Artif Intell Law (2014) 22:337–374
DOI 10.1007/s10506-014-9160-8

Measuring the complexity of the law: the United States Code

Daniel Martin Katz · M. J. Bommarito II

Published online: 16 September 2014

Abstract Einstein's razor, a corollary of Ockham's razor, is often paraphrased as follows: make everything as simple as possible, but not simpler. This rule of thumb describes the challenge that designers of a legal system face—to craft simple laws that produce desired ends, but not to pursue simplicity so far as to undermine those ends. Complexity, simplicity's inverse, taxes cognition and increases the likelihood of suboptimal decisions. In addition, unnecessary legal complexity can drive a misallocation of human capital toward comprehending and complying with legal rules and away from other productive ends. While many scholars have offered descriptive accounts or theoretical models of legal complexity, most empirical research to date has been limited to simple measures of size, such as the number of pages in a bill. No extant research rigorously applies a meaningful model to real data. As a consequence, we have no reliable means to determine whether a new bill, regulation, order, or precedent substantially effects legal complexity. In this paper, we begin to address this need by developing a proposed empirical framework for measuring relative legal complexity. This framework is based on "knowledge acquisition", an approach at the intersection of psychology and computer science, which can take into account the structure, language, and interdependence of law. We then demonstrate the descriptive value of this framework by applying it to the U.S. Code's Titles, scoring and ranking them by their relative complexity. We measure various features of a title including its structural size, the net flow of its intra-title citations and its linguistic entropy. Our framework is flexible, intuitive, and transparent, and we offer this approach as a first step in developing a practical methodology for assessing legal complexity.

D. M. Katz (✉)
Michigan State University, East Lansing, MI, USA
e-mail: katzd@law.msu.edu

M. J. Bommarito II
Lex Predict, LLC, Wayne, NJ, USA
e-mail: mike@lexpredict.com

Keywords Legal complexity · Measuring complexity · Political economy · Artificial intelligence and law

1 Optimal legal complexity

Einstein's razor, a corollary of Ockham's razor, is often paraphrased as follows: make everything as simple as possible, but not simpler (Einstein 1934). This rule of thumb describes the challenge that designers of a legal system face—to craft simple laws that produce desired ends, but not to pursue simplicity so far as to undermine those ends. Whether considered by legal theorists or scholars in subfields such as tax, corporate or environmental law, significant debate addresses how law might achieve particular societal ends "as simply as possible".

There can be no doubt that detailed legal rules are sometimes required to produce desired outcomes. Through an exacting specification, drafters can group similar types of behavior while also distinguishing instances where differential treatment is or might be appropriate. But there are limits. With every new distinction or caveat come the costs associated with looking up, learning, and complying with relevant rules (Tullock 1995; Kaplow 1995). Even for drafters, too many trees make it difficult to see the forest, and even more difficult to predict how cutting down or planting new trees will affect the whole. These costs are borne by everyone, either in traditional terms, or more indirectly in the form of poorly-working legal systems that are not well understood and are difficult to reform. Over-specification in the law can also lead to law's general de-legitimization (Schuck 2000) and a misallocation of human capital. At any given moment in time, there exists a finite amount of human capital in a society. Unnecessary legal complexity can drive a misallocation of that human capital toward comprehending and complying with legal rules and away from other productive ends.

Of course, one possible response to law's complexity is to make legal rules less nuanced, shorter and easier to understand. In some instances, such efforts at simplification can be achieved without imposing significant policy consequences.[1] Yet it is critical to emphasize that underspecified rules that fail to achieve their desired ends can also impose significant social costs. Many have observed that the success of law as a mechanism for social, economic and political organization is contingent on optimizing this tradeoff between complexity and simplicity, precisely because the likelihood of legal system's success is a function of its precision (Epstein 1995, 2004; Tullock 1995; Kaplow 1995).

Legal rules are one important mechanism through which the state seeks to increase social welfare. In principle, law can offer the cooperative scaffolding necessary to help solve various social dilemmas by aligning incentives, channeling

[1] Recent evidence points to a potential bipartisan political constituency in favor of at least basic overtures toward simplicity. H.R. 946: Plain Writing Act of 2010 (Signed by President Obama on October 13, 2010) is designed "[T]o enhance citizen access to Government information and services by establishing that Government documents issued to the public must be written clearly, and for other purposes".

behavior and forcing actors to internalize the cost of their respective actions. Justifications such as these are often presented as a basis for the imposition of law. To be effective, it stands to reason that the relevant regulatory apparatus must minimally reflect the nature and intricacy of social and economic exchange in the society to be governed (Rook 1993). In other words, the question of complexity is really a question of necessity. Given a society and a set of normative preferences, how much complexity in the means is necessary to achieve law's desired ends?

This question of necessity is particularly challenging, as law must operate in both a static and dynamic context, both in the moment and across time. While current circumstances typically dictate how legal rules are crafted, innovation in social interaction, economic exchange, and political behavior yield an increasingly complex world whose conduct the law is subsequently called upon to regulate. Thus, as a society or economy becomes more complex, legal rules arguably must in some way adapt to match this complexity. There is likely some form of '*scaling*' relationship. While this idea is clear in theory, when one observes a society and its complex body of legal rules, it is difficult to objectively assess the necessity of this complexity. Observed legal complexity may be driven by a genuine effort to keep pace with ongoing developments in society. Alternatively, it may only be the by-product of politicians' efforts to deliver particularized benefits to specific individuals or interest groups.

Politicians may seek to maximize their own reelection function by modifying legal rules in a manner consistent with the preferences of their core constituency. As a general matter, that constituency is interested in their specific goals and thus does not have a sufficient interest in the impact that this rent-seeking behavior imposes upon the overall complexity of the law. While there is sometimes mention of legal complexity in political discourse, the pursuit of a less complex legal system rarely extends beyond mere rhetorical bluster.

This implies that legislative processes have a directional bias or drift. Indeed, it is rarely the case that legal systems become complex overnight. Instead, through a time-evolving process, institutions slowly creep toward complexity. Even if each addition, markup, amendment or particularized modification to existing law represented only a slight addition to its overall complexity, over longer windows of time, the aggregate impact of these small movements is a body of legal rules that is exceedingly unwieldy.

The consideration of these and other questions would be enhanced by both a conceptual framework and an empirical strategy designed to better understand legal complexity. Though both academic scholarship and policy rhetoric have often invoked "complexity" to argue for or against some particular outcome, many of the existing claims regarding complexity would benefit from a more rigorous treatment of the question. In other intellectual domains, scholars have increasingly advanced our collective understanding by leveraging computational tools and associated methods to evaluate a variety of previously underexplored questions (Lazer et al. 2009). Such tools should be applied to study legal systems (Ohm 2009). Applying an analogous approach, we focus upon the specific task of developing complexity measures for legal rules.

The scientific study of complexity is itself a complex question. (Mitchell 2009; Page 2008; Bonanno and Collet 2007; Stoop et al. 2004; Feldman and Crutchfield 1998; Bates and Shepard 1993; Landauer 1988; Lloyd and Pagels 1988). In turn, the complexity of legal systems is, unsurprisingly, also complex. However, several worthy efforts have been undertaken (Bourcier and Mazzega 2007a; Slemrod 2005). In the existing literature discussing legal systems, there are many inconsistent uses of the term "complexity". For our purpose, we focus upon the human capital expended by a society when an end user is required to review and assimilate a body of legal rules. While this is hardly the only version of legal complexity, it does represent a rough conception of the typical understanding of legal complexity as invoked by lay people.

Using our specific view of complexity, this paper measures the complexity of one particular body of law—the United States Code. Organized into Titles, including the commonly referenced Title 11—Bankruptcy and Title 26—Internal Revenue Code, the United States Code represents a large and substantively important body of law familiar to many legal scholars and at least some laypersons. In published form, it contains hundreds of thousands of provisions and tens of millions of words. Though the Code is clearly complex, measuring this complexity is a non-trivial task. To do so, we borrow concepts and tools from a range of academic disciplines, including computer science, linguistics, physics, and psychology.

Our conceptual framework is centered upon a hypothetical individual engaging in a knowledge acquisition process. Knowledge acquisition, a field at the intersection of psychology and computer science, studies the protocols individuals use to acquire, store, and analyze information (Iria 2009; Schnotz and Kürschner 2008; Ferstl and Von Cramon 2007; Cimiano et al. 2005; Sanderson and Croft 1999; Kintsch and Van Dijk 1978). Using ideas developed in this field, we develop and apply an acquisition protocol for the Code. In this context, many in the literature commonly use the words "cost" and "complexity" interchangeably. We follow this idea and conceptualize complexity as the cost of carrying out the acquisition protocol. Our protocol contends that the three primary qualitative features of the Code that contribute to its complexity are *structure*, *language*, and *interdependence* as identified in the literature (Bourcier and Mazzega 2007a; Bommarito and Katz 2010).

The U.S. Code is a document structured as a hierarchical network or tree. The depth of an element in this tree typically corresponds to its level of detail or specificity. Not surprisingly, the active Titles of the Code sit at the top of the tree, while hundreds of thousands of provisions are organized in the many branches below. These provisions also contain language, and, in total, the Code features millions of words with varying lengths and diverse meanings. Finally, citations from one provision to another are often contained within the text of various provisions. These citations represent interdependence and allow sections to reference definitions or processes contained elsewhere within the Code. Taken together, these references form a citation network that is not constrained by the hierarchical structure of the Code (Bommarito and Katz 2009).

Having developed this conceptual framework, we use it to empirically measure the structure, language, and interdependence of the Code's active Titles.

Throughout this paper, we utilize an XML representation of the forty-nine active Titles of the U.S. Code in 2010.[2] We use measurements based on this data set to calculate a composite measure that offers a comprehensive score for the relative complexity of each of these Titles. This composite measure simultaneously takes into account contributions made by the structure, language, and interdependence of each Title through the use of "weighted ranks". Weighted ranking is an approach commonly used to pool or score objects with multidimensional or nonlinear attributes. We support the use of a "weighted ranks" framework, as it is flexible, intuitive, and entirely transparent, thus allowing other researchers to quickly replicate or extend our work. Using this framework, we provide simple examples of empirically supported claims about the relative complexity of Titles.

Beyond its application in measuring Titles or the Code at any given scope, we believe that this high level framework is useful for many academic and policy questions. In law and social science, this framework can contribute a common set of words and tools to comparative and normative analyses of complexity. In policy spheres, this framework can provide a concrete mechanism for evaluating and comparing single pieces of legislation, snapshots of documents over time or even the outputs of entire legal systems. Though in this paper we focus only on the Code at a single time, the framework developed herein can be successfully applied in order to longitudinally measure the Code or any of its pieces.

In support of these goals, the balance of this paper is structured as follows: Sect. 2 describes in detail the Code and the process that produces it; Sect. 3 outlines prior research on complexity, develops our knowledge acquisition protocol, and formalizes our conceptual framework for legal complexity; Sects. 4, 5 and 6 carry out measurement of the structure, language, and interdependence, respectively; Sect. 7 pools these measurements into a composite measure of complexity for each of the active Titles of the Code; and finally, Sect. 8 concludes the paper and outlines directions for future research.

2 The United States Code

The United States Code is the substantively important corpus that constitutes the compiled federal statutory law of the United States. In published form, the Code spans many volumes and contains hundreds of thousands of provisions and millions of words. This content is organized as a *hierarchical document*, with Titles serving as the initial unit of organization. The Code features forty-nine active Titles, including well-known Titles such as Title 26—Internal Revenue Code, Title 11—Bankruptcy, and Title 18—Crimes and Criminal Procedure. Beyond the initial partitioning, the Code carves out additional topical subdivisions resulting in units of decreasing size. Although the specifics vary, each Title begins with a "root node" and its broad topical

[2] This data set was provided by the Cornell Legal Information Institute and can be accessed at http://hula.law.cornell.edu/uscode_xml_dist/usc-xml-2010-10-28/. The United States Code features a total of fifty Titles. However, *Title 34—Navy* has been repealed. With the recent approval of *Title 51—National and Commercial Space Programs* the United States Code will once again feature a total of fifty active Titles. All code and additional replication materials are available here https://github.com/mjbommar/us-code-complexity.

label. Next, it includes one or more of the following hierarchical elements: Subtitle, Chapter, Subchapter, Part, Subpart, Section, Subsection, Paragraph, Subparagraph, Clause, or Subclause. In many instances, the hierarchical elements produce Titles with elaborate structures. As an example, Fig. 1 both hierarchically visualizes the structure of Title 2 (The Congress) and highlights up to ten discrete hierarchical layers of provisions that fall below the Title 2 "root node".

Although end users[3] engage the Code at a given moment, it is a dynamic rather than static object. The United States Code is a constant *work-in-progress* and its observed changes are driven by at least two distinct processes: a legislative process and a lesser-known codification process. There is a host of scholarship devoted to the mechanics of the legislative policy-making process, including the fields of public choice (Black 1948; Arrow 1963; Buchanan and Tullock 1965; McKelvey 1976; Ostrom 1990) and legislative politics (Riker 1962; Achen 1978; Becker 1983; Weingast and Marshall 1988; Poole and Rosenthal 1991; Cox and McCubbins 2007; Ansolabehere et al. 2001).

Indeed, analyses of actions leading to the provision of public goods have a rich history. While the classic scholarship emphasizes questions of preference aggregation under various institutional frameworks, (Harsanyi 1955; Downs (1957); Arrow 1963; Sen 1970; Gibbard 1973; McKelvey 1986; Austen-Smith and Banks 1996) more recent work leverages principles from computer science and physics to consider how observed policy choices are affected by a series of complex interactions by heterogeneous agents operating on a time-evolving landscape (Rothkopf et al. 1998; Kirman and Zimmermann 2001; Shoham and Leyton-Brown 2009). Such agents include individuals, organizations, corporations and even other countries, all encouraging the relevant political actors to support their preferred policy outcomes.

In much the way Ronald Dworkin (Dworkin 1986) described the common law as a chain novel, it is important to emphasize the temporal and decentralized process that generates the United States Code. At any given moment, the United States Code represents the aggregation of policy choices made by various Congresses operating under different political and economic conditions. While the Code could, in principle, be completely reset by each respective Congress, as a practical matter, drafters of legislation typically build upon the existing document by writing legislation that edits the existing corpus in a manner consistent with the drafters' preferences. With an aggregate document produced under such decentralized conditions, even under ideal circumstances, complete internal coherence is unlikely to be obtained. Rather, complexity, inconsistency and indeterminacy are almost certain to follow.

While the vast majority of the policy offerings proposed by legislative actors do not survive, those that do persist are enacted into positive law. Following passage, the final version of legislation is submitted to the Government Printing Office (GPO) and is distributed as a slip law[4] (Tress 2009). Along with other sources such

[3] End users of the Code are actors who interact directly with its text. End users include not only sophisticated parties, such as lawyers and lawmakers, but also laypersons, public interest groups, and businesses.

[4] A "slip law" is the first print of a new law in pamphlet form, usually available 2–3 days after enactment. The Government Printing Office (GPO) offers a useful description of this process *see* http://www.gpoaccess.gov/plaws/about.html.

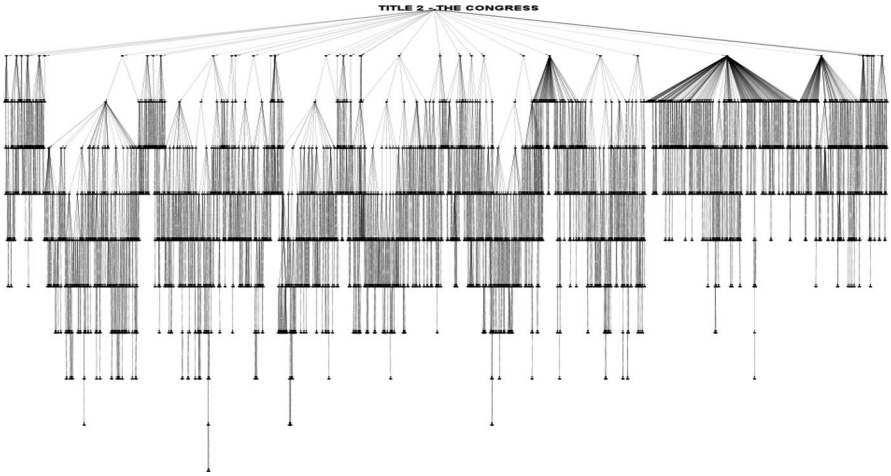

Fig. 1 Hierarchical visualization of Title 2—The Congress

as private laws and treaties, slip laws are published annually in order of passage in the *Statutes at Large* (Tress 2009). While the *Statutes at Large* are the law of the United States, their chronological ordering does not allow for convenient or categorical access by interested parties. To aid those who seek access to federal law, the U.S. House of Representatives—Office of the Law Revision Counsel (LRC) compiles the *Statutes at Large* into the United States Code. In this compilation, the LRC removes expired provisions, incorporates amendments and groups elements with similar topics into hierarchical elements of increasing detail. Though elements at or below the section level may contain text, the subdivisions above the section level such as Subtitles, Chapters and Subchapters are purely containers for elements and feature no substantive text.

Although often cited by judges and lawmakers, as a technical matter, the United States Code is merely *prima facie* evidence of federal law. In the event that a discrepancy between the United States Code and the *Statutes at Large* develops, the latter serves as the ultimate authority. Disputes between these sources are rare and given its useful organization, most scholars, judges and practicing lawyers consult the United States Code. While the Code is a repository of compiled federal statutory law, it is important to note that it does not represent the exclusive repository of federal legal materials. Specifically, other valid sources of law including regulations offered by administrative agencies such as the Environmental Protection Agency, the Food and Drug Administration and the Treasury Department supplement the Code and are published chronologically in the Federal Register.[5] Notwithstanding

[5] While we do not incorporate these regulations into our analysis, we recognize that their incorporation would paint a more complete picture of the relevant legal landscape. Through a process similar to the compilation of the United States Code, federal regulations are subsequently compiled by topic in the *Code of Federal Regulations* (C.F.R.). Of course, administrative regulations and the United States Code are not the only sources of federal legal materials. There also exist additional materials such as judicial decisions, executive orders, revenue rulings, etc.

the existence of these additional sources of federal law, we believe the significant scope and attractive properties such as compilation and substantive organization make the United States Code an appropriate corpus with which to explore the measurement of legal complexity.

3 Complexity, knowledge acquisition, and the U.S. Code

Using the United States Code as an emblematic example of a large and substantively important body of law, we develop both a conceptual and empirical framework designed to measure its complexity. The complexity of the law is a topic long considered by scholars. (Long and Swingen 1987; Schuck 1992; White 1992; Flournoy 1994; Feltovich et al. 1995; Kaplow 1995; Tullock 1995; Epstein 1995, 2004; Kades 1997; Wright 2000; Bourcier and Mazzega 2007a; Ruhl 2008; Phelan 2009; Frisch 2011). Although it is a question with important positive and normative dimensions, the study of legal complexity has arguably taken a narrow approach to the question. As noted by Professor Schuck "[F]or all the broad-gauged interest in legal complexity, legal scholars have largely confined their analyses of this phenomenon to two aspects: (1) its transaction costs, including legal uncertainty; and (2) certain of its sources, especially litigation incentives, judicial decisions, and rule form" (Schuck 1992). In response to the state of the scholarship, "…one is struck by [its] narrowness. Legal complexity, after all, has consequences that go well beyond transaction costs, and has sources besides litigation incentives and judicial lawmaking. Its forms, costs, structural causes, limits, and possible reform raise important social issues that need to be explored in greater depth and breadth" (Schuck 1992). These themes are later echoed in a subsequent survey of legal complexity, "[L]egal scholars have not had an easy time defining complexity, and some have been disarmingly honest about this difficulty. One author admitted that he was tempted to define complexity by averring, 'I know it when I read it'" (Kades 1997).

In addition to the large body of traditional legal scholarship, there has been a variety of theoretical and applied work that has either conceptualized or attempted to measure the complexity of the law. On the theoretical side, two intellectual communities have displayed a particular interest in legal complexity. Specifically, scholars in the subfields of *law and economics* (White 1992; Kaplow 1995; Tullock 1995; Epstein 1995, 2004; Kades 1997; Frisch 2010) and *artificial intelligence and* law (Bibel 2004; Pagallo 2010; Bourcier and Mazzega 2007a) have offered contributions to various dimensions relevant to the broader question.

While most of the theoretical work has been concentrated in a couple of subfields, the applied scholarship asserting or considering the complexity of the law has been quite varied with subject matter experts in a number of substantive fields making contributions. Among this broader set, tax features some of the most significant efforts, to date, to consider and measure legal complexity (Surrey 1969; Long and Swingen 1987; Eustice 1989; Schenk 1989; Koppelman 1989; McCaffery 1990; Paul 1997; Cecil 1999; Donaldson 2003). Despite this success, early work in the tax subfield took a pessimistic view about efforts at measurement. For example,

an early commentator noted "neither tax 'simplification' nor its mirror image, complexity, is a concept that can be easily defined or measured. I know of no comprehensive analytic framework for these ideas, nor any empirical study that supplies a 'simplicity index' in particular areas of tax law practice" (Bittker 1974). The intervening years have witnessed a number of notable attempts to measure the complexity and compliance costs associated with the tax law (Long and Swingen 1987; Pitt and Slemrod 1989; Slemrod 2005). Indeed, with respect to measurements of complexity, tax scholarship arguably offers the most concrete attempts to determine the complexity of a written body of law. Notwithstanding, it is clear that even within the sub-field of tax significant work remains.

While scholars have identified various factors associated with complexity, there is little agreement on relevant standards and/or a reasonably concrete approach to measurement. Indeed, outside of tax virtually all dimensions related to law's complexity are still yet to be comprehensively measured. These sentiments are also echoed by one commentator who notes "[t]he empirical work on legal complexity has been relatively limited." (Barton 2008) Surveying the literature, there is a general lack of agreement and a variety of inconsistent definitions of complexity offered throughout the literature. (Barton 2008)

Undoubtedly, law's complexity has many facets. Given the scope and difficulty of the problem, it is thus important for researchers to identify the components of the question they are considering. In this vein, our analysis does not directly consider various costs associated with legal compliance, such as record keeping costs (Slemrod and Blumenthal 1996). We do not consider litigation costs including those generated by statutory ambiguity or other forms of legal uncertainty. While these classes of transaction costs are important and worthy of study, uncertainty is not conceptually equivalent to complexity (Page 2008). In terms of considering complexity, we believe developing measures that simulate the structure and content of the relevant written bodies of law is an appropriate place to begin any form of broader empirical analysis.

In an effort to remedy some of the conceptual confusion and return to first principles, we consider the decision making process of a hypothetical end user experiencing the Code and determining whether to comply with its mandates. For a given individual (or their learned intermediary), the initial step on the path to compliance (or non-compliance) is determining the particular legal rule, if any, which touches upon the given matter in question. Yet, this is only one step of a much broader process—a process that includes various components such as a heuristic or data informed assessment of whether compliance is net beneficial. The following steps roughly memorialize the process generating the observed level of compliance behavior.

Rule Search	How complex is the task of determining the rule or set of rules generally applicable to the conduct in question?
Rule Assimilation	How complex is the process of assimilating the information content of a body of legal rules?

Rule Uncertainty	How much latent uncertainty is contained within the rule?
Cost Assessment	How costly is it to comply with the rule?
Comply or Not	On balance, do the respective costs and benefits favor compliance or non-compliance?

With respect to the complexity of the United States Code, our analysis is devoted to a segment of this broader process, most notably rule search and rule assimilation. We believe that together the *structure*, *language* and *interdependence* of the Code collectively impact the complexity of the law as experienced by an end user seeking to determine whether certain conduct is covered by a particular legal rule. In order to formalize this experience, we have chosen to model the Code as the object of a knowledge acquisition process (Boose 1989; Boose and Gaines 1990; Spiro and Jehng 1990; Holsapple et al. 2008; Francesconi 2011). As noted earlier, the study of knowledge acquisition is an important field that combines features of psychology and computer science and is primarily focused on how individuals acquire, analyze, and store information. A common mode of analysis in this subfield is to study the *protocols* and *schemas* that individuals develop as they learn to perform difficult tasks or acquire complex information (Sweller and Chandler 1994; Pollock et al. 2002; Halford and Busby 2007). Such protocols are akin to heuristics or inductive rules that individuals employ to handle complex environments.

Reflecting upon the contours of a knowledge acquisition protocol for the United States Code, we formalize the experience of an individual trying to obtain information about a particular element of the Code. Whether applied to the entire Code or at a more selected scope, our goal is to reasonably simulate the experience faced by our hypothetical end user. Though we will later provide a complete ontology of the Code's elements, here an element could correspond to a portion of the Code of any size. For example, an end user interested in learning broadly about tax could begin by consulting *Title 26* as the initial element of analysis while an individual interested in filing their personal income taxes might consult *Title 26, Subtitle A—Income Taxes*. Focusing the scope of resolution, an individual interested in the specific rules associated with tax exempt organizations might consult 26 U.S.C. § 501. In all of these scenarios, the steps associated with acquiring the relevant knowledge are reflected in the following protocol:

1. Select an initial element of the Code corresponding to a concept of interest.
2. Beginning from this initial element, recursively assimilate the content of all sub-elements.
3. When a citation is encountered, apply this protocol recursively to the cited element.

In applying this acquisition protocol, the complexity associated with acquiring knowledge is driven by three qualitative features of the United States Code:

structure, *language* and *interdependence*, each of which we describe in *infra* Sects. 4–6.

4 Structure: the U.S. code as hierarchical document

The basic structural feature of the United States Code is inherited from its representation as a hierarchical object. In this hierarchy, the Office of Law Revision Counsel organizes the Code so that similar topics are grouped together. While their groupings are likely imperfect in some respects, the general principle is evidenced in the division of the Code into Titles and smaller substantively related elements such as Chapters or Parts. Although we recognize many plausible alternative orderings could, in principle, be developed, we hereafter assume that the hierarchical divisions and groupings created by the Office of the Law Revision Counsel generally correspond to hierarchical divisions within the underlying concepts.

The selection of a hierarchical form of organization allows related content to be grouped and set forth at increasing levels of resolution. Following the work of Bourcier and Mazzega (2007b); Bommarito and Katz (2010); Boulet et al. (2011); Mazzega et al. (2011), we formalize the Code as a tree with more distant levels of the tree typically addressing more specific concepts. Each Title or other non-terminal element of the Code, e.g., Chapter, can be separately represented as a disjoint "subtree". For example, Fig. 2 below offers the complete tree based representation of the United States Code's smallest Title, *Title 9-Arbitration*.[6] Within Fig. 2, each circle or node corresponds to an **element** of the respective Title. All trees have one **root** element and *at least one* **leaf** element. In Fig. 2, the root element or "root node" is found at the top of the tree and is labeled "Title 9". In a tree style representation, leaf elements are terminal points that occur at the end of branches, such as 9 U.S.C. § 9 or 9 U.S.C. §16(a)(1)(A) as labeled below.

In addition to differentiating between root and leaf elements, there are a number of additional substantive distinctions between elements of the Code. In the codification process discussed in Sect. 2, *infra*, the hierarchy chosen imposes labels onto elements that reflect their resolution with respect to the Title. These labels vary across Titles, but as described earlier, any given element will be one of the following: Title, Subtitle, Chapter, Subchapter, Part, Subpart, Section, Subsection, Paragraph, Subparagraph, Clause, or Subclause. Thus, Title 9 is divided into three Chapters: *Chapter 1—General Provisions*, *Chapter 2—Convention on the Recognition and Enforcement of Foreign Arbitral Awards*, and *Chapter 3—Inter-American Convention on International Commercial Arbitration*.[7] Under each of these Chapters, there are also a number of additional subdivisions including sections and subsections. Sections are of particular importance because they are both the first level at which substantive text appears and the first level at which the hierarchy can

[6] As Title 9 is the smallest Title in the United States Code, it allows us to clearly indicate these distinctions that would otherwise be obscured by the size of the tree for other Titles.

[7] While Chapter 1 is explicitly labeled, the remaining Chapters are located at the same horizontal level of the hierarchy.

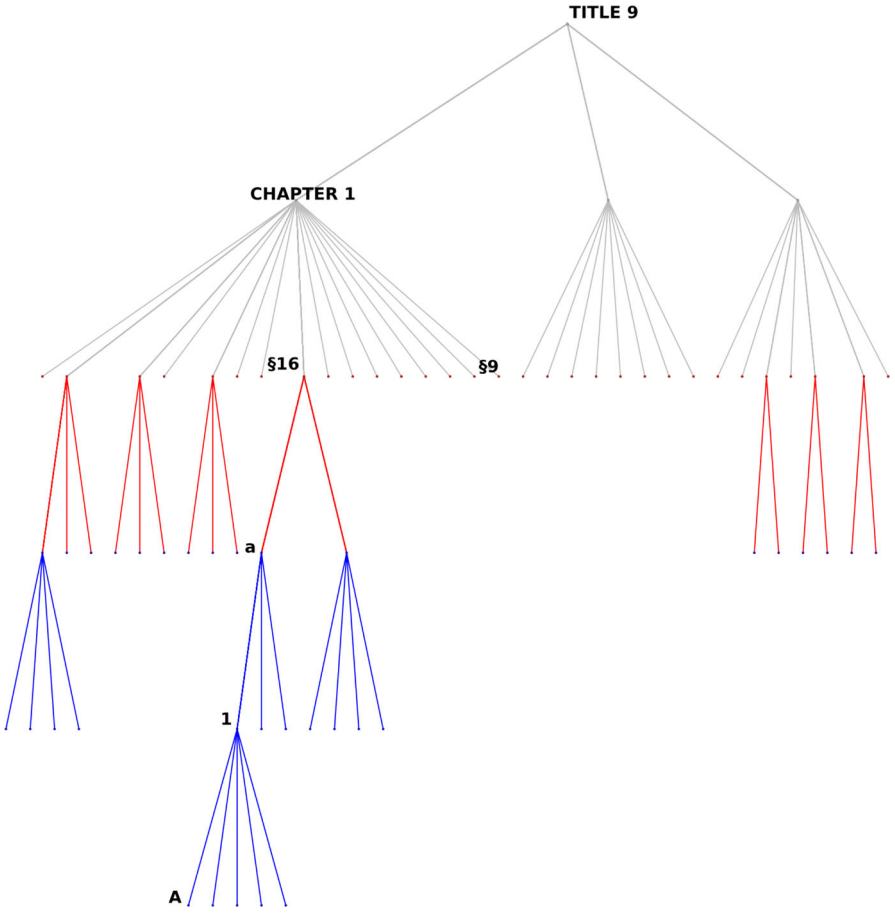

Fig. 2 Hierarchical visualization of Title 9—Arbitration

terminate. Sections, however, need not be the point of termination for the tree. In Fig. 2, for example, 9 U.S.C. § 9 is a leaf node, 9 U.S.C. §16 is *not* a leaf node because it features a number of elements below the section level such as 9 U.S.C. §16(a), 9 U.S.C. §16(a)(1), or 9 U.S.C. §16(a)(1)(A).

To formalize these structural classifications, we distinguish between elements above the section level, section elements, and elements below the section level. As noted above, elements above the section level, such as Titles or Chapters, do not contain substantive text and are denoted by V^*. Elements at the section level such as 26 U.S.C. §501 or below the section level such as 26 U.S.C. §501(c)(3) typically contain text, and are denoted by V^s for sections and V_* for elements below the sections. Thus, the set of all elements is given by $\mathbf{V} = V^* \cup V^s \cup V_*$.

These elements are linked by the set of structural arcs A_s. which in Fig. 2 are the lines connecting the various hierarchical elements. Using nodes and structural arcs, we can represent the structure of the nth Title as a mathematical graph $T_n = (V,$

A_s).[8] Table 1 highlights the significant scope of the Code's structure by providing raw counts for structural arcs A_s and each of the types of vertices in **V**.

A researcher could employ a number of legitimate strategies to measure the structural properties of the Code. Following the approach offered in prior scholarship (Bommarito and Katz 2010), Table 1 measures the structural size of the full United States Code at a given temporal snapshot. However, it is also possible to unpack the aggregate measure and apply it to the Code's primary sub-components (i.e., Titles). Much like the full United States Code, the structural complexity of a given Title is driven by its size, depth and the relationship between these two features. As a general proposition, we assume *ceteris paribus* increases in each of these measures typically correspond to increases in the effort associated with knowledge acquisition.

4.1 Structural size

Consistent with the approach offered in Table 1, the size of a given Title can be measured by simply counting the number of vertices $| V |$ located therein. Applying this approach to *Title 9—Arbitration* yields a count of 68 vertices $| V |$ distributed as $| V^* | = 4$, $| Vs | = 31$ and $| V_* | = 33$. Although previously displayed in Fig. 2, Title 9 offers a highly unrealistic portrait of the structural size of an average Title within the Code. For example, Fig. 1 displayed earlier is a much more representative U.S. Code Title. With $| V | = 7{,}873$, *Title 2—The Congress* is two orders of magnitude larger than *Title 9—Arbitration* and provides a glimpse of the significant variation in Title size. To capture this variation, we measure the raw size of each of the Code's forty-nine active Titles and report results for the five largest and smallest Titles in Table 2. Table 2 also subdivides and reports the total elements $| V |$ into those falling above section $| V^* |$ section $|Vs |$, and below section $| V |$. While measures for each of the forty-nine active Titles can be found in our online appendix, Table 2 demonstrates that the disparity in the structural size between the Codes' Titles is on orders of magnitude.

Both Table 2 also highlights significant differences in the ratio of all elements $| V |$ to above section elements $| Vs |$. *Title 23—Highways*, for instance, has 6 Chapters above the section level but 3,809 elements at or below the section level, indicating that each of these sections contains a large amount of structure embedded within the section and below section elements. *Title 13—Census*, on the other hand, contains 22 Chapters and Subchapters for only 250 elements at or below the section level. The disparity between T_{23} and T_{13} indicates the latter features much more structure defined in hierarchical elements above the section.

As they are units most commonly invoked in policy debates, we focus our analysis upon scoring the relative complexity of the Code's forty-nine active Titles. Given the substantial variation in the sizes of Titles, we initially explore the relationship between section and total elements to determine whether Titles are an appropriate unit to compare. Considering the boundary cases, Title 9 defines a small number of rules for arbitration, whereas Title 42 creates and manages hundreds of agencies with sizeable scope and authority. Title 9 contains $O(10^1)$ elements and $O(10^5)$ words,

[8] Since T_n is a tree as in Fig. 1, A_s must be $|V| - 1$.

Table 1 The structural elements of the United States code

Element	Notation	Count
Above section nodes	V^{*}	8,077
Section nodes	V^{s}	51,922
Below section nodes	V^{*}	522,055
Total nodes	V	582,054

Table 2 Five largest and smallest titles by structural size

Title	V	V^{*}	V_S	V_*
Public Health and Welfare (Title 42)	110,605	1,405	7,321	101,879
Internal Revenue Code (Title 26)	51,553	532	2,083	48,938
Conservation (Title 16)	33,062	473	4,704	27,885
Agriculture (Title 7)	29,191	349	2,701	26,141
Education (Title 20)	28,096	665	2,124	25,307
Arbitration (Title 9)	68	4	31	33
General Provisions (Title 1)	84	4	39	41
Flag and Seal, Seat of Government, and the States (Title 4)	221	6	47	168
Intoxicating Liquors (Title 27)	224	13	45	166
Census (Title 13)	272	22	70	180

whereas *Title 42—Public Health and Welfare* contains $O(10^5)$ elements and $O(10^7)$ words. Given that over 100 copies of Title 9 could thus be contained within Title 42, we examined the overall relationship between sections and elements to determine whether a reasonable scaling relationship exists. Our analysis indicates a fairly strong log–log relationship of the form $log_{10}(|Vs|) \cong 0.76\ log_{10}(|V|) - 0.07$ with an $R = 0.94$ and $p = 0.04$. Based upon this analysis, we believe it is appropriate, from a structural perspective, to consider smaller Titles such as Title 9 as simply scaled versions of larger Titles such as Title 42. Indeed, the presence of this scaling relationship indicates that Title trees are constructed subject to some natural constraints that in turn produce the scaling relationship. Therefore, although Titles are clearly far from identical in size, for a class of general questions, we believe it is reasonable to engage directly in comparative analysis.

4.2 Element depth distribution

The depth of an element corresponds to the hierarchical "level" on the Title tree upon which a given element lies. By assigning depth zero to the root node of the Title tree, the depth of other elements can be measured by counting the number of "steps" required to reach a given element for the Title node. Referring back to Fig. 1, we observe the initial Title 9 element has depth zero, while Chapter 1 has depth one, §16 has depth two, §16(A) has depth three, and so on. Given our assumption that the Law Revision Counsel's recursive division of concepts generally corresponds to increasing levels of detail, the depth of an element is a measure of the specificity

of its contents relative to the overall Title. Some elements, like Chapters, deal with broad concepts and occur higher in the tree (lower depth). Elements below the section, such as individual clauses, tend to focus on very specific definitions or conditions, and thus fall much lower in the tree (higher depth). For example, consider the well-known provision defining tax-exempt organizations—26 U.S.C. §501(c)(3). Working from shallow to greater depth, *Chapter 1—Normal Taxes and Surtaxes* lies at depth two, whereas §501(c)(3) lies at depth seven.

Using this concept of depth, we can generate a basic understanding of the average specificity of a Title's elements (Rook 1993). If the mean element depth is high, then as an average proposition, more of the Title's elements feature a high level of specificity relative to the Title's concept. Likewise, if the mean element depth is low, the Title's elements are generally less specific. Building from this basic idea, Table 3 offers the five largest and smallest Titles based upon mean element depth.[9] Table 3 demonstrates that the Titles with the highest levels of depth are *Title 26—Internal Revenue Code*, *Title 5—Government Organization and Employees*, and *Title 8—Aliens and Nationality*, and the Titles with the lowest mean element depths are *Title 9—Arbitration*, *Title 1—General Provisions* and *Title 4—Flag and Seal, Seat Of Government, and the States*. As a matter of external validation, both Table 3 and the online appendix[10] appear to distinguish between Titles with greater substance and those featuring more ministerial content. For example, tax scholars often highlight the intricacy of federal tax law. Thus, it is not terribly surprising to find the Internal Revenue Code features the highest mean element depth. At the same time, *Title 9—Arbitration*, as displayed in Fig. 1, unsurprisingly features the lowest mean depth.

4.3 Comparing size and mean element depth

We also explored the relationship between a Title's size and its mean element depth. Representing the forty-nine active Titles, Fig. 3 is a scatter plot memorializing the relationship between size and depth. This diagnostic figure indicates that size is, at best, weakly correlated with mean element depth. The ordinary least squares relationship has slope 1.00, intercept 1.34, an $R = 0.72$, and a p value of 0.14. While some large Titles are also the most intricate and some small Titles are the least intricate, a non-trivial number of Titles disobey this trend. For example, while Title 42 is by far the largest, it has only the tenth largest mean element depth. In addition, Fig. 3 highlights the heteroskedastic dispersion associated with the increasing size of the title.

5 Language

In addition to its structural features, the United States Code's most recognizable feature is its language. Distributed across its vast architecture, the Code features

[9] All code and additional replication materials are available here https://github.com/mjbommar/us-code-complexity.

[10] The online appendix can be access here: http://computationallegalstudies.com/measuring-legal-complexity-appendix/.

Table 3 Five largest and smallest titles by mean element depth

Title	Avg. depth
Internal Revenue Code (Title 26)	7.80
Govt. Organization and Employees (Title 5)	6.66
Aliens and Nationality (Title 8)	6.51
Education (Title 20)	6.41
Transportation (Title 49)	6.40
Census (Title 13)	3.97
National Guard (Title 32)	3.50
Flag and Seal, Seat of Govt. and the States (Title 4)	3.23
General Provisions (Title 1)	2.85
Arbitration (Title 9)	2.82

millions of words. It is, of course, these text strings within the Code that are organized and presented to the end user. The content of this language is a second factor that contributes to the complexity of the United States Code. Thus, it must be included in any serious model of its complexity.

While it is hardly controversial to argue that language increases the cost of knowledge acquisition, measuring how language contributes to the complexity of the knowledge acquisition process is, of course, challenging. Even a cursory review of the Code demonstrates that the amount of text, as well as the size and distribution of different words, can vary significantly across the Code's respective elements. To capture these differences, we offer a measure of linguistic complexity that seeks to model the "cost" of assimilating the language contained within each element of the Code. Although likely imperfect in some respects,[11] we are interested primarily in the number of words, the average length of those words and, most importantly, how the distribution of these word frequencies varies across multiple elements of the Code.[12]

Returning to the knowledge acquisition process outlined in Sect. 3, *supra*, the second step in the protocol reads "Beginning from this initial element, recursively assimilate the content of all sub-elements". Here, "assimilate" refers to the process of reading and understanding the actual text of each of these elements. The manner in which individuals synthesize textual information is a question at the forefront of research in cognitive psychology and learning theory. Without doing deep violence to research in this domain, we can think of language as imposing some "cost" on the individual, and thus we would like to measure how the actual text of an element corresponds to its corresponding per word cost in the knowledge acquisition process. We believe the "cost function" is driven not only by the volume of words

[11] One obvious weakness with our proposed measure of linguistic complexity is its failure to capture the underlying semantics. As this is an introductory effort, we would invite future work focused upon this particular dimension of the question.

[12] We acknowledge the extensive literature on text complexity (e.g., Flesch and Gould 1949; Kincaid et al. 1975; Si and Callan 2001). Much of this work, however, is directed at reading comprehension of standard sentences and paragraphs. The United States Code is a specialized document with its passages separated by the unique presentation formatting used to display statutes.

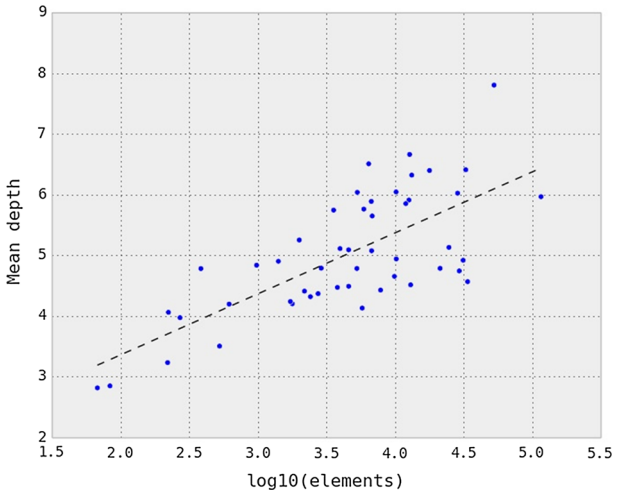

Fig. 3 Scatter of mean element depth and log_{10} (|*V*|) with *OLS*

but also by features of those words and the distribution of the word frequencies across the respective corpus.[13] After providing per-Title word counts, our analysis shifts to consideration of the various linguistic properties such as the average word length and the entropy of a Title's word distribution.

5.1 Size

To measure the size of a given Title's linguistic content, we simply count the number of tokens contained within all elements of a Title. Here, "tokens" refer to contiguous strings of text, which are often words but may also be numbers, citations, or abbreviations not formally considered words.[14] Using token counts, Table 4 shows the five largest and smallest Titles. In addition, Table 4 highlights the average number of tokens per section for each of these Titles. This table confirms the suspicion that Title 42 and Title 26 are the largest Titles by content size. Yet, Title 16 is surprisingly similar in raw size to Title 26. In a manner similar to raw structural size, Table 4 demonstrates that the cross-Title variation in token counts is on orders of magnitude with the largest Title, featuring roughly $O(10^6)$ tokens, and the smallest Title, which features $O(10^3)$. In addition, Table 4 demonstrates a significant amount of variation in average tokens per section with Title 26 featuring an average of 487 tokens while Title 9 contains only 80 tokens per section.

[13] We offer this as a *ceteris paribus* proposition across the millions of words contained in the United States code.

[14] We selected tokens rather than other alternative length measures, such as pages, as we believe these are far less likely to be impacted by formatting conventions.

5.2 Average word length

When considering the process of assimilating the information contained within a given element of the Code, the volume of words is important but by no means the exclusive property driving its complexity. One important refinement is to consider the average length of the words contained within each Title. As a general proposition, we believe longer words are more difficult to comprehend. However, the relationship is not strict as word utilization rates can impact the complexity as experienced by the end user. While word length is often inversely proportional to frequency of use, there are a number of short words that are used infrequently and a number of longer words that are easily understood by the average reader.[15] While we believe length is often associated with complexity, we recognize the assumption does not always hold in every discrete instance. However, given the scope of information in question, we believe that as a basic proposition, Titles with longer average word length will prove more costly to review.

In order to measure the average word size across Titles, we again tokenized the text and generated a "bag of words" distribution for every section within a Title. Since we are interested in the length of contextually relevant nouns and verbs, we remove "stop-words" such as "and," "or," and "but" from the text (Manning et al. 2008). These "stop-words" are typically short words that decrease the average word length in proportion to the amount of grammar in the text. The resulting "bag of words" therefore represents the relevant nouns and verbs in each Title. Using this approach, Table 5 highlights the ten Titles with the highest average word size. These Titles with the highest average word size seem to represent some of the more technical and specific Titles in the Code. Titles 6 and 50 pertain to security and defense, whereas Titles 44 and 41 cover Education and Public Contracts, respectively. In general, a review of Table 5 demonstrates the variation within these Titles is fairly minimal. Thus, while this measure seems intuitive, it does not actually meaningfully distinguish between Titles.

5.3 Word entropy

In addition to the volume and average word length of each Title, there exists additional cross-Title linguistic variation that is important to capture. Specifically, we are interested in the diversity of language and concepts within Titles as, all else equal, it is more difficult for an individual to assimilate information in a corpus with high concept variance than one comprised of largely homogeneous material. Although a number of potential measures could be employed to consider the question of language and content variation, we borrow the concept of *entropy* from the field of information theory (Shannon 1948, 1951; Jaynes 1957; Kolmogorov 1965; Hamming 1986; Landauer 1996) as we believe that it is well suited to our question.

[15] There exist additional potential complications. For example, in some instances, longer words are more specific and thus their use can result in less ambiguity.

Table 4 Five largest and smallest titles by token count

Title	Tokens	Tokens per section
Public Health and Welfare (Title 42)	2,732,251	369.22
Internal Revenue Code (Title 26)	1,016,995	487.07
Conservation (Title 16)	947,467	200.48
Commerce and Trade (Title 15)	773,819	336.88
Agriculture (Title 7)	751,579	274.00
President (Title 3)	7,564	120.06
Intoxicating Liquors (Title 27)	6,515	144.78
Flag and Seal, Seat of Govt. and the States (Title 4)	5,598	119.11
General Provisions (Title 1)	3,143	80.59
Arbitration (Title 9)	2,489	80.29

Table 5 Ten titles with highest average word size

Title	Avg. word size
Domestic Security (Title 6)	6.90
War and National Defense (Title 50)	6.83
Public Printing and Documents (Title 44)	6.74
Foreign Relations and Intercourse (Title 22)	6.74
Public Contracts (Title 41)	6.73
Crimes and Criminal Procedure (Title 18)	6.16
Intoxicating Liquors (Title 27)	6.15
Internal Revenue Code (Title 26)	6.10
Flag and Seal, Seat of Govt. and the States (Title 4)	6.10
Bankruptcy (Title 11)	6.07

In information theory, the use of entropy has a long intellectual history. While entropy measures have roots in thermodynamics, the use of entropy traces back to the early and influential work of Claude Shannon (1948, 1951). Shannon's work has proven important to a number of fields including cryptography, machine learning and artificial intelligence (Nigam et al. 1999; Lall et al. 2006; Bose 2002; Ganapathi et al. 2012). In addition, entropy has enjoyed appeal in other domains where information processing is relevant including physics, statistics, economics, as well as many other related disciplines (Jaynes 1957; Csiszar 1991; Golan et al. 1997; Soofi 2000; Schennach 2005; Tang et al. 2008).

Entropy is a statistical measure designed to characterize the uncertainty or variance of a signal, message or body of text. The basic intuition that underlies entropy is related to prediction. In a signal environment, prediction of the future information content of a message is more difficult when that message is drawn from an environment with greater variation. Imagine an individual were to observe a certain percentage of a full message and were then interested in predicting the

balance of the message. Entropy measures what percentage of the information content is contained within the partially observed message. In the low entropy environment prediction is far more likely than in an environment with high entropy.

While the notion of entropy is relatively straightforward, it has led to significant advances in engineering, computer science, and the physical sciences. Consider the case of compressing images such as digital photographs.[16] From a data compression standpoint the two images displayed in Fig. 4 represent the low entropy and high entropy boundary cases. The image on the left requires the least memory to store as it is a canvas comprised of a single color. This pseudo-code for compression relies upon two items: (1) a numerical value corresponding to the color of any single pixel[17] and (2) the dimensions of the canvas. From an entropy perspective, this is a "uniform signal" and its reduced form representation is trivial. At the other end of the spectrum lies a picture generated by a "random signal." In this instance, each individual pixel on the canvas is assigned a random color.[18] As displayed in Fig. 4, this random pixel image offers the opportunity for only minimal compression as its reduced form representation closely approximates its original representation.[19]

What is true for image compression also holds in the case of text-based messages. The more information that is known about what a given message source will produce, uncertainty will lower, entropy will lower, and less information can be obtained from additional segments of the message. In the case of a randomly encoded message such as "orange in of going the not large kick more end to …" prediction is difficult as the previously received signals offer little insight into the content of future signals.[20] By contrast, a uniform message such as "dog, dog, dog, dog, dog…." contains a single repeated signal. Observation of any slice of the signal provides fairly clear insight as to the content of future messages.[21]

As displayed in Fig. 5, each possible message falls somewhere on the spectrum between a uniformly encoded signal and a randomly encoded signal. Entropy is a measure designed to identify where a given message falls on this continuum. Given that the content of most language is neither uniform nor random, most messages are quite distant from either of these boundary cases.

As applied to a distribution of words and concepts such as those contained within the Code, higher entropy indicates that it is harder to predict the language and

[16] It should be acknowledged that compression ratios are a common alternative to entropy measures. However, due to the large variation in compression algorithms and their implementation-specific behaviors, we felt that simple Shannon entropy was a more reproducible measure than compression.

[17] This is the *red, green, blue* or RGB value. A pure black canvas has an RGB value = #000000.

[18] In the random signal case, each pixel is assigned a random color assignment. The pseudocode for this assignment requires a randomly generated string of numbers where the assigned number corresponds to an RGB value and the length of the string is equal to the number of pixels on the canvas.

[19] In expectation, given an initial random assignment of pixel colors and a reasonably large canvas, there is likely to be at least some clustering of RGB values. This implies that at least some form of reduced representation is possible. However, this compression will be nominal.

[20] In the context of message compression, the fragment "orange in of going the not large kick more end to …" does not easily lend itself to reduced form representation.

[21] In the case of the uniform signal, the first fragment "dog" is the only new information content that is imparted to the end user. With only the first fragment and the total length of the message the signal could be quickly compressed.

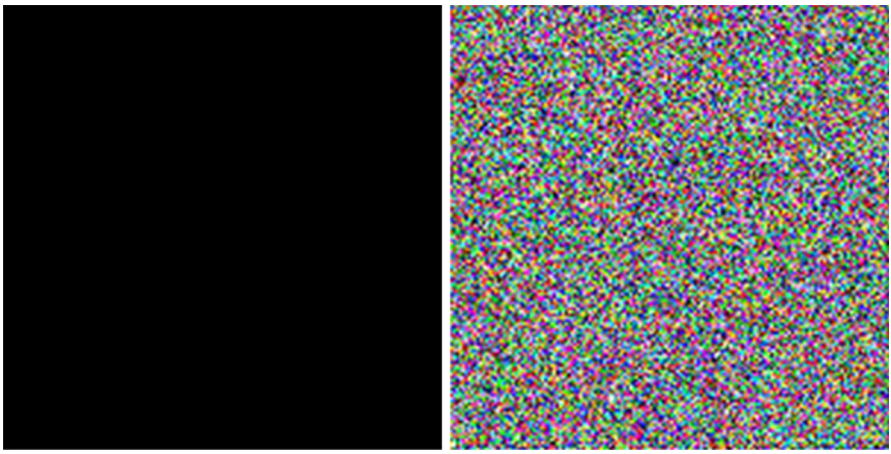

Fig. 4 A uniform image and a randomly generated image

concepts that will be contained in a given sample of the Code. Distributions with lower entropy are likely more cohesive, whereas distributions with higher entropy are likely less cohesive. When an individual is reading a Title with higher entropy, that individual is less likely to have already been exposed to a randomly sampled word or concept from that Title. In turn, higher entropy indicates the end user will more often encounter new language and new concepts. We believe measuring entropy is useful as it likely captures how linguistic diversity contributes to the cost of knowledge acquisition.

In order to calculate the Shannon entropy of a Title,[22] we tokenize the entire text of a Title and store these tokens in a "bag of words". Next, we remove all tokens known as "stopwords", such as "the" "it" and "am".[23] These words serve primarily grammatical purposes and do not represent concepts. Thus, their presence in the distribution can artificially skew the results. For all remaining tokens W in this "bag of words" for each Title, we then calculate the empirical probability of each token's occurrence $P(W = w) = p_w$. With a probability distribution for each token within each Title, we can then calculate the Shannon entropy where Shannon entropy is given by the following formula:

$$-\sum_{w \in W} p_w \log_2(p_w)$$

Applying the Shannon entropy measure, Table 6 highlights the five Titles with the highest levels of entropy, and the five Titles with the lowest levels of entropy.

[22] While a number of alternative and more sophisticated forms of entropy exist, the original Shannon entropy measure is the most straight-forward measure and is still commonly used in the information science literature. Thus, for the purpose of comparing the distribution of words within Titles, we apply the Shannon entropy. For additional work on entropy *see* Tsallis (1988) and Rényi (1961).

[23] Following upon common practice in the field of information retrieval and computational linguistics, we use the stopword list from the Natural Language Toolkit (NLTK) available at http://www.nltk.org/.

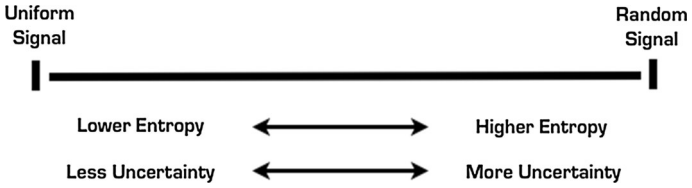

Fig. 5 The entropy spectrum

Upon qualitative review, entropy appears to succeed in separating Titles with central clustered topics from those embracing a far more diverse set of subjects. For example, the Titles with the highest entropy are *Title 15—Commerce and Trade*, *Title 42—Public Health and Welfare*, and *Title 16—Conservation*. Title 15 is an excellent example of a high entropy Title, containing over 100 Chapters ranging from Children's Bicycle Helmet Safety and the Year 2000 Computer Date Change to Collection of State Cigarette Taxes and the Transportation of Firearms. Title 42 displays a similar pattern as it contains over 150 disparate Chapters such as Leprosy, Social Security, National Flood Insurance, United States Synthetic Fuels Corporation and International Child Abduction Remedies. Although a number of these Chapters are nominally related to the umbrella topic "Public Health and Welfare," at closer inspection the full Title appears bound together by little more than the binding.

6 Interdependence

The third factor that contributes to the complexity of the United States Code is the interdependence between its many elements. This interdependence is explicitly indicated through the inline citations contained within its various provisions.[24] Consistent with the acquisition protocol outlined in Sect. 3, *supra*, an end user who attempts to review the Code or even selected portions thereof will likely encounter a citation to another segment of the Code. These citations contribute to the cost of knowledge acquisition as an individual following the acquisition protocol must expend effort traversing the citation network[25] and incorporating the information contained within any of the cited provisions. In a non-trivial number of cases, the protocol results in an extended "walk" across the United States Code when the cited provision also contains a citation.[26]

[24] There has been a significant amount of recent work on statutory citations including but not limited to Bourcier and Mazzega (2007b); Bommarito and Katz (2010); Boulet et al. (2011); Mazzega et al. (2011).

[25] Of course, if an element contains no citations whatsoever, then the protocol above collapses to only the first two rules. However, given many elements of the Code do contain citations, we embed this consideration into our analysis.

[26] This "walk" is by no means a random walk. Rather, it could better be described as a special case of graph traversal. These extended citation paths can grow to be quite lengthy. The maximum path length from 46 USC §51510 and 7 USC §87e requires thirty-two separate steps to complete.

Table 6 Five highest and lowest titles by word entropy

Title	Word entropy
Commerce and Trade (Title 15)	10.80
Public Health and Welfare (Title 42)	10.79
Conservation (Title 16)	10.75
Navigation and Navigable Waters (Title 33)	10.67
Foreign Relations and Intercourse (Title 22)	10.67
Intoxicating Liquors (Title 27)	9.01
President (Title 3)	8.89
National Guard (Title 32)	8.50
General Provisions (Title 1)	8.49
Arbitration (Title 9)	8.24

Citations are used in a variety of ways, including referencing definitions, qualifying conditions, or pointing to well-defined processes. Regardless of their individual functions, citations occur throughout the Code and must be acknowledged in any rigorous study of its complexity. To give a concrete example, consider 11 U.S.C. §101(12A)(A) and (B):

> *(12A) The term debt relief agency, means any person who provides any bankruptcy assistance to an assisted person in return for the payment of money or other valuable consideration, or who is a bankruptcy petition preparer under* ***section 110****, but does not include,*
> *(A) any person who is an officer, director, employee, or agent of a person who provides such assistance or of the bankruptcy petition preparer;*
> *(B) a nonprofit organization that is exempt from taxation under* ***section 501(c)(3)*** *of the Internal Revenue Code of 1986;*

This example highlights the two classes of citations within the Code—the "within-Title" citations and the "cross-Title" citations. The first occurs within (12A) and simply reads "under section 110"—this citation is to §110 of the same Title, 11 U.S.C. §110, though the Title must be assumed from omission. The second occurs within (12A)(B) and reads "under section 501(c)(3) of the Internal Revenue Code of 1986"—this citation is to the well-known 26 U.S.C. §501(c)(3), which falls under Title 26.[27]

Building from the example offered above, the presence of either form of inline citation implies that in order to fully understand 11 U.S.C. §101(12A)(A) and (B) one needs to also consider the provisions it references. Given the cost associated with executing both the lookup and assimilation of the referenced provision, the properties of the United States Code's citation network are meaningful for those interested in the complexity of the broader object.

Taken together, the two forms of citation within the Code generate a citation network (or dependency graph) whose structural properties can be explored using

[27] As an additional complication, note that when a named Act like the IRC of 1986 is cited, one must consult a short name list in order to determine where the Act was codified.

the tools of network science. To formalize a representation of this interdependence, we consider a mathematical graph like the T_n above. Similar to the formalization provided earlier, the vertices of the graph are the elements $|V|$ of the Code. This time, however, the graph is not a tree.[28] For analytical ease, we use the section elements $|V|^s$ as the units of analysis and attribute all citations made by any elements below the section level back up to its respective parent section. In the example from Title 11 above, both the citations in 11 U.S.C. §101(12A) and 11 U.S.C. §101(12A)(B) would be attributed to 11 U.S.C. §101. Using this formulation, we can encode the citations by letting A^C be the set of directed arcs from one section to another. This yields the citation graph $G = (V^s, A_c)$. Unlike the structural graphs, this graph is not indexed by Title, as this would ignore the existence of citations made across Titles. Such cross-Title citations are a testament to the interdependence that permeates the Code. The resulting citation graph G is large, and if one discards any sections which neither make nor receive citations in G, then $|V(G)| = 35{,}488$ and $|A(G)| = 145{,}091$. The largest weakly and strongly connected components likewise have 31,306 and 6,630 sections, respectively, indicating that the graph is fairly well connected.[29]

6.1 Measuring interdependence across titles

As noted above, there are two forms of interdependence within the Code—interdependence *across* Titles and interdependence *within* Titles. Both forms of interdependence impact the cost of executing the knowledge acquisition protocol and thus must be considered in any composite measure of Code complexity. With respect to the interdependence across Titles, we measure the relative extent of that interdependence, as all else equal, we would anticipate that some Titles are more likely to cite other Titles and some Titles are more likely to be cited by their counterparts. To consider these questions, we define the ideas of *concept importation and exportation*. When a section in Title A cites a section in Title B, this can be characterized as A *importing* some concept from B or as B *exporting* some concept to A. Similar to the measures of flow used in international trade, we can then consider *net importation and exportation* as an aggregate flow of these concepts. If a Title is cited by other Titles more than it cites other Titles, then it is a *net exporter*, and likewise a *net importer* if the opposite condition holds.

Table 7 displays the results of this calculation. The Titles that most rely on other Titles are *Title 42—Public Health & Welfare*, *Title 22—Foreign Relations and Intercourse*, and *Title 50—War and National Defense*. The Titles that are most relied upon by other Titles are *Title 5—Government Organization and Employees*, *Title 18—Crimes and Criminal Procedure* and *Title 31—Money and Finance*. In some respects, this net flow measure is likely impacted by the sheer size of a Title. Therefore, Table 7 also reports the normalized "net flow per section"—where net

[28] Instead, the citation graph disobeys the hierarchical or vertical tree and memorializes various horizontal connections between elements.

[29] In the strongly connected component of the graph, there is a directed path from each vertex in the graph to every other vertex. In the weakly connected component, there is an undirected path from each vertex in the graph to every other vertex.

flow is measured relative to Title size. This approach offers slightly different qualitative results than those obtained using the normalized results.

Although we only provide it with passing attention, another interesting method of analysis is to examine the actual dyadic pairs between Titles that operate to generate this flow. The strongest of these directed arcs are shown in Table 8. Note that the strongest dyadic pair corresponds to the link formed between the largest exporter, Title 5, and the largest importer, Title 42.

A slightly different version of this question considers the most interdependent sections. This analysis reveals which individual sections are most interdependent, either by relying on other sections or by being relied upon themselves.[30] The results are shown in Table 9. These sections all represent concepts that are repeatedly relied on, both within their own Title as well as by sections in other Titles.

6.2 Measuring interdependence within titles

In order to measure interdependence *within* a Title, we need to consider an indexed version of the citation graph G used above. For a Title N, we consider the special sub-graph G_n representing *only citations between sections in Title N*. One way of measuring the interdependence within a Title is to ask the following question: what proportion of a Title's citations is contained in G_n? Since any citations in G that spanned across different Titles are removed in G_n, this must be a proportion between 0 and 1. Table 10 shows the Titles with the five highest and lowest proportions. Title 26, the "Tax Code", stands out as almost entirely self-contained, as 97 % of its citations are to other portions of Title 26 itself. On the other hand, more than half of the citations that Title 6 generates are directed at other Titles.

7 Building a composite measure for code complexity

We have previously argued that features such as *structure*, *interdependence*, and *language* collectively characterize the United States Code. In *infra* Sects. 4, 5 and 6, we have provided various measurements relevant to each of these important dimensions. However, in the analysis presented in previous sections, we have not simultaneously considered these features of the Code. In this section of the paper, we present a composite measure that ranks the Code's forty-nine active Titles based on multiple simultaneous criteria. Unlike previous work on the United States Code, our approach yields a composite measure for comparing Titles based on their overall complexity, not just a single property. While our effort is directed toward measuring the complexity of Titles, the framework offered could be applied to segments of the Code at any size or scope including *any* unique combination of provisions that maps to the lived experiences of practitioners or scholars. Regardless of the contours of the object being measured, the broader purpose of this effort is to estimate the

[30] A given section can feature internal references to other internal provisions. For purposes of this measurement, we do not distinguish this case from the more general case of interdependence.

Table 7 Five largest importing and exporting titles

Title	Net flow	Net flow per section
Govt. Organization and Employees (Title 5)	2,654	2.58
Crimes and Criminal Procedure (Title 18)	836	0.62
Money and Finance (Title 31)	751	1.59
Judiciary and Judicial Procedure (Title 28)	659	0.83
Internal Revenue Code (Title 26)	576	0.28
Banks and Banking (Title 12)	−514	−0.28
Conservation (Title 16)	−534	−0.11
War and National Defense (Title 50)	−561	−0.78
Foreign Relations and Intercourse (Title 22)	−719	−0.25
Public Health and Welfare (Title 42)	−846	−0.11

manner in which these factors collectively influence the costs of knowledge acquisition.

Although we are interested in developing a measure to proxy for the amount of energy one would need to expend in order to review and acquire knowledge regarding the content of the United States Code, it is important to note that results offered here are drawn from an extensive set of possible composite measures.[31] Indeed, a wide combination of functions that behave reasonably over our criteria could, at least in principle, be chosen. We have at least partially constrained the set of possible composite measures by requiring that any measurement framework meet the two following goals. First, the composite measure should be simple to understand and easy to replicate. As one goal of this paper is to open broader discussion on the measurement of complexity for written bodies of law, we believe the best way to foster such a discussion is to make our results completely accessible. Second, we believe a composite measure should be flexible enough to evaluate simple competing statements within its framework. For example, two individuals might disagree as to whether a given piece of legislation increases or decreases the overall complexity of the law. Alternatively, legal theorists as well as policy advocates might argue about whether, in the aggregate, law is growing more or less complex. While likely imperfect in some respects, the simplest composite measure that meets these two goals is the method of *weighted ranks*. There exist a variety of highly sophisticated approaches using some form of weighted ranks (Fainmesser et al. 2005; Buckley 1984; Quade 1979; Eckenrode 1965; Tukey 1957). To score a multidimensional object such as the United States Code using weighted ranks, the methodology described below is often implemented:

1. Choose the set of measurement criteria that are important for the question of interest.

[31] It is really important to highlight the wide set of potential composite complexity measures that one could contemplate. The purpose of this article is to set forth some of the core components that might be contemplated in a future application.

Table 8 Five strongest title citation dyads

Title A	Title B	Citations from A to B
Public Health and Welfare (Title 42)	Govt. Organization and Employees (Title 5)	535
War and National Defense (Title 50)	Armed Forces (Title 10)	403
Foreign Relations and Intercourse (Title 22)	Govt. Organization and Employees (Title 5)	330
Banks and Banking (Title 12)	Public Health and Welfare (Title 42)	326
Labor (Title 29)	Internal Revenue Code (Title 26)	302

2. Calculate the raw values for each of these selected criteria.
3. Convert the raw scores into a ranking (*most to least* or *least to most*).
4. Choose a scheme designed to weight the rank assigned to each of these criteria. For example, the simplest possible weighting scheme is to average the ranks across the respective criteria.
5. Calculate the *weighted rank* for each object. In the case of averaging, this is simply the sum of all ranks divided by the total number of total criteria. However, many alternatives are possible depending upon one's theory regarding the relative contribution of each criterion to overall complexity.
6. Using the *weighted rank* for each object, re-rank each object from most to least (or least to most) using the composite measure calculated in (5).

7.1 Two forms of composite measurements

Varying the selected input criteria, we consider two alternative implementations of the framework described above. Both approaches are designed to measure a distinct experience that a hypothetical end user could encounter. The two composite measures we present are constructed using either "unnormalized" or "normalized" inputs.

The unnormalized score is designed to simulate the complexity of reading and assimilating the entire content of a given Title. Imagine that each Title was separately bound and presented to an end user for consideration. The unnormalized score measures the amount of complexity the end user would encounter in the knowledge acquisition process. Under such conditions, factors such as sheer size are directly relevant to the analysis as, all else equal, we believe reviewing larger Titles will prove more costly than reviewing their shorter counterparts. Therefore, the unnormalized measure is composed of measures that do not control for Title size.

By contrast, the normalized measure controls for Title size, thereby capturing the complexity of the experience of an end user who encounters a random provision within a given Title. Again, assume an individual was presented with a copy of each Title. Then, further assume the individual was instructed to open to a random page of his or her choosing. The normalized complexity score is designed to measure the

Table 9 Five most cited sections

Section	Citations received
26 U.S.C. § 501	679
8 U.S.C. § 1101	508
26 U.S.C. § 401	432
5 U.S.C. § 552	345
42 U.S.C. § 1395x	341

Table 10 Five highest and lowest proportions of intra-title citation

Title	Citations received
Internal Revenue Code (Title 26)	0.97
Bankruptcy (Title 11)	0.96
Copyrights (Title 17)	0.95
Flag and Seal, Seat of Government, and the States (Title 4)	0.92
General Provisions (Title 1)	0.92
Public Printing and Documents (Title 44)	0.59
Patriotic Societies and Observances (Title 36)	0.59
Public Buildings, Properties, and Works (Title 40)	0.59
National Guard (Title 32)	0.58
Domestic Security (Title 6)	0.55

expected level of complexity for the provision found on that particular page. While there are a set of questions for which normalization is wholly inappropriate, we believe for many substantive applications the normalized complexity score will often prove to be most appropriate. In either case, it is important to remember that both the unnormalized and normalized composite measures offer relative (ordinal) and not absolute (cardinal) measures of complexity.

7.1.1 Calculating the unnormalized measure

To calculate the unnormalized score, we select a measurement from each of the dimensions we have previously argued contribute to Code complexity. We use total provisions as a measure of size, entropy as a linguistic measure, and net flow as a measure of interdependence. Relying upon the unnormalized raw data reported in the online appendix[32] and with respect to each measure, we rank each Title from 1 (most complex) to 49 (least complex). Table 11, offered below, reports the ranks for each of the Codes' forty-nine active Titles.

With a rank for each dimension, it is possible to pool these measures into a composite measure using a weighted ranking scheme. For purposes of simplicity,

[32] The online appendix can be access here: http://computationallegalstudies.com/measuring-legal-complexity-appendix/.

we apply the most naïve of possible weighting schemes, thereby taking a simple average across these respective ranks.[33] Although we select this approach, we recognize many plausible alternatives could potentially be offered. However, in deviating from simple averaging, we believe it is necessary to demonstrate (1) the departure is justified on theoretical grounds and (2) the alternative weighting scheme yields results that are qualitatively distinct from those offered under the more naïve scheme.[34]

Table 11 reports both a composite score and a composite rank. This composite highlights the complexity of Titles such as *Title 42—Public Health & Welfare* and *Title 16—Conservation*. Of important note is that, using this measure, *Title 26—The Internal Revenue Code*, often decried for its complexity, is far from the most complex. Instead, based on this unnormalized composite approach, it is not even in the five most complex Titles. Tuning the weights, it is certainly possible to raise the relative complexity score of *Title 26*. However, it requires a rather particular configuration of weighting in order to increase the score of this well-known Title significantly. In a similar vein, a mere ocular review demonstrates the relative ranking of *Title 42* is fairly robust to a wide class of alternative weighting schemes.

7.1.2 Calculating the normalized measure

An alternative to the unnormalized weighted scheme considers the complexity of an emblematic or average provision within each Title. As noted earlier, if an end user were presented with a random provision found within a Title, the normalized complexity score is designed to measure the expected level of complexity of that provision.[35] Similar to the approach offered in the other composite measure, we start by selecting a measurement from each of the three dimensions we believe contribute to complexity.[36] We use average depth as a structural measure, entropy for language, and net flow per section as a measure for interdependence. In addition, given that we are shifting the unit of analysis from aggregate Titles to emblematic sections, we are thus interested in not only the expected depth of the average section but also the expected number of tokens contained therein. Thus, for purposes of the normalized complexity score we add an additional factor—tokens per section. This offers four separate measures whose raw scores are recorded in our online appendix.[37] Similar to the approach applied to the unnormalized measures, each of the respective columns are then ranked from 1 (most complex) to 49 (least complex). Those rankings are presented in Table 12. The far columns in Table 12

[33] In this case, this is akin to assigning each measure a weight of $\frac{1}{3}$.

[34] While mere averaging has a certain attraction, it also represents a somewhat arbitrary approach. Given that we do not have any specific theoretical grounds that justify a departure, we have chosen this naïve approach.

[35] In this case, "normalization" implies that in all components that comprise the composite measure the size of the Title is controlled for in one respect or another. Therefore, the measured highlighted Table 12 all measures feature a "per section" or some other analogous form of standardization.

[36] Again, these are *structure*, *interdependence* and *language*.

[37] The online appendix can be access here: http://computationallegalstudies.com/measuring-legal-complexity-appendix/.

Table 11 Unnormalized ranking from most to least complex

Title	Vertices	Entropy	Flow	Composite score	Composite rank
Public Health and Welfare (Title 42)	1	2	2	1.67	1
Conservation (Title 16)	3	3	9	5.00	2
Foreign Relations and Intercourse (Title 22)	9	5	5	6.33	3
Agriculture (Title 7)	5	6	11	7.33	4
Commerce and Trade (Title 15)	6	1	19	8.67	5
Crimes and Criminal Procedure (Title 18)	16	8	3	9.00	6
Banks and Banking (Title 12)	8	15	10	11.00	7
Government Organization and Employees (Title 5)	13	23	1	12.33	8
Internal Revenue Code (Title 26)	2	29	7	12.67	9
Armed Forces (Title 10)	7	13	21	13.67	10
Education (Title 20)	4	19	20	14.33	11
Indians (Title 25)	12	14	18	14.67	12
Money and Finance (Title 31)	22	18	4	14.67	12
Transportation (Title 49)	10	9	25	14.67	12
Customs Duties (Title 19)	14	7	28	16.33	15
Labor (Title 29)	11	16	22	16.33	15
Navigation and Navigable Waters (Title 33)	18	4	29	17.00	17
War and National Defense (Title 50)	21	22	8	17.00	17
Judiciary and Judicial Procedure (Title 28)	27	20	6	17.67	19
Veterans' Benefits (Title 38)	15	25	17	19.00	20
Congress (Title 2)	19	26	14	19.67	21
Mineral Lands and Mining (Title 30)	28	11	23	20.67	22
Food and Drugs (Title 21)	17	12	35	21.33	23
Domestic Security (Title 6)	24	27	16	22.33	24

Table 11 continued

Title	Vertices	Entropy	Flow	Composite score	Composite rank
Public Buildings, Properties, and Works (Title 40)	32	21	15	22.67	25
Public Lands (Title 43)	25	10	34	23.00	26
Railroads (Title 45)	33	30	12	25.00	27
Telegraphs, Telephones, and Radiotelegraphs (Title 47)	26	17	37	26.67	28
Aliens and Nationality (Title 8)	23	28	31	27.33	29
Public Contracts (Title 41)	35	35	13	27.67	30
Shipping (Title 46)	20	24	43	29.00	31
Highways (Title 23)	31	31	26	29.33	32
Patriotic Societies and Observances (Title 36)	29	39	33	33.67	33
Territories and Insular Possessions (Title 48)	39	36	27	34.00	34
Copyrights (Title 17)	36	32	36	34.67	35
Pay and Allowances of the Uniformed Services (Title 37)	34	43	30	35.67	36
Postal Service (Title 39)	37	33	38	36.00	37
National Guard (Title 32)	43	47	24	38.00	38
Hospitals and Asylums (Title 24)	42	42	32	38.67	39
Bankruptcy (Title 11)	30	38	49	39.00	40
Public Printing and Documents (Title 44)	38	34	46	39.33	41
Patents (Title 35)	41	40	40	40.33	42
Coast Guard (Title 14)	40	37	45	40.67	43
Census (Title 13)	45	41	41	42.33	44
President (Title 3)	44	46	43	44.33	45
General Provisions (Title 1)	48	48	39	45.00	46
Intoxicating Liquors (Title 27)	46	45	46	45.67	47
Flag and Seal, Seat of Government, and the States (Title 4)	47	44	48	46.33	48
Arbitration (Title 9)	49	49	41	46.33	48

Table 12 Normalized ranking from most to least complex

Title	NetFlow rank	Token rank	Entropy rank	Depth rank	Composite score	Composite rank
Public Health and Welfare (Title 42)	2	8	2	10	5.5	1
Internal Revenue Code (Title 26)	7	2	29	1	9.75	2
Government Organization and Employees (Title 5)	1	21	23	2	11.75	3
Transportation (Title 49)	25	11	9	5	12.5	4
Money and Finance (Title 31)	4	17	18	12	12.75	5
Labor (Title 29)	22	7	16	6	12.75	5
Banks and Banking (Title 12)	10	9	15	18	13	7
Education (Title 20)	20	12	19	4	13.75	8
Food and Drugs (Title 21)	35	3	12	7	14.25	9
Crimes and Criminal Procedure (Title 18)	3	25	8	22	14.5	10
Customs Duties (Title 19)	28	13	7	11	14.75	11
Agriculture (Title 7)	11	19	6	23	14.75	11
Commerce and Trade (Title 15)	19	10	1	30	15	13
Aliens and Nationality (Title 8)	31	4	28	3	16.5	14
Telegraphs, Telephones, and Radiotelegraphs (Title 47)	37	5	17	8	16.75	15
Foreign Relations and Intercourse (Title 22)	5	32	5	27	17.25	16
Domestic Security (Title 6)	16	14	27	14	17.75	17
Conservation (Title 16)	9	29	3	32	18.25	18
Armed Forces (Title 10)	21	31	13	9	18.5	19
Veterans' Benefits (Title 38)	17	22	25	13	19.25	20
War and National Defense (Title 50)	8	27	22	21	19.5	21
Railroads (Title 45)	12	15	30	26	20.75	22
Public Buildings, Properties, and Works (Title 40)	15	33	21	15	21	23
Judiciary and Judicial Procedure (Title 28)	6	30	20	28	21	23

Table 12 continued

Title	NetFlow rank	Token rank	Entropy rank	Depth rank	Composite score	Composite rank
Navigation and Navigable Waters (Title 33)	29	26	4	31	22.5	25
Mineral Lands and Mining (Title 30)	23	23	11	34	22.75	26
Highways (Title 23)	26	1	31	35	23.25	27
Indians (Title 25)	18	36	14	33	25.25	28
Postal Service (Title 39)	38	20	33	17	27	29
Copyrights (Title 17)	36	6	32	37	27.75	30
Public Contracts (Title 41)	13	24	35	39	27.75	30
Congress (Title 2)	14	35	26	36	27.75	30
Shipping (Title 46)	43	37	24	16	30	33
Bankruptcy (Title 11)	49	16	38	19	30.5	34
Public Lands (Title 43)	34	40	10	43	31.75	35
Pay and Allowances of the Uniformed Services (Title 37)	30	18	43	38	32.25	36
Patents (Title 35)	40	28	40	25	33.25	37
Patriotic Societies and Observances (Title 36)	33	48	39	20	35	38
Territories and Insular Possessions (Title 48)	27	45	36	40	37	39
Coast Guard (Title 14)	45	44	37	24	37.5	40
National Guard (Title 32)	24	34	47	46	37.75	41
President (Title 3)	43	41	46	29	39.75	42
Public Printing and Documents (Title 44)	46	39	34	41	40	43
Hospitals and Asylums (Title 24)	32	49	42	42	41.25	44
Census (Title 13)	41	43	41	45	42.5	45
Intoxicating Liquors (Title 27)	46	38	45	44	43.25	46
Flag and Seal, Seat of Government, and the States (Title 4)	48	42	44	47	45.25	47
General Provisions (Title 1)	39	46	48	48	45.25	47
Arbitration (Title 9)	41	47	49	49	46.5	49

report both the normalized score and the normalized rank. In this instance, the weights on each factor are set at 0.25 rather than 0.33.

A review of Table 12 indicates the positions of many Titles such as *Title 42—Public Health and Welfare* remain qualitatively unchanged. The normalized results presented in Table 12, however, do reveal some important departures from Table 11. Most notably, Title 26 has risen from the 9th to 2nd most complex Title. This is the result of its high scores in token count and average element depth. In addition, the complexity of several medium sized Titles such as *Title 47—Telegraph, Telephone and Radiotelegraphs* and *Title 29—Labor* is revealed once we control for Title size. Given the detailed nature of these forms of commercial regulation, the results are perhaps not terribly surprising. Namely, the forms of electronic communication that are regulated by Title 47 are complex and thus rules governing this important sector are in turn complex. As displayed in Table 12 as well as in the online appendix,[38] Title 47 displays a significant level of structural depth implying the presence of many internal distinctions within each section. Furthermore, Title 47 features a significant number of tokens within each of its sections.

For some Titles, controlling for Title size has the opposite impact. Under the normalized measure, *Title 22—Foreign Relations and Intercourse* and *Title 16—Conservation* experience a significant decline in their relative position.

8 Conclusion

Make the law as simple as possible, but not simpler. In theory, this should be a guiding principle for designers of legal systems. Understanding complexity, the antithesis of simplicity, is therefore important for both theoretical and practical reasons. Claims invoking the concept of complexity are common in both legal scholarship and in policy debates; however, the supporting evidence is often quite scant. In this paper, we present a framework for measuring legal complexity motivated by the specific contours of the *United States Code*. Though the Code is only a small portion of existing law, it is an important and representative body of law. We believe that our framework is appealing because it is both conceptually rigorous and empirically measurable. Our framework is conceptually rigorous because it is anchored to a model of the Code as the object of a *knowledge acquisition protocol*. By examining this protocol, we find that the *structure*, *language*, and *interdependence* of the Code determine its complexity. This conceptual justification allows us to move discussion of legal complexity past assessments akin to the adage "I know it when I see it".

Having identified these three aspects of complexity, we empirically measure them by applying computational techniques that scale to the scope of this large body of information. We combine these measurements to calculate a composite measure that scores the *relative* complexity of these Titles. This composite measure simultaneously takes into account contributions made by the structure, language, and interdependence of each Title through the use of *weighted ranks*. Weighted

[38] The online appendix can be access here: http://computationallegalstudies.com/measuring-legal-complexity-appendix/.

ranks are commonly used to pool or score objects with multidimensional or nonlinear attributes. Using this framework, scholars can evaluate various competing empirical claims. Furthermore, our weighted rank framework is flexible, intuitive, and entirely transparent, allowing other researchers to quickly replicate or extend our work.

The complexity one typically observes in modern legal systems may sometimes be motivated by an effort to particularize the law and thereby make it more effective. However, as a general matter, more complex legal systems are likely, all else equal, to be less well-designed and result in higher compliance costs and/or lower levels of compliance. At any given moment in time, there exists a finite amount of human capital in a society. Unnecessary legal complexity can drive a misallocation of that human capital toward comprehending and complying with legal rules and away from other productive ends. We believe this has serious implications for democratic theory and should be of serious concern to the designers of legal and political institutions.

Thus, while legal complexity is itself a complex question and thus it is probably not perfectly measurable down to its neurons, we believe legal complexity can be *approximated* using the tools from complex systems and computational linguistics. This paper presents the first conceptually rigorous and empirical framework for measuring the complexity of a legal system. While this represents an exploratory foray into measuring the complexity of legal corpora, we believe our methods could be applied to other legal documents, such as treaties, administrative regulations, municipal codes, state laws, or corporate policies. Furthermore, our analysis naturally lends itself to studies of legal systems over time. This is not the last word on the question but rather the first word. Thus, we believe these and other explorations into the political economy of legal complexity are warranted including the relationship between legal complexity and societal complexity, the relationship between complex laws and special interest lobbying, and the differential time evolving complexity of various areas of law.

References

Achen CH (1978) Measuring representation. Am J Polit Sci 22:475–510

Ansolabehere S, Snyder JM Jr, & Stewart C III (2001) Candidate positioning in US House elections. Am J Polit Sci 45:136–159

Arrow KJ (1963) Social choice and individual values. Yale University Press, New Haven

Austen-Smith D, Banks JS (1996) Information aggregation, rationality, and the Condorcet jury theorem. Am Polit Sci Rev 90:34–45

Barton BH (2008) Judges, lawyers, and a predictive theory of legal complexity. University of Tennessee Legal Studies Research Paper No. 31

Bates JE, Shepard HK (1993) Measuring complexity using information fluctuation. Phys Lett A 172(6):416–425

Becker GS (1983) A theory of competition among pressure groups for political influence. Q J Econ 98(3):371–400

Bibel LW (2004) AI and the conquest of complexity in law. Artif Intell Law 12(3):159–180

Bittker BI (1974) Tax reform and tax simplification. U Miami L Rev 29:1

Black D (1948) On the rationale of group decision-making. J Polit Econ 56(1):23

Bommarito MJ II, Katz DM (2010) A mathematical approach to the study of the United States code. Physics A 389(19):4195–4200
Bommarito II, Michael J, Katz DM (2009) Properties of the United States code citation network. arXiv preprint arXiv:0911.1751
Bonanno C, Collet P (2007) Complexity for extended dynamical systems. Commun Math Phys 275(3):721–748
Boose JH (1989) A survey of knowledge acquisition techniques and tools. Knowl Acquis 1(1):3–37
Boose JH, Gaines BR (1990) The foundation of knowledge acquisition. Academic Press Professional, San Diego
Bose R (2002) Information theory, coding and cryptography. Tata McGraw-Hill Education
Boulet R, Mazzega P, Bourcier D (2011) A network approach to the French system of legal codes—part I: analysis of a dense network. Artif Intell Law 19(4):333–355
Bourcier D, Mazzega P (2007) Toward measures of complexity in legal systems. In: Proceedings of the 11th international conference on artificial intelligence and law. ACM, pp 211–215
Bourcier D, Mazzega P (2007b) Codification law article and graphs. In: Lodder AR, Mommers L (eds) Legal knowledge and information systems. IOS Press, pp 29–38; ISBN 978-1-58603-810-6
Buchanan JM, Tullock G (1965) The calculus of consent: logical foundations of constitutional democracy, vol 100. University of Michigan Press, Ann Arbor
Buckley JJ (1984) The multiple judge, multiple criteria ranking problem: a fuzzy set approach. Fuzzy Sets Syst 13(1):25–37
Cecil MA (1999) Toward adding further complexity to the internal revenue code: a new paradigm for the deductibility of capital losses. U Ill L Rev 1083–1139
Cimiano P, Hotho A, Staab S (2005) Learning concept hierarchies from text corpora using formal concept analysis. J Artif Intell Res 24:305–339
Cox GW, McCubbins MD (2007) Legislative leviathan: party government in the House. Cambridge University Press, Cambridge
Csiszar I (1991) Why least squares and maximum entropy? An axiomatic approach to inference for linear inverse problems. Ann Stat 19(4):2032–2066
Donaldson SA (2003) Easy case against tax simplification. Va Tax Rev 22:645
Downs A (1957) An economic theory of democracy. Harper & Brothers, New York
Dworkin R (1986) Law's empire. Harvard University Press, Cambridge
Eckenrode RT (1965) Weighting multiple criteria. Manag Sci 12(3):180–192
Einstein A (1934) On the method of theoretical physics. Philos Sci 1(2):163–169
Epstein RA (1995) Simple rules for a complex world. Harvard University Press, Cambridge
Epstein RA (2004) The optimal complexity of legal rules. Law School, University of Chicago. Olin Working Paper No. 210
Eustice JS (1989) Tax complexity and the tax practitioner. Tax L Rev 45:7
Fainmesser I, Fershtman C, Gandal N, Panunzi F (2005) A consistent weighted ranking scheme with an application to NCAA college football rankings. Centre for Economic Policy Research
Feldman DP, Crutchfield JP (1998) Measures of statistical complexity: why? Phys Lett A 238(4):244–252
Feltovich PJ, Spiro RJ, Coulson RL, Myers-Kelson A (1995) Reductive bias and the crisis of text (in the law). J Contemp Legal Issues 6:187
Ferstl EC, von Cramon DY (2007) Time, space and emotion: fMRI reveals content-specific activation during text comprehension. Neurosci Lett 427(3):159–164
Flesch R, Gould AJ (1949) The art of readable writing. Harper, New York, p 196
Flournoy A (1994) Coping with complexity. Loyola of Los Angeles Law Rev 27(3):809
Francesconi E (2011) A learning approach for knowledge acquisition in the legal domain. In: Sartor G, Casanovas P, Biasiotti M, Fernández-Barrera M (eds) Approaches to legal ontologies. Springer, Netherlands, pp 219–233
Frisch D (2011) Commercial law's complexity. Geo Mason L Rev 18:245
Ganapathi V, Vickrey D, Duchi J, Koller D (2012) Constrained approximate maximum entropy learning of markov random fields. arXiv preprint arXiv:1206.3257
Gibbard A (1973) Manipulation of voting schemes: a general result. Econometrica 41(4):587–601
Golan A, Judge G, Perloff J (1997) Estimation and inference with censored and ordered multinomial response data. J Econom 79(1):23–51
Halford GS, Busby J (2007) Acquisition of structured knowledge without instruction: the relational schema induction paradigm. J Exp Psychol Learn Mem Cogn 33(3):586
Hamming RW (1986) Coding and information theory. Prentice-Hall, Englewood Cliffs

Harsanyi JC (1955) Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. J Polit Econ 63:309–321
Holsapple CW, Raj V, Wagner WP (2008) An experimental investigation of the impact of domain complexity on knowledge acquisition (KA) methods. Expert Syst Appl 35(3):1084–1094
Iria J (2009) A core ontology of knowledge acquisition. In: Aroyo L, Traverso P, Ciravegna F, Cimiano P, Heath T, Hyvönen E, Mizoguchi R, Oren E, Sabou M, Simperl E (eds) The semantic web: research and applications. Springer, Berlin, pp 233–247
Jaynes ET (1957) Information theory and statistical mechanics. Phys Rev 106(4):620
Kades E (1997) Laws of complexity and the complexity of laws: the implications of computational complexity theory for the law. Rutgers L Rev 49:403
Kaplow L (1995) A model of the optimal complexity of legal rules. J Law Econ Organ 11:150
Kincaid JP, Fishburne RP Jr, Rogers RL, Chissom BS (1975) Derivation of new readability formulas (automated readability index, fog count and flesch reading ease formula) for navy enlisted personnel (No. RBR-8-75). Naval Technical Training Command Millington TN Research Branch
Kintsch W, Van Dijk TA (1978) Toward a model of text comprehension and production. Psychol Rev 85(5):363
Kirman AP, Zimmermann JB (2001) Economics with heterogeneous interacting agents, vol 503. Springer, Heidelberg
Kolmogorov AN (1965) Three approaches to the quantitative definition ofinformation. Probl Inf Transm 1(1):1–7
Koppelman SA (1989) At-risk and passive activity limitations: can complexity be reduced. Tax L Rev 45:97
Lall A, Sekar V, Ogihara M, Xu J, Zhang H (2006) Data streaming algorithms for estimating entropy of network traffic. ACM SIGMETRICS Perform Eval Rev 34(1):145–156. ACM
Landauer R (1988) A simple measure of complexity. Nature 336:306–307
Landauer R (1996) The physical nature of information. Phys Lett A 217(4):188–193
Lazer D, Pentland AS, Adamic L, Aral S, Barabasi AL, Brewer D, Van Alstyne M (2009) Life in the network: the coming age of computational social science. Science (New York, NY) 323(5915):721
Lloyd S, Pagels H (1988) Complexity as thermodynamic depth. Ann Phys 188(1):186–213
Long SB, Swingen JA (1987) An approach to the measurement of tax law complexity. J Am Tax Assoc 8(2):22–36
Manning CD, Raghavan P, Schütze H (2008) Introduction to information retrieval. Cambridge University Press, Cambridge
Mazzega P, Bourcier D, Bourgine P, Nadah N, Boulet R (2011) A complex-system approach: legal knowledge, ontology, information and networks. In: Sartor G, Casanovas P, Biasiotti M, Fernández-Barrera M (eds) Approaches to legal ontologies. Springer, Netherlands, pp 117–132
McCaffery EJ (1990) Holy grail of tax simplification. Wis L Rev 1267–1322
McKelvey RD (1976) Intransitivities in multidimensional voting models and some implications for agenda control. J Econ Theory 12(3):472–482
McKelvey RD (1986) Covering, dominance, and institution-free properties of social choice. Am J Polit Sci 30:283–314
Mitchell M (2009) Complexity: a guided tour. Oxford University Press, Oxford
Nigam K, Lafferty J, McCallum A (1999) Using maximum entropy for text classification. In: IJCAI-99 workshop on machine learning for information filtering, vol 1, pp 61–67
Ohm P (2009) Computer programming and the law: a new research agenda. Vill L Rev 54:117
Ostrom E (1990) Governing the commons: the evolution of institutions for collective action. Cambridge University Press, Cambridge
Pagallo U (2010) As law goes by: topology, ontology, evolution. In: Casanovas P, Pagallo U, Sartor G, Ajani G (eds) AI approaches to the complexity of legal systems. complex systems, the semantic web, ontologies, argumentation, and dialogue. Springer, Berlin, pp 12–26
Page SE (2008) Uncertainty, difficulty, and complexity. J Theor Polit 20(2):115–149
Paul DL (1997) Sources of tax complexity: how much simplicity can fundamental tax reform achieve. NCL Rev 76:151
Phelan DR (2009) The effect of complexity of law on litigation strategy. In: Masson A, Shariff MJ (eds) Legal strategies. Springer, Berlin, pp 335–351
Pitt MM, Slemrod J (1989) The compliance cost of itemizing deductions: evidence from individual tax returns. Am Econ Rev 79:1224–1232
Pollock E, Chandler P, Sweller J (2002) Assimilating complex information. Learn Instr 12(1):61–86

Poole KT, Rosenthal H (1991) Patterns of congressional voting. Am J Polit Sci 35:228–278
Quade D (1979) Using weighted rankings in the analysis of complete blocks with additive block effects. J Am Stat Assoc 74:680
Rényi A (1961) On measures of entropy and information. In: Fourth Berkeley symposium on mathematical statistics and probability, pp 547–561
Riker WH (1962) The theory of political coalitions, vol 578. Yale University Press, New Haven
Rook LW (1993) Laying down the law: canons for drafting complex legislation. Or L Rev 72:663
Rothkopf MH, Pekeč A, Harstad RM (1998) Computationally manageable combinational auctions. Manag Sci 44(8):1131–1147
Ruhl JB (2008) Law's complexity: a primer. Ga St UL Rev 24:885
Sanderson M, Croft B (1999) Deriving concept hierarchies from text. In: Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval, ACM, pp 206–213
Schenk DH (1989) Simplification for individual taxpayers: problems and proposals. Tax L Rev 45:121
Schennach SM (2005) Bayesian exponentially tilted empirical likelihood. Biometrika 92(1):31–46
Schnotz W, Kürschner C (2008) External and internal representations in the acquisition and use of knowledge: visualization effects on mental model construction. Instr Sci 36(3):175–190
Schuck PH (1992) Legal complexity: some causes, consequences, and cures. Duke Law J 42:1–52
Schuck PE (2000) The limits of law. Westview Press, Boulder
Sen A (1970) Collective choice and social welfare. Holden Day, San Francisco
Shannon CE (1948) A mathematical theory of communication. Bell Syst Tech J 27:379–423
Shannon CE (1951) Prediction and entropy of printed English. Bell Syst Tech J 30(1):50–64
Shoham Y, Leyton-Brown K (2009) Multiagent systems: algorithmic, game-theoretic, and logical foundations. Cambridge University Press, Cambridge
Si L, Callan J (2001) A statistical model for scientific readability. In: Proceedings of the tenth international conference on Information and knowledge management. ACM, pp 574–576
Slemrod J (2005) The etiology of tax complexity: evidence from US state income tax systems. Public Financ Rev 33(3):279–299
Slemrod JB, Blumenthal M (1996) The income tax compliance cost of big business. Public Financ Rev 24(4):411–438
Soofi ES (2000) Principal information theoretic approaches. J Am Stat Assoc 95(452):1349–1353
Spiro RJ, Jehng JC (1990) Cognitive flexibility and hypertext: theory and technology for the nonlinear and multidimensional traversal of complex subject matter. Cogn Educ Multimed Explor Ideas High Technol 163–205
Stoop R, Stoop N, Bunimovich L (2004) Complexity of dynamics as variability of predictability. J Stat Phys 114(3–4):1127–1137
Surrey SS (1969) Complexity and the internal revenue code: the problem of the management of tax detail. Law Contemp Probl 34:673–710
Sweller J, Chandler P (1994) Why some material is difficult to learn. Cogn Instr 12(3):185–233
Tang A, Jackson D, Hobbs J, Chen W, Smith JL, Patel H, Beggs JM (2008) A maximum entropy model applied to spatial and temporal correlations from cortical networks in vitro. J Neurosci 28(2):505–518
Tress W (2009) Lost laws: what we can't find in the United States code. Golden Gate UL Rev 40:129
Tsallis C (1988) Possible generalization of Boltzmann–Gibbs statistics. J Stat Phys 52(1–2):479–487
Tukey JW (1957) Sums of random partitions of ranks. Ann Math Stat 23:987–992
Tullock G (1995) On the desirable degree of detail in the law. Eur J Law Econ 2(3):199–209
Weingast BR, Marshall WJ (1988) The industrial organization of Congress; or, why legislatures, like firms, are not organized as markets. J Polit Econ 96:132–163
White MJ (1992) Legal complexity and lawyers' benefit from litigation. Int Rev Law Econ 12(3):381–395
Wright RG (2000) Illusion of simplicity: an explanation of why the law can't just be less complex. Fla St UL Rev 27:715