, we settled on 10 not on theoretical ground but because it produced the strongest signal.

We then compute the centroid vector by computing the mean of all the vectors within the set.

$$f^{(c)}(S) \equiv \frac{1}{|S|} \sum_{f_j \in S} f_j$$

The centroid vector represents a typical opinion in each circuit in each window of time. We define the novelty as:

$$\mathcal{N}_w(\mathrm{i}) \equiv D_{KL}\left(f^i \parallel f_{prev}^{(c)}\right)$$

And transience as:

$$\mathcal{T}_w(i) \equiv D_{KL}\left(f^i \parallel f_{sub}^{(c)}\right)$$

Finally, resonance is simply the difference between the novelty and the transience:

$$\mathcal{R}_w = \mathcal{N}_w - \mathcal{T}_w$$

Given the structure of the United States judiciary, and that courts are most influenced by cases within their jurisdiction, we calculate the novelty and transience of each opinion with a window of 10 years and compare opinions to opinions within the same circuit.

INFORMATION THEORETIC RESULTS

Transience Versus Novelty

We begin by examining the relationship between novelty and transience. One would expect to find a strong correlation between novelty and transience; opinions with a high novelty tend to have a correspondingly high transience. What is new is often forgotten. For the most part, the opinions in the corpus follow this trend very closely. **Figure 8** plots the relationship between novelty and transience. We see highly novel opinions do not tend to leave a lasting impact on the jurisprudential landscape. The slope of a linear regression indicates the bias toward novelty or transience. A fit with a slope <1 indicates a novelty bias, whereas a fit with a slope >1 indicates a transience bias. A linear regression preformed on the corpus yields a slope of 0.97, close to 1. **Figure 8** also plots the slope of the regression line for each circuit as well as the Supreme Court. All circuits have a slope of slightly less than 1. This indicates a slight bias toward novelty. The Supreme Court is the closest to 1, with a slope of 0.97. The 10th Circuit has the largest novelty bias, with a slope of 0.87. Overall, the United States court system lacks a strong novelty or transience bias. This strongly implies that change moves slowly within the court system, as implied by *stare decisis*. Furthermore, the novelty bias is consistent with the notion that law becomes settled over time and individuals tend to bring new cases and controversies to a court.

Novelty Over Time

Looking at novelty over time, we see two effects: the mean novelty increases over time, and the 95-percent bootstrapped confidence interval shrinks over time. **Figure 9** plots the mean novelty for all cases over time. Each circuit follows a similar pattern. From 1920 to 1970 the mean novelty increases over time and has a large confidence interval, meaning the data is scattered. After 1970 the mean novelty has a much lower confidence interval and stops increasing. The significance of these trends were tested using the standard Mann-Kendall test with level set to 0.05 and were found to be significant, matching our assumptions from examining the graph. The presence of a statistically meaningful trend suggests that the decreasing confidence interval results from a larger number of opinions per year, not from a true variance in the mean novelty.

This finding that novelty increases over time perhaps suggests an acceleration in the pace of law. The low-mean novelty in the early 20th century suggests that the pace of law was comparatively slow; new opinions did not stray from what came before. Does this suggest the pace of law is increasing? Why would this be? Perhaps law scales with population, the pace of life increases with the size of a city. This is explored in more detail, *infra*.

ArticleRank and Resonance

Another way to quantify the influence of an opinion is to use ArticleRank. ArticleRank is a centrality measure closely related to Google's PageRank algorithm. The ArticleRank algorithm has proven useful in the analysis of citation measures. Unlike the traditionally used measure of times cited, ArticleRank does not weight all citations equally. Citations from other influential opinions are weighted more heavily [36].

Do our purely semantic influence measures correlate with the ArticleRank?

We find that the ArticleRank and resonance are largely orthogonal measures. **Table 3** shows the Pearson correlation

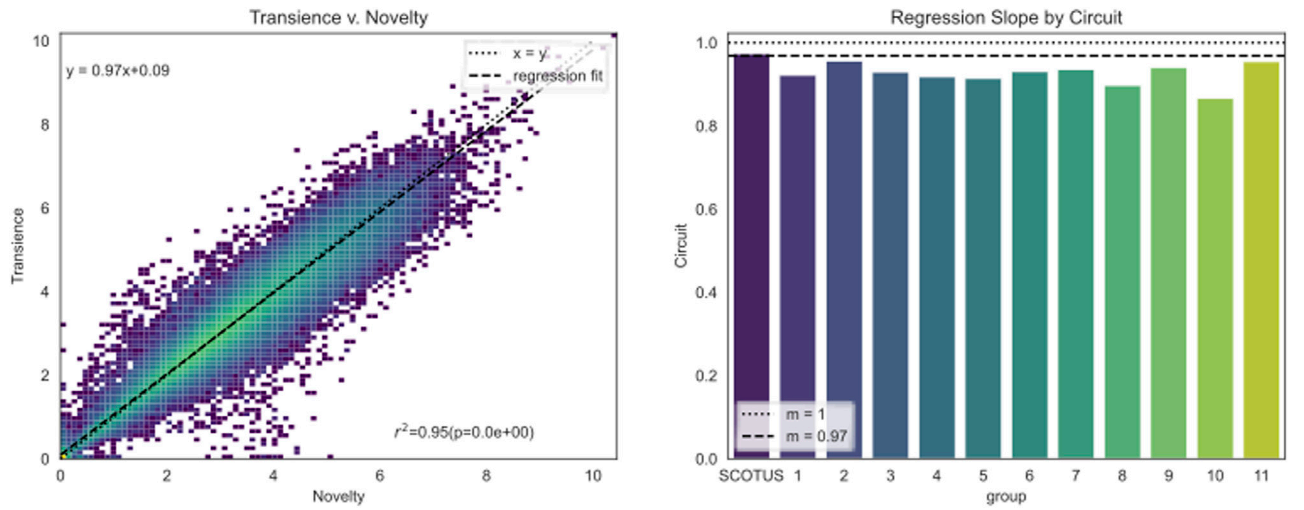


FIGURE 8 | The relationship between transience and novelty.

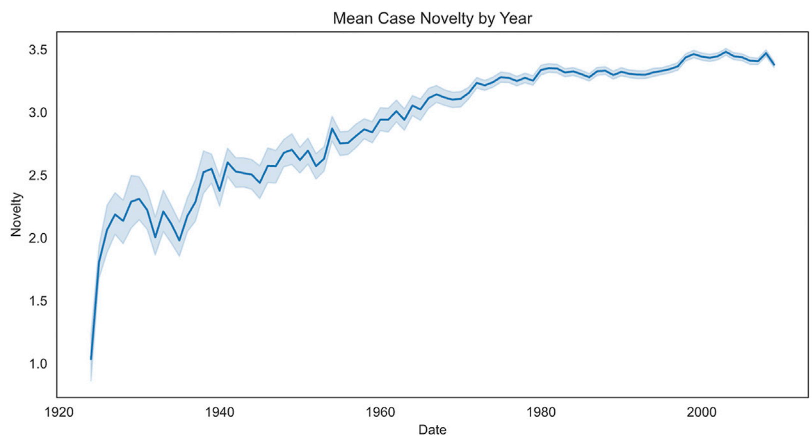


FIGURE 9 | Mean case novelty by year.

TABLE 3 | Correlation between ArticleRank and semantic measures.

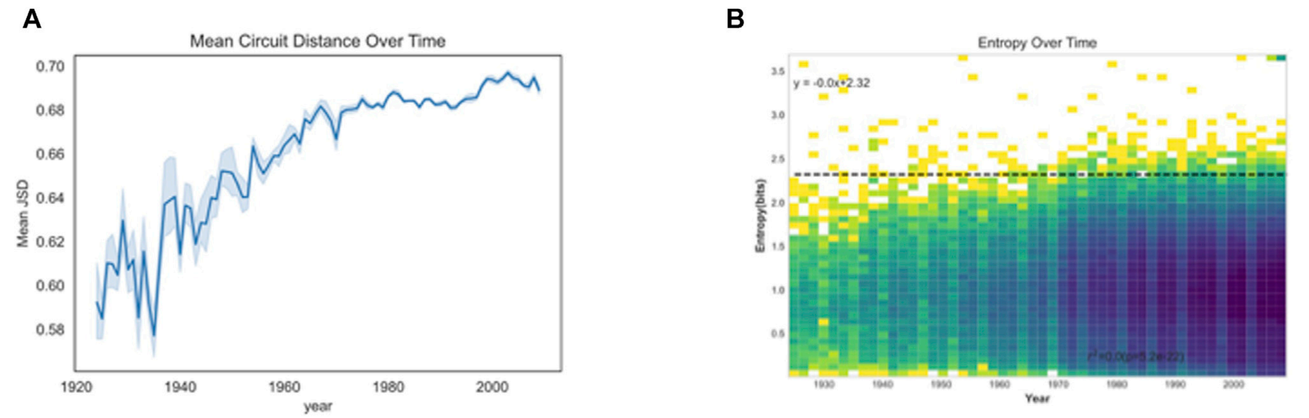
Correlation quantities	Pearson correlation	*p*-value
ArticleRank-Novelty	0.014	1.3E-12
ArticleRank-Transience	0.010	5.0E-7
ArticleRank-Resonance	0.010	5.0E-7

between ArticleRank and Novelty, Transience, and Resonance. There appears to be no meaningful correlation between citation-based measures of influence and semantic measures of influence within our corpus. A high ArticleRank does not imply an opinion will have a high or low resonance.

Jensen Shannon Distance

In the analyses above, we compared trends within circuits. In the subsequent sections we want to examine trends between circuits. To quantify the distance between circuits, we compute the Jensen-Shannon distance (JSD) between pairs of circuits. Unlike KLD, which is a type of f-divergence, JSD distance is true distance metric and is therefore symmetric. The JSD between two vectors p and q is defined as the square root of the mean of the KLD between each vector and their mean vector. We use JSD here rather than KLD because there is less of a sense of “direction” between circuits. With, for example, novelty, there is a directionality to the question, so there is a specific way to apply KLD. However, when comparing the 6th Circuit to the 7th Circuit, there is no specific directionality to the analysis, thus the need to use JSD rather than KLD. The specific JSD calculation is:

$$JSD = \sqrt{\frac{D(p\|m) + D(q\|m)}{2}}$$



Mean JSD among circuits over time.

Entropy over time.

FIGURE 10 | Mean JSD among circuits over time and associated entropy: **(A)** Shows increasing mean JSD and decreasing variance over time among circuits; **(B)** Show relatively little change in entropy over time among circuits.

Distance Between Circuits

To examine how the circuits change over time we compute the mean JSD between each pair of circuits for each year. This data is displayed in **Figure 10A**. Once again, the trend is checked using the Mann-Kendall test and found to be significant to a level of 0.05. The results indicate a significant increase in the distance between courts over time. The distance begins to increase more slowly in the 1970s.

The most obvious cause of an increase in semantic distance would be specialization. Many have noted specialization occurs among judges, even those on courts without special jurisdiction, see generally Baum [37,38] and Wasserman and Slack [39]. This hypothesis seems reasonable since circuits loosely correspond to geographic location and as societies specialize in activities geospatially (for example, banking and equity trading in New York City), the courts will disproportionately hear cases associated with these social activities. Given this dynamic among individual judges and the role of precedence and *stare decisis*, one might expect this to cause a more generalized specialization over time within a jurisdiction. If this were the case, one would also expect the mean opinion entropy to drop over time. However, **Figure 10B** shows the entropy over time remains, essentially, constant. The Pearson correlation between the year and the entropy is −0.022 with a p-value of 5.2E-22. No correlation exists between the entropy and time, and therefore no evidence exists of any increasing specialization.

We can also see this lack of specialization for circuits over time by looking at the frequency of the most common topics within circuits. In the context of the topic model, a court specializing in a particular type of case should appear as a specific topic having a high value in the document probability vectors for a progressively larger proportion of cases over time. To investigate this possibility, for each case in a circuit we identified the topic that had maximum probability in the topic probability vector and then plotted the distributions of most likely topics over the period between 1950 and 2010. **Figure 11** shows these distributions for the 9th circuit. The distributions are grouped by decade and displayed for each of the 20-, 40-, and 80-topic models. From the graphs it is apparent that no single topic is ever the most likely for more than 25% of cases in that decade, and for the 40- and 80-topic models it was extremely rare for a single topic to be most likely in more than 10% of cases in that decade. Graphs of the other 10 circuits demonstrate broadly similar behavior to the to **Figure 11**.

The lack of consolidation within the topical distributions of jurisdictions demonstrates a lack of specialization over time, but that still leaves the question of what is causing the increasing distance between circuits. One possibility is that the increasing circuit distance is a result of what Smith labels "legal clustering." The cause for clustering in the citation network is evident. He writes: "A court is likely to cite what is jurisdictionally relevant. The judge will prefer to cite a case from his own court or from a higher court in its jurisdiction than from some remote jurisdiction" [10]. High correlation exists between distance in the citation network and semantic distance. It is not unreasonable to conclude that judges also tend to use jurisdictionally relevant language and concepts. The increasing semantic distance between circuits could be explained by an increase in the amount of clustering within the network. This hypothesis could be tested in several ways. The correlation between the inter-circuit semantic distance and the clustering within the citation network could be measured. Additionally, if semantic clustering is the cause of the increasing semantic distance, one would expect the semantic distance between lower courts and higher courts in the same jurisdiction to be consistently lower than the semantic distance between lower courts in different jurisdictions. These are analyses queued up for our next study.

Investigating Scaling and the Pace of Law

Many complex systems have been shown to exhibit power-law scaling relationships. These relationships have been found in biological systems, like the scaling of metabolic rate with weight

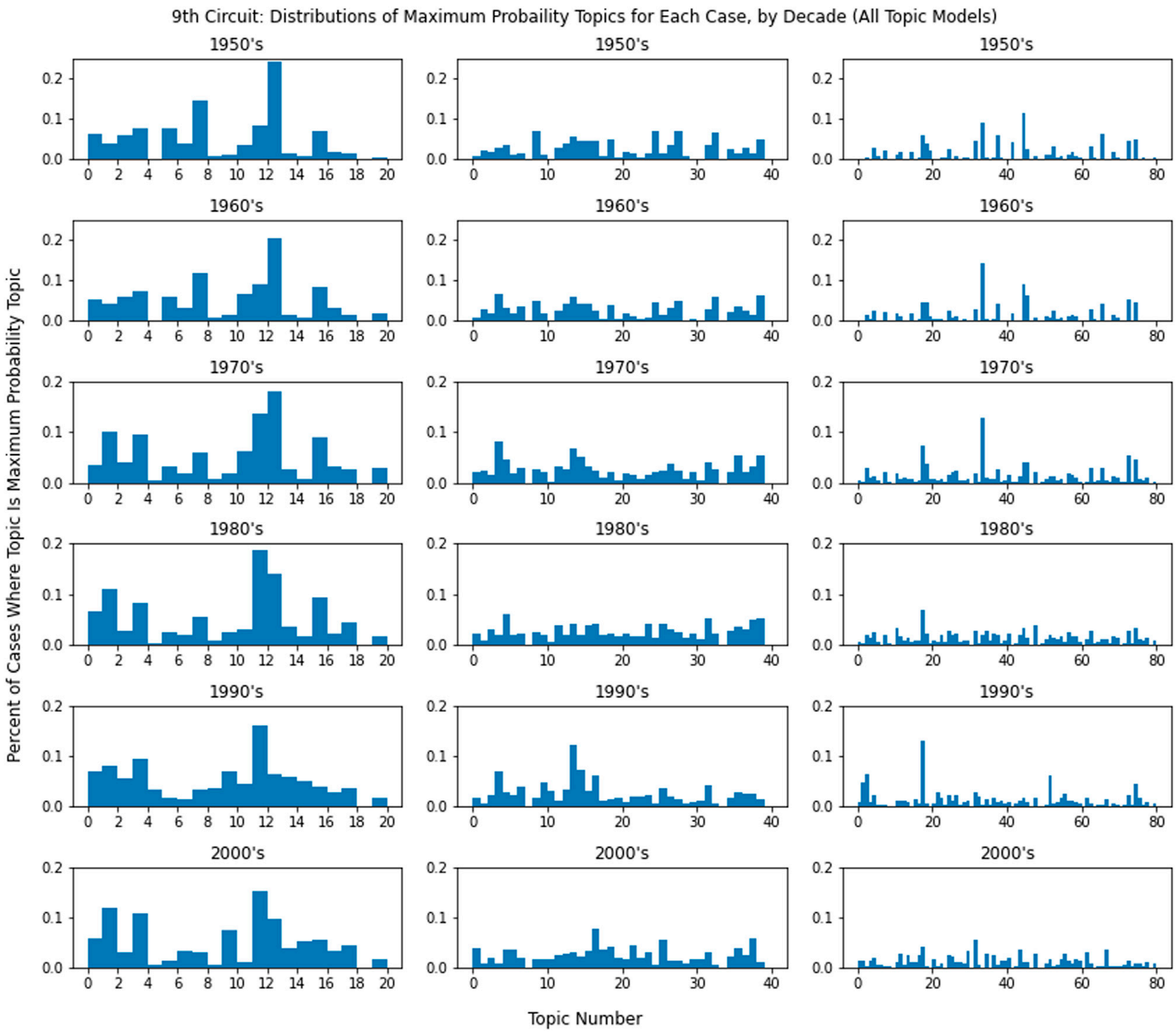


FIGURE 11 | Distributions of max probability topics by decade for each topic model.

[40]. Importantly, the scaling relationships within biological systems are sublinear, larger organisms metabolize more efficiently. We find scaling laws in social systems like cities, as well. Some aspects of cities related to infrastructure scale sub-linearly [41]. Larger cities exploit efficiencies and require less infrastructure per resident than smaller cities. Unlike biological systems, social and economic aspects of cities scale super linearly. Research by Bettencourt, West, and others indicates that these super-linear scaling relationships result from the properties of cities' social networks and the increased interaction among the residents.

Next, we investigate the scaling properties of the United States legal system. On the one hand the United States legal system acts like infrastructure, a publicly funded institution intended to provide justice for its citizens. From this perspective one might imagine the legal system scales sub-linearly like other pieces of infrastructure. On the other hand, the law is a deeply social activity, and the cases it sees arise from the social interactions between individuals. In this case we would expect the law to scale super-linearly. It should also be noted that we limited the dataset used for this analysis to the United States Supreme Court (as the most influential) and Circuits 1–11 (as having some geospatial connection via jurisdictional boundaries). Jurisdictions based upon subject were removed as they have very little geospatial connection to a particular region.

We use two different metrics to quantify the scaling of the legal system. The first is the total number of opinions filed per year. This simply measures the productivity of a circuit. We also consider the ArticleRank, which quantifies the amount of influence a given jurisdiction has over the legal system. We begin by examining the dynamics of the case count by year and the population separately. For population data we used

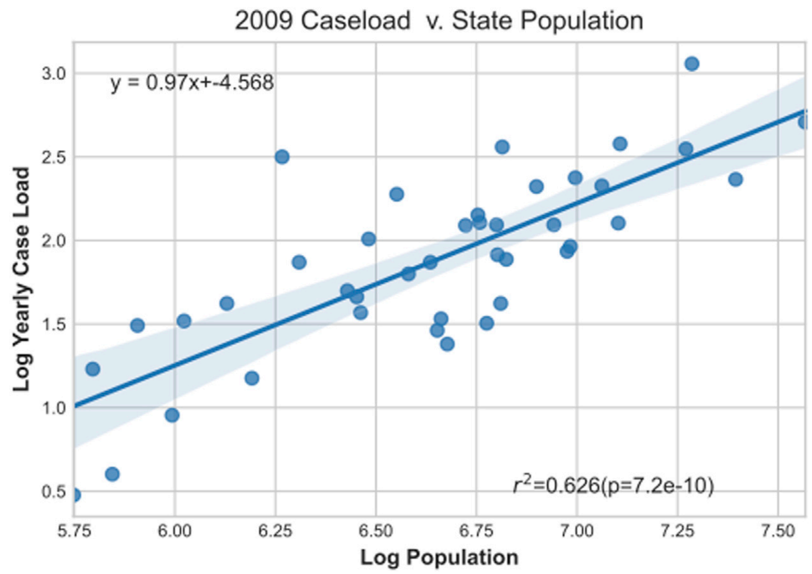


FIGURE 12 | Example case load versus population plot for 2009.

United States Census Annual Estimates for the United States population by state.[3] Not surprisingly, the state population grows quite differently than the opinions of the surrounding judicial jurisdiction.

Scaling–Linear Regressions

We use linear regressions to quantify these scaling properties. We plot the number of cases published in each state district court versus the population of the state, both on linear axes as well as on a log-log scale. If the linear regression explains the data well on the log-log plot, with $a > 1$, this will lend credence to the idea that the number of cases scales exponentially with the state population. If the data is better explained by a regression on the data with linear axes, or if neither regression explain the data well, one may conclude the data does not follow a power-law like scaling relationship. **Figure 12** plots an example of the caseload versus population data on a log-log plot; the regression is displayed as the blue line. The regression suggests the scaling relationship is almost linear, with a slope of 0.97, which indicates that law scales somewhere between social infrastructure and social interaction.

We perform this same analysis of the data for every year in the data set, from 1926 to 2020, on both linear as well as log-log axes. The linear regressions on linear axes result in slopes that increase significantly over time from near 0 to over 1.7. The log-log axes linear regression slopes suggest a relationship between population and caseload that changes over time from highly sublinear, with slopes near 0.4, to close to linear. The r-squared values of the linear axes regressions begin to decrease in the 1970s while the r-squared values of the log-log regressions begin to steadily increase at the same rate. These results, taken together, suggest the linear scaling properties of the United States Federal Case Law only begin to emerge in the 1970s. The near-linear scaling suggests that the law as whole does not benefit from efficiencies as would infrastructure that scales sub-linearly. Neither does the law benefit from the snowballing effect of social interaction. Furthermore, one might expect the largest jurisdictions to produce a disproportionate number of opinions. Instead, the number of opinions appears to be directly proportional to the population of the state. Again, this suggests law is a "hybrid." Meaning, it is both an outgrowth of social interaction, as well as infrastructure supporting the general functioning of society.

TEMPORAL SCALING BY CIRCUIT

Previously we examined the scaling relationship between a state's population and the number of opinions authored. These scaling relationships are static, examining the relationship between states at a single moment in time. Next, we examine the scaling of each circuit over time. We group the states by circuit and sum the population of each state within a circuit to find the total historical population of each circuit. **Figure 13** plots the results and shows the statistics for each fit. We find that the scaling of each circuit is super-linear. The minimum and maximum slopes of the linear regression are 1.9 and 7.3, with a mean of 3.96. The r-squared values range from 0.52 to 0.94, with a mean of 0.8. The only true outlier is Circuit 10. The distribution for circuit 10 visually does not fit the regression line, and an r-squared value of 0.52 lends further support. What is different about circuit 10? Circuit 10 contains much of the American west and southwest and has a small population relative to its size. Perhaps this lower population

[3]Data from Annual Estimates of the Population for the U.S. and States, and for Puerto Rico | FRED | St. Louis Fed.

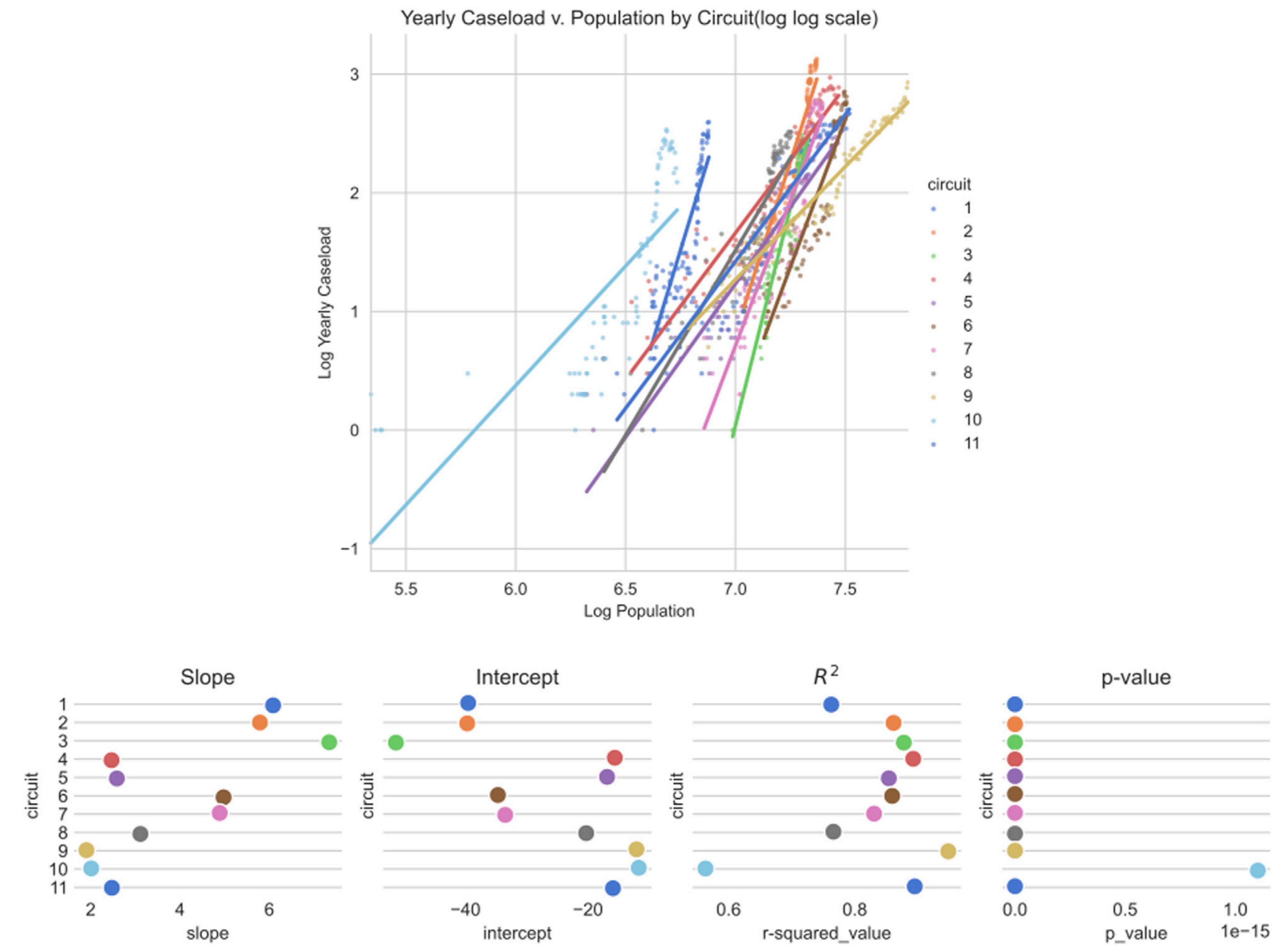


FIGURE 13 | Graph of yearly caseload versus population by circuit and statistics for linear regression fits of yearly caseload versus population by circuit.

density causes fewer sociolegal collisions and interactions, thereby making this circuit an outlier.

Interestingly, when examined as a whole, i.e., aggregating across all circuits, one finds sublinear scaling. These results, when coupled with the previous results, paint a strange picture. The scaling relationship across states at a given moment in time is sub-linear. The scaling relationship within the same circuit over time is super-linear. The same super-linear relationships that exist withing circuits over time exist within individual states over time as well. This is likely a function of the level of aggregation used within the analysis, an example of the Simpson's Paradox. Comporting with the "cases and controversies" clause, one would expect court productivity to be closely related to social activities contained within its jurisdiction and, thus, scale super-linearly. In addition to meeting this prior expectation, intra-circuit activity rather than inter-circuit activity controls precedence. Thus, though circuit jurisdiction is not tightly coupled to its surrounding populations, on balance, circuit-by-circuit super-linear scaling seems to best model the pace of law over time.

ARTICLE RANK POWER-LAW SCALING

Description

In their 2007 paper "Web of Law," T. Smith showed that American case law is a scale-free network, in which the number of citations follows a power-law distribution [10]. Smith conjectured that this structure is a result of so-called preferential attachment. Highly cited opinions are more visible and thus more likely to receive more citations in the future. In this section we expand this work, fitting a power law not to the number of citations but to the ArticleRank of a case. ArticleRank is less sensitive to the dynamics of preferential attachment. A relatively unknown case may achieve a high ArticleRank by being cited by several landmark cases. We also examine the ArticleRank distribution in federal district, federal appellate, and United States Supreme Court cases separately to compare the dynamics. **Figure 14** plots the complementary cumulative distribution function (CCDF) of the of the ArticleRank distributions as solid lines and their corresponding power law fits as dotted lines.

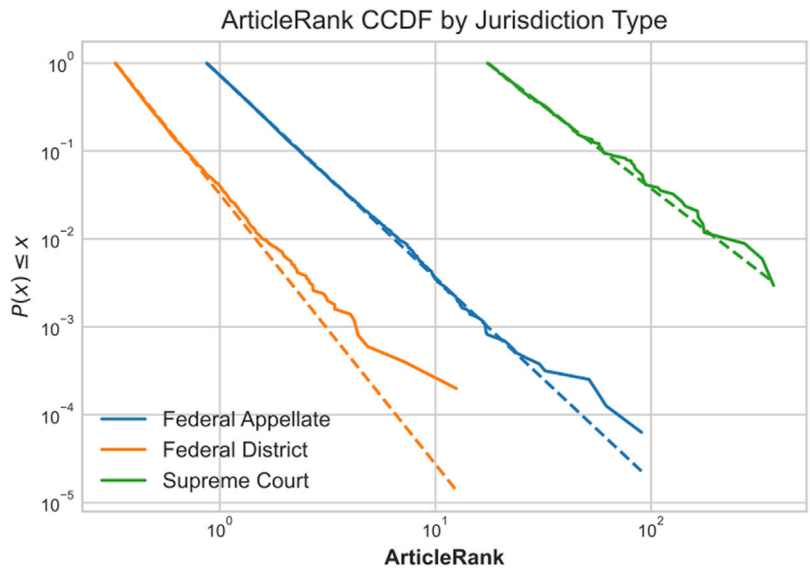


FIGURE 14 | Article Rank CCDF by jurisdiction type.

TABLE 4 | Scaling exponent for each court type.

Court type	Alpha
Federal District	4.072855
Federal Appellate	3.302189
Supreme Court	2.879686

RESULTS

We use the statistical methods developed by [42] and implemented in python by [43]. Examining the CCDF of the ArticleRank distributions within **Figure 14**, the major difference between the three types of federal courts is immediately obvious. Supreme Court opinions have a much higher ArticleRank than federal appellate opinions, and federal appellate opinions have a much higher ArticleRank than federal district opinions. This result is to be expected. Higher courts have more influence than lower courts. The scaling exponent, alpha, of each distribution is displayed in **Table 4**. The scaling exponent determines the heaviness of the distribution tail. A low scaling exponent implies the distribution has more weight in its tail, or more high-rank cases. A high scaling exponent implies a distribution has fewer high-rank cases. The table shows that higher courts have lower scaling exponents, or more high-rank cases.

As power laws are mathematical objects, one does not expect to find perfect power-law fits from data gathered from a real-world system. This being the case, it is difficult to say definitively if a power-law fit is statistically significant. Typically, one would look at the fits of several different skewed distributions, with the assumption that the best fit of the set is the true underlying distribution. Following from Alstott, we use comparative methods to assess the goodness of fit of the distributions. We fit several candidate distributions to the data and compute the log-likelihood ratio of each to determine what distribution explains the data best. The results of the power-law fit as well as the log-likelihood ratios are shown in **Table 5**. A positive value of the log-likelihood ratio indicates the first distribution is a better fit than the second, while a negative value indicates the reverse. The exponential distribution is an exceedingly poor candidate. The results are inconclusive for the two higher courts, the Supreme Court and federal appellate courts; for both courts, the log-likelihood ratio between the power law and the log-normal distribution is close to zero. The district court has a power-law and log-normal log-likelihood ratio of −0.47, indicating the log-normal distribution is a significantly better fit than a power law. Once more, the high p-values indicate these results are not conclusive as to the power law fit.

T. Smith showed that the number of citations in the United States Federal Case Law follows a clear scale-free distribution [10]. Our results are performed on another metric, ArticleRank, and performed on a significantly larger sample of the corpus. Our results, in contradistinction to T. Smith, show that the ArticleRank, while clearly a heavy-tailed distribution, is not definitively a power law. One of two causes could explain this discrepancy. First, ArticleRank is not strongly dependent on the *number* of citations an opinion receives. It has been shown both empirically and analytically that the degree distribution of scale-free networks follows a power law [15], but others have shown analytically that the PageRank distribution in scale-free networks departs from a power law for large and small values [44]. If this result holds for ArticleRank as well as PageRank, it could explain the deviations in our data. Second, our fits were conducted on a much larger sample than Smith's. It is possible the United States Federal Case Law network is not in fact scale-free, although more work would be needed to confirm or reject this hypothesis. For example, when a landmark case is published it may be initially heavily cited per a preferential

TABLE 5 | Goodness of fit for power-law distribution by court-type. *p*-values of log-likelihood ratios are in parenthesis.

Court type	KS statistic for power law fit	Power Law/Log-Normal	Power law/Exponential	Log-normal/Exponential
		Log-likelihood ratio	Log-likelihood ratio	Log-likelihood ratio
Federal Appellate	0.006119	0.073231 (0.778403)	2,228.198 (1.64E-27)	2,228.124 (1.45E-27)
Federal District	0.015566	−0.47829 (0.371466)	548.2672 (2.46E-11)	548.7454 (2.93E-11)
Supreme Court	0.026405	0.003575 (0.733253)	63.86727 (0.000175)	63.8637 (0.000173)

attachment dynamic; however, as that part of the law becomes well-settled it may be litigated less often, resulting in fewer citations and thus breaking the preferential attachment dynamic and truncating the power law into something more closely resembling an exponential distribution.

CONCLUSION

This paper explored the structure and dynamics of modern United States Federal Case Law. The dataset was larger than previous work of which we are aware. While not all of the analyses performed were novel, e.g., others have analyzed the citation network structure of some cases, we demonstrated the ability to perform these analyses at very large scales, and our analyses confirmed previously hypothesized features of the citation network: namely sparsity and degree centrality that is both highly skewed and proportional to the hierarchy of courts represented. Other analyses performed on this data were novel in this space and demonstrated that these techniques can highlight change occurring within society that are then reflected in changes to the society's legal system. Adding the structure of topic modeling to the citation network gave a grounding for tracing the development of legal doctrine through time. In future analyses we hope to investigate additional specific threads of legal doctrine development [akin to the constitutional analyses undertaken by 47]. The results of analysis with topic modeling also give evidence against the common intuition that jurisdictions have specialized in particular types of cases over time, even if the circuits have grown further apart as measured by certain distance metrics. We find signatures of *stare decisis* and precedence within the data via increasing ArticleRank within the court hierarchy and a distribution not clearly scale-free nor exponential. This also provides quantitative evidence that cases are built very purposefully. While not a shocking insight, it is instructive to know that these data contain the correct signals. We also find the impact of the "cases and controversies" clause in the scaling dynamics of opinion production, scaling super-linearly with population growth suggesting courts are directly impacted by social interaction. Furthermore, this analysis highlighted the impact of controlling precedence, by showing that the "correct" level of analysis is likely the circuit rather than the whole of the courts. Additionally, our analysis indicates that a change occurred in the dynamics of the federal courts in the 1970s. It is our intention to investigate this further in a subsequent study.

DATA AVAILABILITY STATEMENT

All raw data was obtained via bulk download from www.courtlistener.com.

AUTHOR CONTRIBUTIONS

KA, JH, MK, JK, SM, CP, ZS, WT, and RZ performed analyses, MK and BB designed the project, all authors contributed to the creation of the manuscript equally.

FUNDING

This research was funded by the MITRE Innovation Program.

REFERENCES

1. Klingenstein S, Hitchcock T, DeDeo S. The Civilizing Process in London's Old Bailey. *Proc Natl Acad Sci USA* (2014) 111(26):9419–24. doi:10.1073/pnas.1405984111
2. Livermore MA, Rockmore DN. *Law as Data: Computation, Text, and the Future of Legal Analysis*. Santa Fe, NM: SFI Press (2019).
3. Bommarito MJ, Katz DM, Zelner JL, Fowler JH. Distance Measures for Dynamic Citation Networks. *Physica A: Stat Mech its Appl* (2010) 389(19):4201–8. doi:10.1016/j.physa.2010.06.003
4. Bommartio M, Katz D, Isaacs-See J. An Emperical Survey of the Population of U.S. Tax Court Written Decisions. *Va Tax Rev* (2011) 30(3):523–58.
5. Bommarito M, Katz D, Zelner J. 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 (ICAIL '09). IEEE (2010).
6. Dadgostari F, Guim M, Beling PA, Livermore MA, Rockmore DN. Correction to: Modeling Law Search as Prediction. *Artif Intell L* (2020) 29:1. doi:10.1007/s10506-020-09264-2
7. Ash E, Chen DL. Mapping the Geometry of Law Using Document Embeddings. Ash, Elliott and Chen, Daniel L., Mapping the Geometry of Law Using Document Embeddings (2018). Available at: https://ssrn.com/abstract=3305761 (Accessed July 6, 2020).
8. Fowler JH, Johnson TR, Spriggs JF, Jeon S, Wahlbeck PJ. Network Analysis and the Law: Measuring the Legal Importance of Precedents at the U.S. Supreme Court. *Polit Anal* (2007) 15:324–46. doi:10.1093/pan/mpm011
9. Leibon G, Livermore M, Harder R, Riddell A, Rockmore D. Bending the Law: Geometric Tools for Quantifying Influence in the Multinetwork of Legal Opinions. *Artif Intell L* (2018) 26(2):145–67. doi:10.1007/s10506-018-9224-2
10. Smith TA. The Web of Law. *San Diego L Rev* (2007) 44:309–54. doi:10.2139/ssrn.642863
11. Coupette C, Beckedorf J, Hartung D, Bommarito M, Katz DM. Measuring Law over Time: A Network Analytical Framework with an Application to Statutes

and Regulations in the United States and Germany. *Front Phys* (2021) 9:1. doi:10.3389/fphy.2021.658463
12. Katz DM, Coupette C, Beckedorf J, Hartung D. Complex Societies and the Growth of the Law. *Sci Rep* (2020) 10:18737. doi:10.1038/s41598-020-73623-x
13. Staudt CL, Sazonovs A, Meyerhenke H. NetworKit: A Tool Suite for Large-Scale Complex Network Analysis. *Net Sci* (2016) 4(4):508–30. doi:10.1017/nws.2016.20
14. Simon HA. On a Class of Skew Distribution Functions. *Biometrka* (1955) 42(3-4):425–40. doi:10.1093/biomet/42.3-4.425
15. Baraba´si A-L, Albert R. Emergence of Scaling in Random Networks. *Science* (1999) 286:509–12. doi:10.1126/science.286.5439.509
16. Mohr JW, Bogdanov P. Introduction-Topic Models: What They Are and Why They Matter. *Poetics* (2013) 41(6):545–69. doi:10.1016/j.poetic.2013.10.001
17. Blei DM, Ng AY, Jordan MI. Latent Dirichlet Allocation. *J Learn Res* (2003) 3: 993–1022.
18. McCallum AK. MALLET: A Machine Learning for Language Toolkit (2002). Retrieved from http://mallet.cs.umass.edu/ (Accessed October 11, 2019).
19. Hoffman M, Bach F, Blei D. Online Learning for Latent Dirichlet Allocation. *Adv Neural Inf Process Syst* (2010) 23:1.
20. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. *J Machine Learn Res* (2011) 12: 2825–30.
21. Hoffman MD, Blei DM, Wang C, Paisley J. Stochastic Variational Inference. *J Machine Learn Res* (2013) 14:1303–47.
22. Kingma DP, Welling M. Auto-Encoding Variational Bayes. In: Proceedings of the 2nd International Conference on Learning Representations (ICLR). IEEE (2014).
23. Rezende DJ, Mohamed S, Wierstra D. Stochastic Backpropagation and Approximate Inference in Deep Generative Networks. In: Proceedings of the 31st International Conference on Machine Learning. PMLR (2014). p. 1278–86.
24. Srivastava N, Sutton C. Autoencoding Variational Inference for Topic Models. In: International Conference on Learning Representations. IEEE (2017).
25. Card D, Tan C, Smith NA. NEural Models for Documents with Metadata. *Proc 56th Annu Meet Assoc Comput Linguistics* (2018) 1:2031–40. doi:10.18653/v1/p18-1189
26. Erickson N, Mueller J, Shirkov A, Zhang H, Larroy P, Li M, et al. *AutoGluon-Tabular: Robust and Accurate AutoML for Structured Data*. arXiv preprint (2020).
27. Sievert C, Shirley KE. LDAvis: A Method for Visualizing and Interpreting Topics. In: Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces. IEEE (2014). p. 63–70. doi:10.3115/v1/w14-3110
28. Mikolov T, Sutskever I, Chen K, Corrado G, Dean J. Distributed Representations of Words and Phrases and Their Compositionality. In: NIPS'13: Proceedings of the 26th International Conference on Neural Information. IEEE (2013). p. 3111–9.
29. Pennington J, Socher J, Manning CD. GloVe: Global Vectors for Word Representation (2014). Retrieved from https://nlp.stanford.edu/pubs/glove.pdf (Accessed September 11, 2019). doi:10.3115/v1/d14-1162
30. Joulin A, Grave E, Bojanoswski P, Mikolov T. Bag of Tricks for Efficient Text Classification. In: Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2. Valencia, Spain: Short Papers (2017). p. 427–31. doi:10.18653/v1/e17-2068
31. Perozzi B, Al-Rfou R, Skiena S. DeepWalk. In: Proceedings of the 20th ACM SiGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM (2014). doi:10.1145/2623330.2623732
32. Itti L, Baldi P. Bayesian surprise Attracts Human Attention. *Vis Res* (2009) 49: 1295–306. doi:10.1016/j.visres.2008.09.007
33. Murdock J, Allen C, DeDeo S. Exploration and Exploitation of Victorian Science in Darwin's reading Notebooks. *Cognition* (2017) 159:117–26. doi:10.1016/j.cognition.2016.11.012
34. Thompson WH, Wojtowicz Z, DeDeo S. *Lévy Flights of the Collective Imagination*. Lévy Flights of the Collective Imagination (2018). Available at: https://arxiv.org/abs/1812.04013.
35. Barron ATJ, Huang J, Spang RL, DeDeo S. Individuals, Institutions, and Innovation in the Debates of the French Revolution. *Proc Natl Acad Sci USA* (2018) 115:4607–12. doi:10.1073/pnas.1717729115
36. Li J, Willett P. ArticleRank: a PageRank-based Alternative to Numbers of Citations for Analysing Citation Networks. *Ap* (2009) 61(6):605–18. doi:10.1108/00012530911005544
37. Baum L. Specialization and Authority Acceptance: The Supreme Court and Lower Federal Courts. *Polit Res Q* (1994) 47(3):1. doi:10.2307/448849
38. Baum L. Probing the Effects of Judicial Specialization. *Duke L J* (2009) 58:1. doi:10.1093/sysbio/syp011
39. Wasserman MF, Slack JD. Can There Be Too Much Specialization? In: Specialization in Specialized Courts. Chicago, IL: Northwestern University Law Review (2021). p. 1.
40. West GB, Brown JH. The Origin of Allometric Scaling Laws in Biology from Genomes to Ecosystems: towards a Quantitative Unifying Theory of Biological Structure and Organization. *J Exp Biol* (2005) 208:1575–92. doi:10.1242/jeb.01589
41. Bettencourt LMA, Lobo J, Helbing D, Kuhnert C, West GB. Growth, Innovation, Scaling, and the Pace of Life in Cities. *Proc Natl Acad Sci* (2007) 104:7301–6. doi:10.1073/pnas.0610172104
42. Clauset A, Shalizi CR, Newman MEJ. Power-Law Distributions in Empirical Data. *SIAM Rev* (2009) 51:661–703. doi:10.1137/070710111
43. Alstott J, Bullmore E, Plenz D. Power law: A python Package for Analysis of Heavy-Tailed Distributions. *PLoS ONE* (2014) 9:e85777. doi:10.1371/journal.pone.0085777
44. Avrachenkov K, Lebedev D. *PageRank of Scale Free Growing Networks*. Valbonne, France: Tech. rep. (2006).
45. Quanquin H. Slavery Cannot Be Covered up with Broadcloth or a Bandanna": The Evolution of White Abolitionist Attacks on the "Patriarchal Institution. *J Early Republic* (2005) 25(3):383–415.
46. Pierson MD. Why Study the Early American Women's Rights Movement? *Am J* (2019) 68(1):1.
47. Rockmore DN, Fang C, Foti N, Ginsburg T, Krakauer DC. The Cultural Evolution of National Constitutions. *J Am Soc Inf Sci Technol* (2018) 69 (3): 483–494. doi:10.1002/asi.23971

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers | Frontiers in Physics

PERSPECTIVE
published: 08 April 2022
doi: 10.3389/fphy.2022.760780

Jurisprudence Meets Physics

Matthew Koehler *

The MITRE Corporation, McLean, VA, United States

For many system level questions jurisprudential data has grown to a size and scale that no longer lends itself to traditional analytic techniques driven by human examination and direct analysis. While there will always be vast numbers of specific questions well within the capabilities of humans, an understanding of the system as a whole is no longer among them. Over the past several decades jurisprudence has begun to use mathematical and other analytic techniques many of which were developed in the physical sciences. It is now time for jurisprudence to embrace more fully the analytic tools of these other disciplines, specifically those coming out of physics, in order to continue to produce new insights to aid in the structure, function, design of judicial systems and the analysis of judicial dynamics.

Keywords: judicial dynamics, jurisprudence, analysis, physics, nontraditional data, generative methods

OPEN ACCESS

Edited by:
Pierpaolo Vivo,
King's College London,
United Kingdom

Reviewed by:
Andrea Rapisarda,
University of Catania, Italy
Michal Shur-Ofry,
Hebrew University of Jerusalem, Israel

***Correspondence:**
Matthew Koehler
mkoehler@mitre.org

Specialty section:
This article was submitted to
Social Physics,
a section of the journal
Frontiers in Physics

Received: *18 August 2021*
Accepted: *26 January 2022*
Published: *08 April 2022*

Citation:
Koehler M (2022) Jurisprudence
Meets Physics.
Front. Phys. 10:760780.
doi: 10.3389/fphy.2022.760780

INTRODUCTION

Jurisprudential data grows monotonically over time. Every law that is passed, every case that is decided, every brief that is filed increases the size of the jurisprudential dataset. In 1756 Blackstone published a version of his lecture notes on English Common Law that later became the four volume set Commentaries on the Laws of England [1]. He was able to do this based upon his efforts to understand the common law developed over approximately the first 20 years of his legal career and with a relatively small number of other individuals. Today, even if one only examines US Federal judicial opinions, this would be a difficult feat. **Figure 1** shows the monotonic increase in opinions over time. The quantity of opinions just within the United States is now well into the millions. Even with very optimistic assumptions (each being only ten pages in length and being able to read a page in 2 minutes for 12 hours a day), it would take a team of one hundred people years to read them all. And, of course, the number of opinions would continue to grow over that period.

Moreover, federal opinions are not the only part of this jurisprudential dataset. There are executive orders, statues, regulations, state court opinions, treaties, constitutions, court transcripts, etc. The task of understanding a nation's judiciary is now beyond the scope of a human, or even a team of humans, over the course of their career. It is now time for jurisprudence to embrace more fully the analytic tools and techniques from other disciplines that are designed to deal with this scale. For example, CERN is able to process approximately one petabyte of data per day, and the Large Hydron Collider alone produces about twenty-five petabytes of data per year. Given that there are tools and techniques from physics and other fields that can handle the scale of jurisprudential data, will these analytic techniques provide any useful insights for jurisprudential study? In what follows, I will argue that not only have these techniques produced meaningful insights for jurisprudential study, but they can also produce insights that are not able to be created by other means.

Is it valid to approach jurisprudence with such a new set of tools? Simply put, yes, it is consistent with concepts of jurisprudence to use tools and techniques from outside the legal discipline to study the law. [2] commented on the decline of law as an autonomous discipline as the field of economics grew in importance within jurisprudence. This was further stressed a few years later when Posner articulated an approach to jurisprudence that is consistent with the perspective taken here:

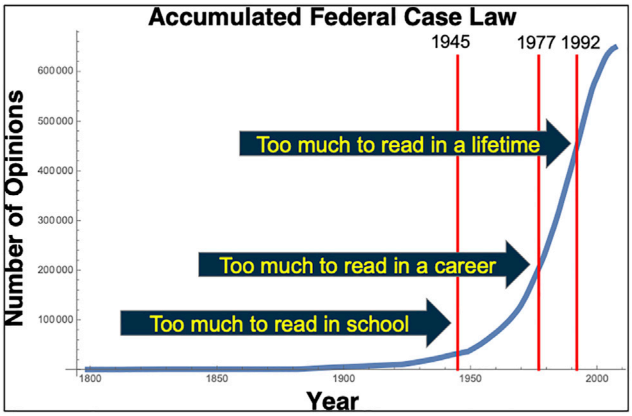


FIGURE 1 | The monotonic accumulation of judial opinions over time.

"By 'jurisprudence' I mean the most fundamental, general, and theoretical plane of analysis of the social phenomenon called law. For the most part it deals with problems, and uses perspectives, remote from the daily concerns of legal practitioners: problems that cannot be solved by reference to or by reasoning from conventional legal materials; perspectives that cannot be reduced to legal doctrines or to legal reasoning" ([3] at xi).

ANALYTIC METHODS

The artifacts generated by a nation's judiciary are becoming more and more accessible. For example, the Free Law Project's CourtListener website provides access to the text of millions of judicial opinions from many Federal and State jurisdictions (www.courtlistener.com). The Harvard Caselaw Access Project has digitized forty million pages of US court opinions spanning 360 years (www.lil.law.harvard.edu). As a final example, transcripts of London's Old Bailey court from 1,674–1913 have been digitized and made available online (www.oldbaileyonline.org). The US Courts even have biographical information available for all justices that have sat and are sitting on a court. And, of course, the code of federal regulations and US code are all available in machine usable formats.

Researchers have made great use of these datasets. Notable early work in this space includes [4] and [5]. These works examined judicial voting behavior through correlation analyses to shed light on the decision making of justices and whether or not a judge's decisions are consistent over time and topic, as well as consistent with institutional traditions such as stare decisis. This line of correlative analysis of judicial dynamics has been significantly extended and broadened over time with works such as [6] that examined the statistical mechanics of the US Supreme Court.

Bias in judicial decision making has also received analytic-based analyses, such as those by [7], and [8]. [9] and [10] both studied law from a geometric perspective created by embedding the text of opinions in a high dimensional space. Finally, researchers are even having success predicting the citations a judicial opinion contains based upon its language [11].

As data about judiciaries and their actions have increased researchers have begun studying different aspects of dynamics of legal systems, from their development to their structure. For example, Barron et al. [12], explored the dynamics of debates from the National Constituent Assembly during the French Revolution during which the new French state was formed. Katz and his many collaborators have studied the dynamics of lawsuits around the US tax code, the movement of clerks from one court to another, the citation structure among the US code and judicial opinions, and created methods of characterizing the temporal dynamics of laws and regulations (e.g. [13,16]).

Many of the aforementioned studies have benefited from tools and techniques developed within Physics, including, inter alia, Ising models, graph analysis, information theory, statistical mechanics, and maximum entropy methods. For example, Barron et al., supra, used a natural language processing technique, topic modeling, to transform the raw text of the speeches into a set of "topics" (here a topic is a collection of co-occurring words) with a numeric value characterizing how likely each topic was contained within the document. Now that concepts of all the texts could be related to each specific text, the authors could use another technique, Kullback-Leibler Divergence, to measure changes in ideas contained within the speeches over time. This is an example of how these new tools can be used by interested researchers to explore collections of documents far too large to read or examine individually. Another example of the utility of these approaches comes from [17]. Here the authors used similar techniques to those of Barron et al., but here the unit of analysis was national constitutions. With these techniques the research team was able to show the flow of concepts from one constitution to another across both time and space and was able to characterize the relative impact of a given constitution based upon its "downstream" influence.

GENERATIVE NUMERIC METHODS

While analytic methods have experienced tremendous growth, so much so that the Santa Fe Institute published a volume on law as data [18], generative methods have experienced much less. For the present discussion I will use generative numeric methods to be representations of judicial processes through time based upon models of their function at the exclusion of numeric approximations used for some analytic methods, e.g., the approximation of stochastic partial differential equations. More specifically, I refer to the use of simulation as a means of testing our understanding of the generative mechanisms at play within a judicial system. The analytic methods discussed supra do an excellent job of producing insights into the current state of a legal system and how that system changed over time, but they do not provide as much insight into why a judicial system produced the observed dynamics or how the system might respond to a perturbation. Here, I argue, progress can be made with the combination of jurisprudential theory and simulation.

Furthermore, as a judiciary is made up of many dynamically interacting heterogeneous components (judges, lawyers, citizens, etc.) who may learn and adapt through time and who are distributed across a meaningful space (jurisdictions), the most efficient way to analyze the system's potential future state is to explicitly represent it and simulate its state changes through time [19]. Typically, this is done via the agent-based model [20,21].

As discussed by [22] the use of agent-based models has dramatically increased over the past 20 years, however they remain largely absent from the jurisprudential literature. As highlighted by the examples discussed by Benthall and Stranburg, when law and agent-based models do collide it is largely in the space of regulation and policy analysis or, more generally, a topic within law and economics. These are truly important uses of agent-based models as they allow for a richer representation of human behavior and decision making than most other methods [23] and have led to many insights, but this is not where the use of agent-based models should end.

In my opinion jurisprudence has before it one of the most fascinating subjects of study available to any discipline. It is studying a complex system that has become self-aware and is now trying to guide itself into specific equilibria, e.g., our society has formalized governing institutions that then created laws and regulations in order to induce its members into particular sets of behavior. The use of agent-based models to examine this aspect of jurisprudence appears to not have been largely embraced . . . , yet.

This is unfortunate as agent-based models provide the jurisprudential scholar with a truly new way to study a society and its judiciary. What if France had a different judicial system? Is a common law judicial tradition a good way to solve hard problems? Ceteris paribus, if the costs associated with courts were to change in manner X what would happen to their utilization? If all judges in a judiciary are slightly biased does that make the system as a whole slightly or greatly biased? If one assumes that a better understanding of a nation's judicial system is critical for the long-term stability and prosperity of a nation, then agent-based modeling provides the jurisprudential scholar with a uniquely powerful way to explore these and many other questions. Unfortunately, I am aware of only three works that specifically use an agent-based model to explore legal or judicial dynamics ([24–26]). These works explored the evolution of stable norms/institutions, the impact of changing information quantities on jurispathic and jurisgenerative judicial decision making, and the evolutionary dynamics of judicial systems respectively. Agent-based models can be particularly useful for abductive exploration, perhaps most famously performed by [27] during his analysis of segregated settlement dynamics in large US cities. In that work Schelling was able to show that even with a society that prefers integrated neighborhoods, if individuals have even a slight bias and do not coordinate their movements, segregated settlement patterns will emerge. And, thus, he was able to create a coherent system from seemingly incoherent signals (individuals prefer integration but create segregated settlement patterns).

BEHAVIORAL AND EXPERIMENTAL JURISPRUDENCE AND AN ANALYTIC COUPLING

Relatively new trends from cognitive psychology and game theory that will aid dramatically the development of our understanding of jurisprudential dynamics are Behavioral Law and Economics [28] and Experimental Jurisprudence [29]. This growing body of literature highlights another example jurisprudence can take from physics: that of the tight coupling of theory and experimentation with perhaps the clearest example those where the existence of a subatomic particle is determined theoretically long before it is discovered experimentally. On the jurisprudential side, one potential example of this coupling could be the Coase Theorem [30]. Coase's theoretical analysis concluded that while the law establishes how negotiations commence it will not impact the conclusion. This analysis was game theory-based and assumed no transaction costs. This conclusion and the relative impact of the rather strong assumption about transaction costs could be experimentally studied within the growing field of experimental jurisprudence.

Another example of the utility of this coupling of techniques relates to the body of jurisprudential literature relating to the notion that common law (judge made law) will evolve to higher levels of efficiency over time largely irrespective of how judges decide the outcome of a case, see generally [31]. Unfortunately, when [32] tested this theory empirically they found no evidence of this increasing efficiency over time. We now have theory and analysis at odds with each other. Here generative techniques may be used to explore how this inconsistency could arise and what it might mean. [26] was able to show through a simulation-based analysis that judicial problem solving may exhibit punctuated dynamics resulting in very short periods of improvement and long periods that resembled random walks. This being the case, a relatively small sample of judicial dynamics would be more likely to show random activity than improvement. In this abductive analysis simulation was used to show that what at first seemed like incoherent results from theory could actually be coherent. Given ongoing advances in cognitive science, experimental jurisprudence, and behavioral law and economics, it is not difficult to imagine using agent-based models to explore the decision-making dynamics of juries, social ideas of justice, or, more tactically, the functioning of a court room or judiciary and how that functioning may be made more efficient.

DISCUSSION

The combination of tools, techniques, and practices of physics and jurisprudence would be a very powerful way to explore judicial dynamics and better understand what drives these dynamics and how we might create more effective judicial systems and reforms to existing systems. Together these fields could observe a particular dynamic, analyze data collected from it, expose statistical regularities in the data, then formulate a closed form expression of the system that is consistent with the regularities found in the data. The abstract representation can be used to understand the basic dynamics of the system and bound its behavior. As understanding improves, the model system can be moved from a closed form representation to that of a simulation allowing for the relaxation of assumptions often needed to express a complex system in closed form. I believe it is the tight coupling of analysis, theory, human-centric experimentation, and generative analyses that will allow jurisprudence to create ever new and more useful insights into judicial dynamics.

While it is likely naive to envision a time when jurisprudence has explanatory and predictive power on par with physics, one can foresee a time when jurisprudence is a science made up of qualitative and quantitative methods; where theory, quantitative analysis, and simulation come together to provide a more complete picture of justice and judicial dynamics. As societies face questions about judicial reform, the impact of bias in decision making, or the use (and impact of) artificial intelligence-based systems within a judicial system these methods become more important, especially when coupled together. These methods provide us with a way to understand the potential impact of changes to, and reforms of, a judiciary. Additionally, these methods provide a way to experiment with a judiciary en silico before making changes that could potentially have negative social or judicial consequences and may be difficult and time consuming to undo.

Until recently it was difficult to conceptualize how these tools could be leveraged within the field of jurisprudence. However, as data continues to become more available, agent-based modeling tools become easier to use, and computer hardware becomes more powerful hopefully this will begin to change. As can be seen in much of the literature highlighted herein, and more generally within this special issue, the narrative data produced by judicial systems can now be fruitfully analyzed and used to produce many new insights and test our intuitions about judicial dynamics in ways unavailable a few years ago. This, combined with emerging academic programs that combine law and data science, will produce the next generation of researchers forging a new jurisprudential science with a tight coupling among theory, experimentation, and analytics all of which will be well informed by other fields such as physics.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: www.uscourts.gov.

AUTHOR CONTRIBUTIONS

The author confirms being the sole contributor of the work and it has been approved for public release, case numbers 13-4197.

REFERENCES

1. Cook DH. Sir William Blackstone: Life and Legacy Set Apart for God's Work. *Regent Univ L Rev* (2000) 13(1):169–78.
2. Posner RA. The Decline of Law as an Autonomous Discipline: 1962-1987. *Harv L Rev* (1987) 100(4):761–80. doi:10.2307/1341093
3. Posner R. *The Problems of Jurisprudence*. Cambridge, MA: Harvard University Press (1990).
4. Segal JA, Spaeth HJ. *The Supreme Court and the Attitudinal Model Revisited*. Cambridge, United Kingdom: Cambridge University Press (2002). doi:10.1017/CBO9780511615696
5. Epstein L, Knight J. *The Choices Justices Make*. Thousand Oaks, California, United States: Sage (1997).
6. Lee ED, Broedersz CP, Bialek W. Statistical Mechanics of the US Supreme Court. *J Stat Phys* (2015) 160(2):275–301. doi:10.1007/s10955-015-1253-6
7. Shayo M, Zussman A. Judicial Ingroup Bias in the Shadow of Terrorism *. *Q J Econ* (2011) 126:1447–84. doi:10.1093/qje/qjr022
8. Gazal-Ayal O, Sulitzeanu-Kenan R. Let My People Go: Ethnic In-Group Bias in Judicial Decisions-Evidence from a Randomized Natural Experiment. *J Empirical Leg Stud* (2010) 7(3):403–28. doi:10.1111/j.1740-1461.2010.01183.x
9. Ash E, Chen DL. Mapping the Geometry of Law Using Document Embeddings (2018). Available at SSRN: https://ssrn.com/abstract=3305761 (Accessed 15 March 2019).
10. Leibon G, Livermore M, Harder R, Riddell A, Rockmore D. Bending the Law: Geometric Tools for Quantifying Influence in the Multinetwork of Legal Opinions. *Artif Intell L* (2018) 26:145–67. doi:10.1007/s10506-018-9224-2
11. Dadgostari F, Guim M, Beling PA, Livermore MA, Rockmore DN. Correction to: Modeling Law Search as Prediction. *Artif Intell L* (2020) 29:1. doi:10.1007/s10506-020-09264-2
12. Barron ATJ, Huang J, Spang RL, DeDeo S. Individuals, Institutions, and Innovation in the Debates of the French Revolution. *Proc Natl Acad Sci USA* (2018) 115(18):4607–12. doi:10.1073/pnas.1717729115
13. Bommarito M, II, Katz D. A Mathematical Approach to the Study of the United States Code. *Physica A: Stat Mech its Appl* (2010) 389:4195–200. doi:10.1016/j.physa.2010.05.057
14. Bommartio M, II, Katz D, Isaacs-See J. An Empirical Survey of the Population of U.S. Tax Court Written Decisions. *Va Tax Rev* (2011) 30(3):523–58.
15. Coupette C, Beckedorf J, Hartung D, Bommarito M, Katz DM. Measuring Law Over Time: A Network Analytical Framework with an Application to Statutes and Regulations in the United States and Germany. *Front Phy* (2021) 9. doi:10.3389/fphy.2021.658463
16. Katz DM, Stafford D. Hustle and Flow: A Social Network Analysis of the American Federal Judiciary. *Ohio State Law J*. (2010) 71: 457.
17. Rockmore DN, Fang C, Foti NJ, Ginsburg T, Krakauer DC. The Cultural Evolution of National Constitutions. *J Assoc Inf Sci Technol* (2017) 69(3):483–94. doi:10.1002/asi.23971
18. Livermore M, Rockmore D, editors. *Law as Data*. Santa Fe, NM: Santa Fe Institute Press (2019).
19. Buss S, Papadimitriou C, Tsitsiklis J. On the Predictability of Coupled Automata: An Allegory about Chaos. In: Proceedings [1990] 31st Annual Symposium on Foundations of Computer Science; Octobert 15–16, 1999; Chicago, IL. IEEE (1990).
20. Epstein J, Axtell R. *Growing Artificial Societies*. Washington, D.C: Brookings Institution Press (1996).
21. Axtell R. Why Agents? on the Varied Motivations for Agent Computing in the Social Sciences. In: Proceedings of the Agents 2000 Conference; 1999 October 15–16; Chicago, IL. Argonne National Lab (2000).
22. Benthall S, Strandburg K. Agent-Based Modeling as a Legal Theory Tool. *Front Phys* (2021) 9. doi:10.3389/fphy.2021.666386
23. Epstein J. *Generative Social Science*. Evanston, IL: Princeton University Press (2006).
24. Picker RC. Simple Games in a Complex World: A Generative Approach to the Adoption of Norms. *Univ Chicago L Rev* (1997) 64(4):1225. doi:10.2307/1600216
25. Koehler M. Legal Agents: Agent-Based Modeling of Dispute Resolution. In: Proceedings of the Agent 2007 Conference; 2007 November 15–17.; Evanston, IL. Argonne National Lab (2007).

26. Koehler M. *The Blind Lawmaker*. PhD Thesis. Fairfax, VA: the George Mason University (2012).
27. Schelling T. Models of Segregation. *Am Econ Rev* (1969) 59(2):488–93.
28. Sunstein C, editor. *Behavioral Law and Economics*. New York: Cambridge University Press (2000).
29. Sommers R. Experimental Jurisprudence. *Science* (2021) 373(65553):394–5. doi:10.1126/science.abf0711
30. Coase RH. The Problem of Social Cost. *J L Econ* (1960) 3:1–44. doi:10.1086/466560
31. Elliott ED. The Evolutionary Tradition in Jurisprudence. *Columbia L Rev* (1985) 85(1):38–94. doi:10.2307/1122404
32. Niblett A, Posner RA, Shleifer A. The Evolution of a Legal Rule. *J Leg Stud* (2010) 39(2):325–58. doi:10.1086/652908

Conflict of Interest: Author MK was employed by the MITRE Corporation.