The Cultural Evolution of National Constitutions

Daniel N. Rockmore
Department of Computer Science, Dartmouth College, Hanover, NH 03755, USA.
E-mail: rockmore@math.dartmouth.edu

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA

The Santa Fe Institute, Santa Fe, NM 87501, USA

The Neukom Institute for Computational Science, Dartmouth College, Hanover, NH 03755, USA

Chen Fang
Department of Computer Science, Dartmouth College, Hanover, NH 03755, USA

Nicholas J. Foti
Department of Statistics, University of Washington, Seattle, WA 98195-4322, USA

Tom Ginsburg
University of Chicago Law School, The University of Chicago, Chicago, IL 60637, USA

David C. Krakauer
The Santa Fe Institute, Santa Fe, NM 87501, USA

We explore how ideas from infectious disease and genetics can be used to uncover patterns of cultural inheritance and innovation in a corpus of 591 national constitutions spanning 1789–2008. Legal "ideas" are encoded as "topics"—words statistically linked in documents—derived from topic modeling the corpus of constitutions. Using these topics we derive a diffusion network for borrowing from ancestral constitutions back to the US Constitution of 1789 and reveal that constitutions are complex cultural recombinants. We find systematic variation in patterns of borrowing from ancestral texts and "biological"-like behavior in patterns of inheritance, with the distribution of "offspring" arising through a bounded preferential-attachment process. This process leads to a small number of highly innovative (influential) constitutions some of which have yet to have been identified as so in the current literature. Our findings thus shed new light on the critical nodes of the constitution-making network. The constitutional network structure reflects periods of intense constitution creation, and systematic patterns of variation in constitutional lifespan and temporal influence.

Received September 30, 2016; revised May 26, 2017; accepted September 6, 2017

 • Published online 0 Month 2017 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.23971

Introduction

Cultural inheritance involves the diffusion of innovations, a process of interest to both biologists (Hart & Clark, 1997) and social scientists (Rogers, 1995). In biology, inheritance is governed by mechanisms of genetic transmission, which have been quantified (Christiansen, 2008). Cultural inheritance takes a variety of forms that can resemble variants of biological inheritance (Sforza & Feldman, 1981; Richerson & Boyd, 2006; Mesoudi, Whiten, & Laland, 2006), including cultural selection (Rogers & Ehrlich, 2007; Pagel, 2013). In cultural domains, complex forms of knowledge are encoded in social norms, legal principles, and scientific theories (Wimsatt, 1999; Kaiser, 2009) and follow complex forms of transmission that involve the coordinated borrowing and learning of constellations of ideas, producing a diversity of phylogenetic patterns (Mace & Holden, 2004).

Now that a large body of the cultural record has been digitized (including books [*The Google books website*, n.d.], music [*International Music Score Library Project*, n.d.], art [*ARTstor*, n.d.], etc.), new techniques of machine learning are making the quantitative analysis of high-dimensional cultural artifacts possible. In analogy with the biological sciences, and genetics in particular, this data-mining approach to the analysis of culture is sometimes referred to as "culturomics" (Michel et al., 2010), a term born of the consideration of the frequency distribution of an *n*-gram in the Google Books corpus over time (*The Google books N-Gram Viewer*, n.d.) as proxy for how memes move in and out of the cultural record. Literature (and text generally) remains a primary focus of such work (see, e.g., Moretti, 2005; Jockers, 2013; Hughes, Foti, Krakauer, & Rockmore, 2012). A fascinating challenge is to supplement these correlation-based approaches to the understanding of cultural evolution with principled causal mechanisms directed at discovering fundamental, extra-biological evolutionary processes.

We consider the notion of *diffusion patterns* in the study of cultural inheritance as a means of tracking the diffusion of *topics* through the documents in a legal text corpus of 591 national constitutions (the full list is given in the Supplementary Materials, Table S1). "Topics" has a technical meaning here (and throughout this paper that is the sense in which the word is used) as probability distributions over words (positive weights that sum to one) that are the output of *topic modeling*, which is a computational and statistical methodology for text analysis that has made great inroads throughout the humanities (see, e.g., Riddell, 2014), to the point of reaching an almost "plug-and-play" form (see, e.g., *Stanford Topic Modeling Toolbox*, n.d.) for easy deployment. A set of topics is "learned" (i.e., automatically derived) from the corpus. The various topic distributions highlight (i.e., attach high weight to) different sets of words. In the best cases those words usually suggest a particular theme and associated labeling of the topic. Texts in the corpus are partitioned into chunks, which are thus represented as varying weighted mixtures of topics. In this way, topics provide a low-dimensional representation of the corpus in terms of higher-level ideas and provide a rigorous operational basis for a meme, to be tested against a suitable dynamics of inheritance. Although we focus on its use in the analysis of text, the topic modeling framework is more general and has been used in a number of areas (Blei, 2012).

Given a topic of some significance in a work, embodied in a set of semantically correlated legal concepts, we track its appearance and prevalence in subsequent constitutions within the corpus, as well as its extinction. While dynamical considerations have been incorporated previously into topic models (Blei & Lafferty, 2006; Wang & McCallum, 2006), this analysis differs in that we account for the diffusion of topics from document to document, and in this way reveal more clearly the patterns of genealogy and the essentially recombinant nature of textual artifacts. These resemble in the parallel domain of invention the recombinant quality of patents (Youn, Strumsky, Bettencourt, & Lobo, 2015). It is our contention that while culture is clearly an active in situ feature of human brains (Boyd & Richerson, 1996), it is also present in material artifacts that afford rich forms of combinatorial manipulation and transmission ex situ.

The corpus of national constitutions is particularly well suited to a framing and analysis as a document corpus composed of units of correlated meaning evolving according to idea diffusion and borrowing. Indeed, scholars have demonstrated that many provisions in constitutions are copied from those of other countries. For example, through *n*-gram analysis Foti et al. (Foti, Ginsburg, & Rockmore, 2014) show that constitutional preambles, which are conceptualized as the most nationally localized part of constitutions, also speak in a universal idiom and include a good deal of borrowing. Law and Versteeg (2011) have shown that rights provisions have spread around the globe. Elkins et al. (Elkins, Ginsburg, & Melton, 2009; Elkins, Ginsburg, & Simmons, 2013) show that some rights, such as freedom of expression, have become nearly universal, while others have not. Some even argue that there is a kind of global script at work, whereby nation-states seek to use constitutions to participate in global discourses (Go, 2003; Boli-Bennett, 1987; Law, 2005). This evolutionary framing of the creation of national constitutions draws on broader biological analogies for legal development across time and space (Watson, 1974). Our use of *diffusion trees* as a framework for the study of this problem (see the Methods section in the Supplementary Materials for details) can be seen as a novel quantification of this biological analogy.

It is important that we are clear that this integration of topic modeling and diffusion networks enables only a quantitative articulation and tracking of instances of thematic similarity over time. The links we demonstrate across texts are consistent with a model in which one text influences another. However, our approach does not demonstrate the specific mechanisms by which influences are transmitted, so we focus instead on the sequential patterns in which textual material flows across time and space. As we demonstrate in our Discussion, this enables an analysis enhancing traditional scholarly opinion as regards the usual notion of "influence," while also at times uncovering temporal connections suggesting further or new investigations.

Results

As mentioned, a *topic* is a probability distribution over a fixed *vocabulary* derived from a text *corpus*. It thus represents a correlated set of words encoding something like a "meme" or stochastic set of associations. (Technically, the preprocessing of the texts may result in some elements of the vocabulary set that are not words per se, but instead word stems, often called "tokens." We will use the more colloquial term "word" in this paper.) The text corpus is partitioned into *documents*, sets of roughly contiguous groupings of 500 words. This is a standard topic modeling document length, short enough to reflect local context and long enough to make sensible the statistical model. In the best case, each constitution would be partitioned into contiguous word-blocks, but processing may remove the odd abbreviation,

DOI: 10.1002/asi

title, etc., besides respecting natural boundaries, such as the end of one constitution and the beginning of another. In the case of our corpus of constitutions, each constitution generally comprises a subset of such documents. The model does not take into account word order, just which words occur and in what frequencies. This is the so-called "bag-of-words" model or representation, which is then encoded as a probability distribution over the vocabulary (the frequencies are positive and sum to one).

Topic modeling is a methodology for learning topics such that each document (represented as a bag of words) is represented as a weighted sum (mixture) of topics. In its generative form, the topic model encodes the creation of each document by first choosing a topic according to the mixture of topics that the document comprises and then choosing a word according to the distribution of that particular topic. In this respect, a constitution can be thought of as a "meme cloud" with the topics encoding the memes. We use the latent Dirichlet allocation (LDA) topic model (see (Blei, Ng, & Jordan, 2003) for a discussion of the various parameters that define the model). LDA is effectively the topic modeling industry standard. We tested several choices for the number of topics and chose 100, which we then validated (cf. the Methods section in the Supplementary Materials for details).

The output of the topic model forms the basis for our results. They include i) the discovery of the topics that make up the corpus of constitutions; ii) the determination of their flow through time ("information cascades"); iii) the reconstruction of cultural diffusion trees; iv) network analysis of diffusion trees; and v) discovery of a very biological pattern of inheritance with a highly skewed pattern of cultural fertility.

Topics

The 100 topics were "hand-labeled" by a constitution expert. Note that hand-labeling of topics is standard. Further elaboration on this can be found in our Discussion. Since generally each constitution comprises a set of corpus documents, we assign an overall constitutional weight for a topic as the average topic weight over the documents that the constitution comprises. In Table 1 we list the 10 topics with largest average topic weight (over all the constitutions), along with the 10 most probable (heavily weighted) words (in decreasing order) for each topic.[1]

Influence and Clustering

The identification of the topics now gives a natural way to represent a constitution as a mixture of probability distributions. With that, we can compare quantitatively constitutions and get at a quantitative notion of influence, completely driven by the data of the words. A first coarse pass at this is to create a constitutional "family tree," where the (unique) immediate ancestor of any given constitution is simply the constitution closest to it among all earlier constitutions. Given that our constitutions are now represented as probability distributions (over topics), a natural measure of distance is the Kullback-Liebler (KL) divergence. Recall that the KL divergence of probability distributions P and Q is defined as $KL(P||Q)=\sum_i P(i)\log\frac{P(i)}{Q(i)}$. KL is inherently nonsymmetric. A standard interpretation[2] is the degree to which a distribution Q approximates another distribution P. So thinking of an earlier constitution as a potential model for a newly written constitution, the KL divergence of their underlying topic probability distributions is a natural measure of similarity.

The "KL Constitution Tree" is shown in Figure 1. Note that Figure 1 is not scaled horizontally for time. The size and form of the representation presents some difficulty for reproducing legibly herein, so a separate pdf document, readily magnifiable, can be found at https://www.math.dartmouth.edu/~rockmore/kl-tree.pdf. We also include a detail.

The KL-tree is a coarse and aggregate articulation of the notion that constitutional ideas flow in time. It is also purely correlative and local. We should also like to explore global patterns of influence and the possibility of causal influence. We approach this by considering the "flow" of topics through constitutions and through time. Each instance of a topic flowing and appearing in a constitution (above some fixed threshold) is treated as a "cascade." We follow standard conventions (Leskovec, McGlohon, Faloutsos, Glance, & Hurst, 2007) and define an *information cascade* as a collection of constitutions and their timestamps where each topic in the constitution makes up a proportion greater than a robust threshold value. When two constitutions (nodes) both express a topic above threshold, then we consider this pair as a candidate for information "cascading" from the earlier to the later.

TABLE 1. Most popular topics across entire corpus, and their corresponding top 10 words.

Topic name	Top 10 words in topic
General rights	right rights citizens freedom law public guaranteed citizen everyone religious
Sovereignty	national people sovereignty law rights state flag language international equal
Public order	law public cases order one property laws authority liberty civil
Separation of powers	congress executive laws power ministers state secretaries order necessary public
Organic law	law government president organization national organic public laws social functioning
Socialism	people socialist country revolution working popular citizens system society development
Legislative sessions	session deputies sessions deputy members elected first vote majority extraordinary
Bureaucracy	papers years state department necessary respective individuals departments body power
Socialist legislature	people organs state supreme work organ presidium elected decisions committees

[1]A full list of the topics, in order of average weight, with the weights of the top 20 words, can be found at https://www.math.dartmouth.edu/~rockmore/topics_weight_order.txt.

[2]See, e.g., https://en.wikipedia.org/wiki/Kullback-Leibler_divergence.

DOI: 10.1002/asi

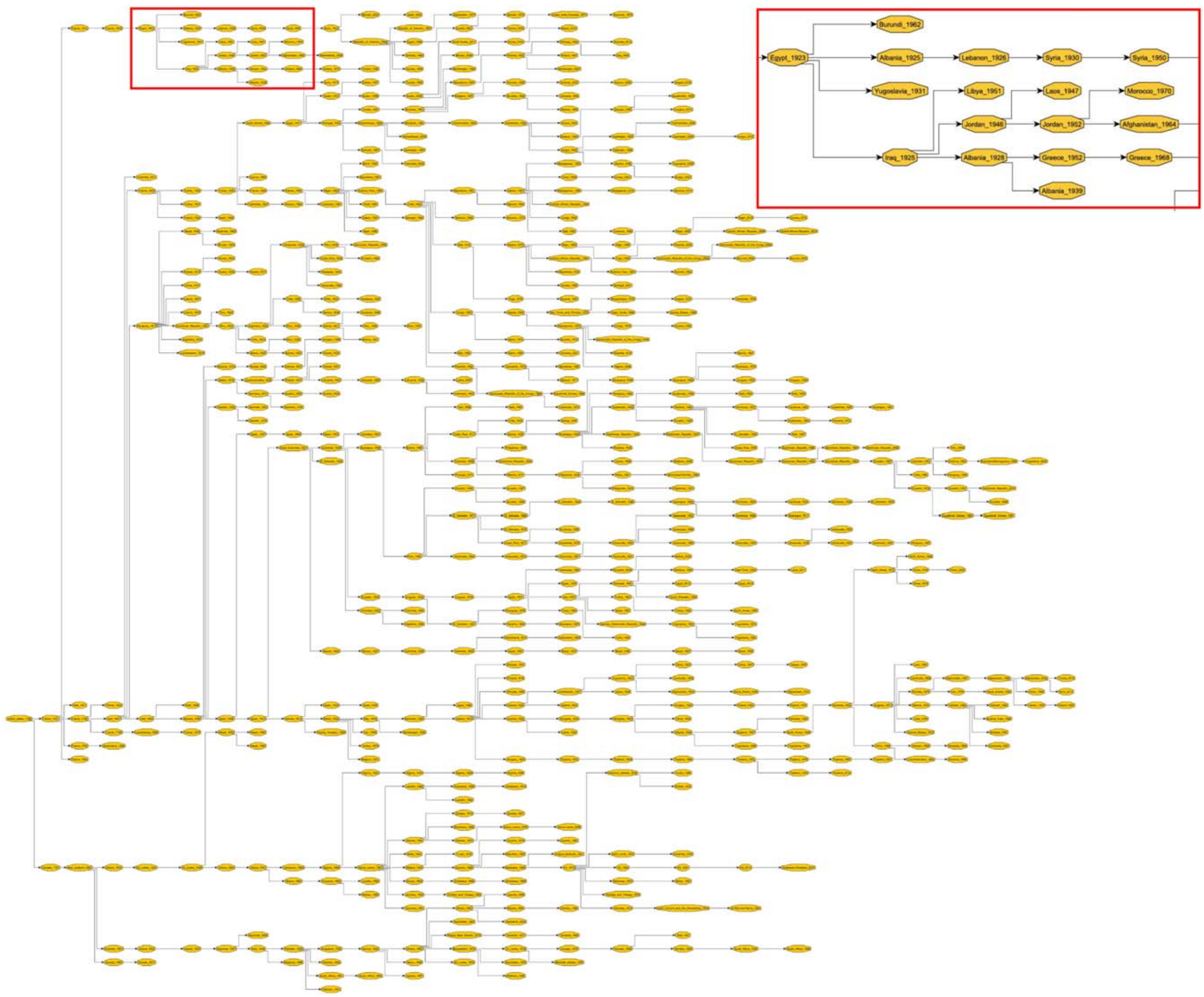

FIG. 1. The "family tree" of constitutions. The United States Constitution of 1789 is the root and thus the "last universal common ancestor constitution." Any other constitutions is deemed as having as its most recent ancestor the closest earlier constitutions where distance is measured as the KL divergence of the former to the latter. The size of this tree makes it difficult to render so that the constitution country and date are legible. A detail of the tree around the Egyptian constitution of 1923 is provided in the upper inset. Note the fertility of that constitution, as well as the sterility of the constitutions of Burundi (1962), Morocco (1970), and Albania (1939). The last of these is particularly interesting, as we see a line of descendants issuing forth from the Albania Constitution of 1925. The Albanian Constitution of 1939 was an imposed, fascist document that drew on earlier models, but had little purchase after World War II. Its most frequent topic, "subnational government," is found in such proportions in only one other, earlier text. So, earlier versions of constitutions can have patterns of transmission that do not include all of their descendants. A PDF document of this tree, easily magnifiable, can be found at https://math.dartmouth.edu/∼rockmore/kl-tree.pdf. [Color figure can be viewed at wileyonlinelibrary.com]

The topic cascades form the underlying data for a mode of inference for how ideas represented by topics are likely to have propagated through the corpus over time. As stated previously, we view the observation of a topic (above some threshold) in two constitutions as a quantitative measure indicating correlation across time. Given the content of the topics and the fact that the constitutions are ordered chronologically and typically clustered spatially (see the Network Analysis subsection below and Figure 2), shared topics may very well have spread from the earlier to the later, and hence are at least consistent with weak causality. In order to learn the most likely propagation structure of the topics (given the data) we estimate an underlying *diffusion network* for the corpus (Gomez-Rodriguez, Leskovec, & Krause, 2012). A diffusion network is a directed graph with nodes corresponding to constitutions and where the edges satisfy the condition that the source constitution predates the destination constitution. This imposes weak causal structure on the correlations. Importantly, we do not observe the diffusion network, but only the cascades that are assumed to diffuse over it and are consistent with it. In brief, a probabilistic model describing the consistency of the observed cascades with respect to a fixed diffusion network is defined. The diffusion network is that which (approximately) maximizes this probability (Gomez-Rodriguez et al., 2012).

DOI: 10.1002/asi

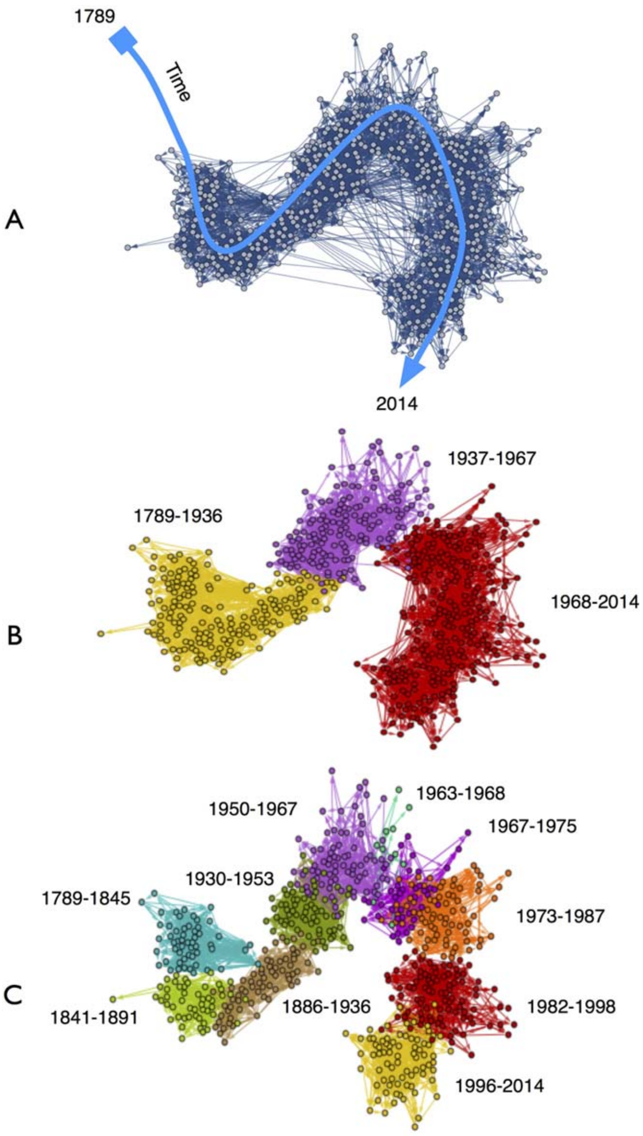

FIG. 2. (A) Spring embedded reconstruction of the constitutional diffusion network. Nodes correspond to constitutions and directed edges encode topic borrowing. The blue arrow traces time forward through the network starting in 1789 and ending in 2014. Time is the dominant factor in explaining the geometric form of the network. (B) Application of a community detection algorithm to the thresholded diffusion tree reveals three clear epochs of constitutional inheritance. The oldest epoch spans 147 years and contains 175 constitutions, giving an average of 1.2 constitutions per year. The second epoch spans 30 years and contains 148 constitutions and an average of 4.9 constitutions per year. The third epoch spans 46 years and contains 267 constitutions and an average of 5.8 constitutions per year. Thus, the rate at which constitutions are being written has increased through time, whereas the temporal influence of constitutions into the future has contracted. (C) Use of more sensitive optimal modularity methods provides a decomposition of each of these epochs into a further three epochs. Each induced cluster preserves the largely temporally contiguous ordering demonstrating that time remains a dominant dimension of variation at the microscopic level. [Color figure can be viewed at wileyonlinelibrary.com]

The presentation of the full diffusion tree on our corpus presents some visualization challenges. To give a sense of what it looks like, Supplementary Materials Figure S1 shows the entire learned diffusion tree on a restricted set of 99 constitutions. Even this is too dense to be inspected visually for information, but the figure at least gives a good sense of the way in which the methodology reifies the phenomena of the idea diffusion. Each of the edges (directed and extending downward) indicate particular topics diffusing forward in time to be taken up by subsequent constitutions. Issues of readability make it impossible to put labels on the various edges. The optimization algorithm that produces the diffusion network only collects a subset of the topics that appear in a constitution. Some diffuse forward, others do not. The "offspring" of a given constitution thus borrow certain "ideas" of the parents, but others are created afresh, presumably depending on legally appropriate contextual factors.

Network Analysis

In order to discern patterns in the diffusion tree, the diffusion network is subjected to a clustering analysis. This picks out communities of constitutions by methods of community detection and optimal modularity in which groups of constitutions share topics—and thereby a directed edge—in an amount above that expected by chance. Such a community constitutes a cluster (Newman, 2006). Figure 2 displays the results of a network reconstruction of the full circuit along with two color codings of the network resulting from the application of two forms of clustering analysis to the network. The network is illustrated using spring embedding, whereby densely connected nodes appear packed together. The network has the form of a "constitutional caterpillar" with a temporal spine threaded through the network spanning 1789 to 2014 (Figure 2A). This temporal structure is very clear in the clustering coloring. Using community structure algorithms (Girvan & Newman, 2002) we observe (Figure 2B) three clear constitutional communities, each of which describes a span of time: epoch 1: from 1789 to 1936; epoch 2: from 1937 to 1967; and epoch 3: from 1968 to 2014. Using a spectral technique for community detection we can further partition (Figure 2C) these network data into higher-order communities (Newman, 2006). This analysis maintains the chronological structure and illustrates the way in which clusters that are growing in absolute size (more constitutions in each) have evolved to encompass roughly decreasing ranges of time.

Each constitution in the diffusion tree can be described in terms of its *transmission motif—"t-motif,"* a visualization of the indegree and outdegree for each constitution. A selection of these motifs is shown in Figure 3 with a full set in Supplementary Materials Figure S2. The motifs demonstrate the variation to be found in balancing in-bound and out-bound influence for each constitution. Early constitutions tends to have few parents (e.g., Canada only has one—the US constitution) whereas subsequent constitutions vary significantly in their ancestry. This variation can be explained thorough a combination of both time (earlier constitutions present more opportunities for imitation) and how representative, novel, and applicable each constitutions is as a model for imitation.

DOI: 10.1002/asi

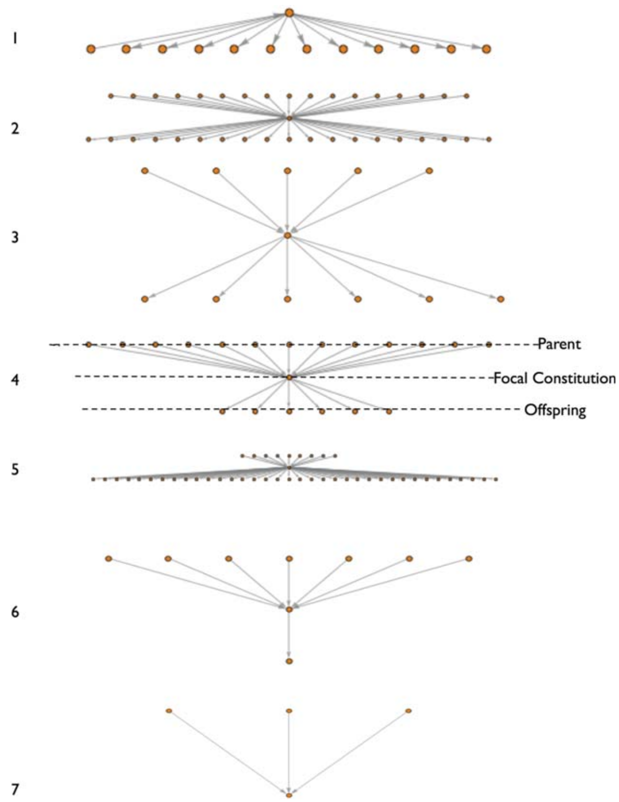

FIG. 3. Each constitution in the diffusion tree can be described in terms of a *transmission motif*, which visualizes the indegree and outdegree for each target constitution. The motifs demonstrate the balance between the in-bound and out-bound influence for each constitution in terms of a threshold number of topics that are borrowed. 1) Early constitutions tend to have few parents, e.g., Canada (1791) only has one (the US (1789) Constitution, the leftmost node in Figure 1). Subsequent constitutions vary significantly in their ancestry: 2) Iceland's (1874) Constitution has many parents and many offspring; 3) Bolivia (1826) Constitution has fewer parents and few offspring; 4) Venezuela (1830) exhibits many parents and few offspring; 5) South Korea (1948) has few parents and many offspring; (6) Albania (1976) has several parents and only one offspring; 7) Montenegro (1992) has no offspring. This variation in parentage and fertility can be explained through a combination of both the time at which they were written and the tendency to preferentially attach to a small number of highly favored models for imitation. [Color figure can be viewed at wileyonlinelibrary.com]

Models for Transmission

We can gain further insights into the patterns of inheritance by studying directly the distributions of indegree and outdegree across the entire data set. Figure 4A,B represent the pdf (probability density function) and cdf (cumulative distribution function) for the indegree for all constitutions. Illustrated in blue are the data and in orange the maximum likelihood parameter estimates for the best-fitting distribution. The indegree distribution is well captured by a Gaussian distribution with a mean of 8.8 and a standard deviation of 2.9. The estimated distribution does tend to slightly underestimate the mean but captures the tails very accurately. A straightforward interpretation is one of independent sampling of possible sources. The outdegree, however, is quite different. Figure 4C,D shows the best-fitting Poisson distribution and the outdegree distribution. Whereas the mean is effectively recovered, the tails of the distribution are poorly fitted; the Poisson underestimates the number of constitutions with few offspring and overestimates the number of constitutions with many offspring. On the other hand, consider Figure 4E,F, where we show the best-fitting negative binomial distribution to the data. This very accurately recovers the entire offspring distribution with maximum likelihood parameter estimates for the two shape parameters of the distribution as $r = 2.5$ and $p = .22$. Recall that for a negative binomial, r describes the number of offspring observed before no more offspring are generated and that the probability of producing an offspring is given by the value of p. We view this as a pure birth process, as constitutions never die—in the sense that they are always available as inspiration for a newly written constitution. Moreover, the negative binomial distributions are well known to be attractors of the Yule process (Karlin & Taylor, 1975), also known as "preferential attachment" (van der Hofstad, 2017). The excellent fit of outdegree to this distribution has broader implications for connections between offspring number and longevity. In short, that we witness a small number of constitutions of relatively early constitutions of enduring influence. All of this—including the attendant modeling considerations—is considered in some greater detail in the Discussion below.

Growth and Lifespans

We are able to track the number of new constitutions written over time. We find statistical evidence for three epochs of authorship reflecting three distinct rates of growth (Figure 5, inset). These three growth phases coincide with the three temporal groupings of the transmission graph determined through community detection. Hence, there is an association between the growth rate and the detailed community structure of the graph. We also find significant variation in the lifespan of constitutions. The lifespan is defined as the first appearance to the last instance of influence. There is a strong association between how early a constitution is written and how long it is observed to live. Unlike biological life spans, nearly all constitutions "die" young (Figure 5).

Discussion

We have searched for regular patterns of transmission in complex cultural artifacts. If there are cultural analogs to genotypes, and perhaps even phenotypes if we were to consider the broader context of constitutional influence, we should be able to observe their signatures in a temporally resolved study of evolving documents. Much like organisms that adapt to local environments, constitutions must be adapted to local cultural and legal conditions to be effective. And as with organisms, a great deal of variability in constitutions has been documented or inferred as derived from ancestral documents.

Our deeper discussion of the results starts with the labeling of the topics. We had an expert in constitutional law inspect the learned topics and provide labels for them

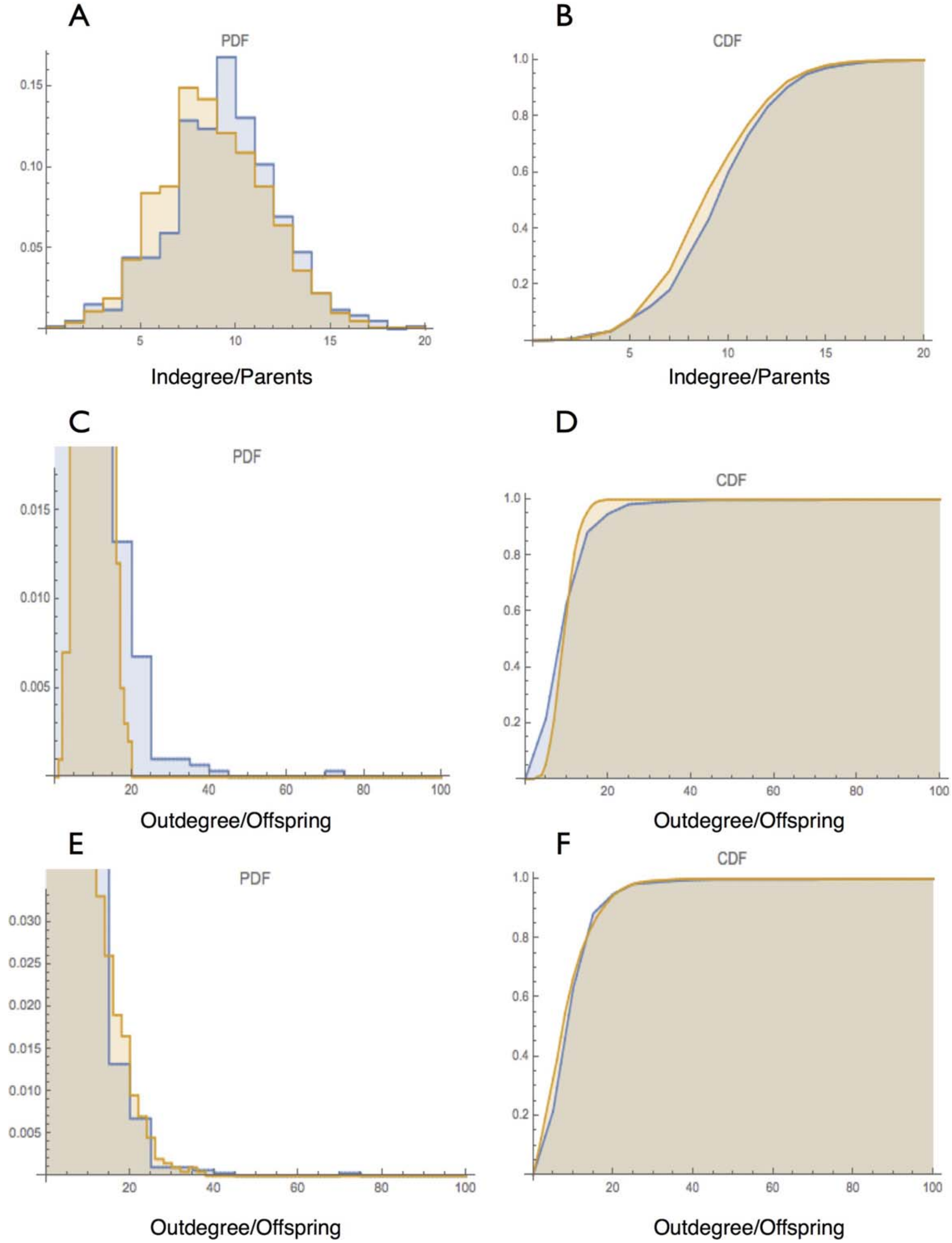

FIG. 4. Illustrated in blue are the inferred connectivity data and in orange the maximum likelihood parameter estimates for the best-fitting distributions for constitutional indegree (A,B) and outdegree (C–F). The indegree is best approximated by a Gaussian distribution with a mean of 8.8 and a standard deviation of 2.9. Figure 4C,D plot the outdegree distributions and the best-fitting Poisson distribution. Whereas the mean is effectively recovered, the tails of the distribution are poorly fitted. The Poisson underestimates the number of constitutions with few offspring and overestimates the number of constitutions with many offspring. In Figure 4E,F, we show the best-fitting negative binomial distribution to the data. This very accurately recovers the entire offspring distribution with maximum likelihood parameter estimates for the two shape parameters of the distribution as $r=2.5$ and $p=.22$. [Color figure can be viewed at wileyonlinelibrary.com]

corresponding to the dominant theme of the most probable words in each topic. We note that providing labels for the learned topics is a challenging task due to the lack of ground truth. Assigning labels to topics in our setting is essentially projecting the learned topics onto one's conception of constitutional law and (admittedly) depends heavily on the individual involved, contributing both bias and variance to the procedure. We assume that an expert in the field mitigates

DOI: 10.1002/asi

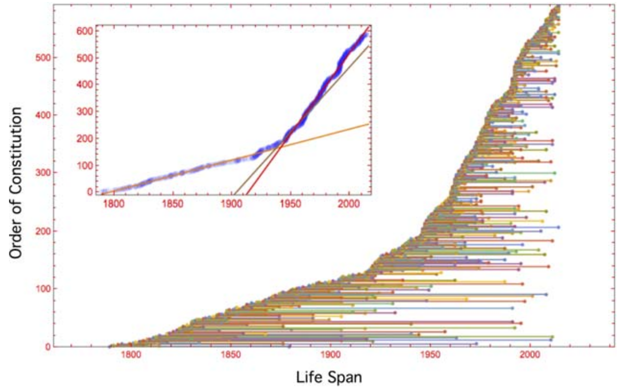

FIG. 5. Growth and lifespan of constitutions. The inset figure superimposes the n best-fitting piecewise linear regressions over the growth rate of constitutions (shown in the larger image). We discover that $n = 3$ and that these three growth rates correspond to the three epochs uncovered through the community structure analysis. We also show the life span of constitutions (first appearance to last recorded influence), with the lifespan plotted against the order of appearance of a constitution in the corpus. We clearly see how the earliest constitutions exert the longest influence on descendant constitutions—a result strongly in accord with the findings supporting a form of preferential attachment rule of influence. [Color figure can be viewed at wileyonlinelibrary.com]

both of these effects and allows us to study the corpus using the learned topics.

Perhaps given the nature of the topic labeling problem (a general lack of ground truth), there is not much prior work on solving it. An early line of research examined whether commonly used predictive measures of topic models correlated with human interpretation of the topics and found that they did not (Chang, Boyd-Graber, Gerrish, Wang, & Blei, 2009). This previous work also was the first to use human experiments to evaluate the interpretation of learned topics. More recent work has focused on incorporating knowledge bases of topics (e.g., WordNet) directly into topic models in order to encourage the model to learn topics that are interpretable by biasing them to look like topics in the knowledge base (Wood, Tan, Das, Wang, & Arnold, 2016). This is an interesting and difficult problem and further progress on it would enhance the results of this paper.

The motifs (Figure 3) illustrate clearly how constitutions are "cultural recombinants," borrowing extensively from their ancestors. Constitutions vary in their hybridicity. The motif variations suggest a constitution taxonomy, of *minor, major, idiosyncratic*, and *innovative*, depending on where in the distribution matrix (divided via the median in both dimensions) the indegree and outdegree lie. As an example of a *minor* constitution, consider Switzerland 1848. It had no descendants and only two parents (Liberia 1847 and El Salvador 1843, both of which are probably explained by temporal proximity.) A *major* constitution, on the other hand, might be Thailand's 1932 Constitution, which established a constitutional monarchy and a European-style administrative system: it had 15 parents and 33 offspring, making it the third most densely networked in the data. *Idiosyncratic* constitutions include those of Burkina Faso 1991 and Lesotho 1983, with 12 and 9 parents, respectively, but only a single offspring each. Some 20% of texts in the data have a child/parent ratio of 0.5 or less, indicating more than twice as many parental relationships as offspring. On the other hand, some 8% of constitutions in the sample have a child/parent ratio of two or more, indicating relatively high levels of *innovation*. Examples include Zambia's 1991 constitution, with 4 parents and 11 offspring, or Micronesia's constitution of 1990, with 8 parents and 24 offspring; in the latter case, it may be that the offspring are in fact those of the United States 1789, which was a very close model for Micronesian drafters. In general, parent–child relationships are temporally proximate, and they are often geographically proximate. This reflects the more general finding in the literature that time and space are powerful determinants of constitutional content. This diversity highlights an important difference from biology, where species of organisms show far less variation in the basic mechanics of transmission.

Returning to the highlighted portion of Figure 1 to illustrate the mechanisms at play, consider Egypt's 1923 Constitution and its relationship with those of its descendants. Examining the top 10 topics in each text, Egypt 1923 shares multiple topics with Albania 1925 ("act" and "public office") and Iraq's 1925 documents ("civil service" and "monarchy"), Burundi ("public office" and "labor"), and one with Yugoslavia 1931 ("mandate"). No other constitutional dyad feature these combinations of topics in the same density. While the influence of Egypt's 1923 Constitution is well known to scholars of the Arab region, it also seems to share similarities with other documents drafted shortly thereafter in neighboring parts of Europe and Africa. This illustrates how our method can point scholars to look at new links that conventional analysis might not identify.

The most fecund constitution in our network is surprising at first glance: Paraguay's 1813 Constitution. It makes sense, however, when one realizes that Latin America is the home to a plurality of constitutional texts, because it is a region of old nation-states and frequent turnover (Elkins et al., 2009). Paraguay's was the first constitution adopted in Latin America after the Spanish Constitution of Cadiz of 1812. That document embodied an ill-fated attempt to establish a liberal constitutional monarchy in Spain, featuring equality under the law and popular sovereignty, and is recognized as a model for the constitutions of Norway of 1814, Portugal of 1822, and Mexico of 1824. The top topic in this Constitution, "language of law," consists of generic legal terms that are, of course, widely used in constitutional texts. So the influence was more formal than substantive.

Conversely, some canonical constitutions do not indicate the same kind of influence in our analysis that conventional analysis would expect. For example, the 1936 Constitution of the Soviet Union is well known as a major step in the ideological development of communism, in that it incorporated many rights that were never implemented. Yet at the level of ideas, much of this involved borrowing from extant models, such as the 1931 Republican constitution of Spain. Perhaps unsurprisingly, there was little new that was in the

DOI: 10.1002/asi

USSR's constitution, and so it has few children. Similarly, the Weimar Constitution of 1919, which was thought to have embodied social democratic ideas (Venter, 2013), in fact was squarely within the topical mainstream of its time. With six parents and nine offspring, it is near the medians and its oldest direct ancestor is only 14 years prior to it. It shares three of its top 10 topics ("geography," "human rights," and "education") with Spain's Republican Constitution of 1931, which is regarded as an important and influential text. Its last direct descendent is the 1936 Constitution of the Soviet Union, with which it shares the topic "social development." This supports the claim that our method emphasizes ideological connections across text, because the Weimar Constitution is generally considered to have been a structural model for France's 1958 Constitution (Skach, 2006), although ideologically it is perhaps closer to that of the USSR.

The notion of "cultural recombination" imports one kind of biological analogy to the evolution of constitutions. The distributions of the indegree and outdegree support different biological analogies. Consider again the striking result of the fit of the outdegree distribution to the negative binomial and the indegree to the Gaussian. A principled way to understand these distributions is to derive them from suitable stochastic processes. The Gaussian distribution arises naturally from the sum of independent random variables with a well-defined mean and variance. Poisson distributions are attractors of the Galton–Watson process, whereas negative binomial distributions are attractors of the Yule process (see, e.g., Karlin & Taylor, 1975). Both Poisson and negative binomial offspring distributions are observed frequently in biological systems. The Galton–Watson process was derived to explain the extinction of family names. The idea is that at each generation a parent can transmit their name to some number of $0, 1, \ldots, n$ offspring. Each parent samples the number of offspring independently from the same distribution. Our data support a negative binomial distribution, so we shall focus on the Yule process. The Yule process is also well known as a preferential attachment process (van der Hofstad, 2017), as it can be derived from an "urn process" in which balls of a given color are sampled in linear proportion to the number of balls already in each urn. The negative binomial distribution is derived by solving a simple recurrence equation describing the temporal evolution of a probability distribution of the form,

$$P'_n(t) = -n\lambda P_n(t) + (n-1)\lambda P_{n-1}(t).$$

Here $P_n(t)$ is the probability of finding n constitutions at time t. The rate of offspring production in some interval δt is parameterized by λ. Hence, at a time t a number n of constitutions will decline through the addition of more offspring proportional to $n\lambda P_n(t)$ and increase through the production of offspring by the class $n - 1$ at a rate $(n-1)\lambda P_{n-1}(t)$. If we establish an initial condition as the number of constitutions at the start of constitutional history as 1, $P_0(0) = 1$, we find that,

$$P_n(t) = \binom{n-1}{n-n_0} e^{-\lambda n_0 t} (1 - e^{-\lambda t})^{n-n_0}.$$

Which takes the form of the negative binomial distribution in which we observe exactly n_0 offspring in n trials with a success probability, $p = e^{-\lambda t}$. For a formal exposition of preferential attachment dynamics illustrating the relationship of negative binomials to the special case of power laws, see Ross (2013).

We can test the assumptions of the Yule process by looking directly at the imitation dynamics of any given constitution. We simply plot the date on which the descendant of a given constitution was created against the order in which it was created. In Figure 6A we look at the evolution of the first 20 constitutions. By far the majority have fewer than 10 offspring and these offspring span a range of under 50 years. However, a few of these constitutions are exceptional. The most remarkable is the 1813 constitution of Paraguay, which has provided material for 70 descendant constitutions in a temporal range extending 200 years. This is followed by the original constitution of the Unites States of America from 1789 that produces 20 descendant constitutions, over a span of 80 years. The Canadian constitution of 1791 produces 11 descendants over 150 years. Figure 6B includes the first 100 constitutions, Figure 6C the first 200, and Figure 6D all 591 in the data set. A clear relationship between offspring number and longevity emerges, consistent with preferential attachment in which a small number of constitutions are of dominant influence; these appeared early in constitutional history, gaining a significant foothold, and with the vast majority of constitutions both short-lived and producing fewer than 10 offspring.

The analysis of cultural recombination through a principled decomposition of textual artifacts suggests new domains of cultural inheritance. Unlike simple Mendelian systems, or simple learning models with homogeneous rules, we observe diverse patterns of variation in the way in which nations encode important moral and legal principles. Moreover, we can obtain a principled definition of a meme—or unit of cultural transmission—that goes beyond the single "word" and captures highly linked sets of words expressing a functional, legal category—much the way a gene, composed of linked sets of nucleotides—contributes to a function. Nations differ in their debt to the past and their original contributions to the future. This allows us to speak in a rigorous fashion about phylogenetic concepts like analogy and homology when it comes to a cultural artifact. This has been an area of active research that includes the formal analysis of cultural and symbolic systems (Sforza & Feldman, 1981; Nowak & Krakauer, 1999; Nowak, Plotkin, & Krakauer, 1999), experimental approaches to cultural transmission (Henrich & McEalreath, 2003; Mesoudi & Whiten, 2008), and qualitative frameworks of integration (Mesoudi et al., 2006). At this point in time the status of key phylogenetic concepts applied to culture is in flux (Mace & Holden, 2004); we favor an instrumental approach defining cultural analogy and homology strictly in phylogenetic terms.

DOI: 10.1002/asi

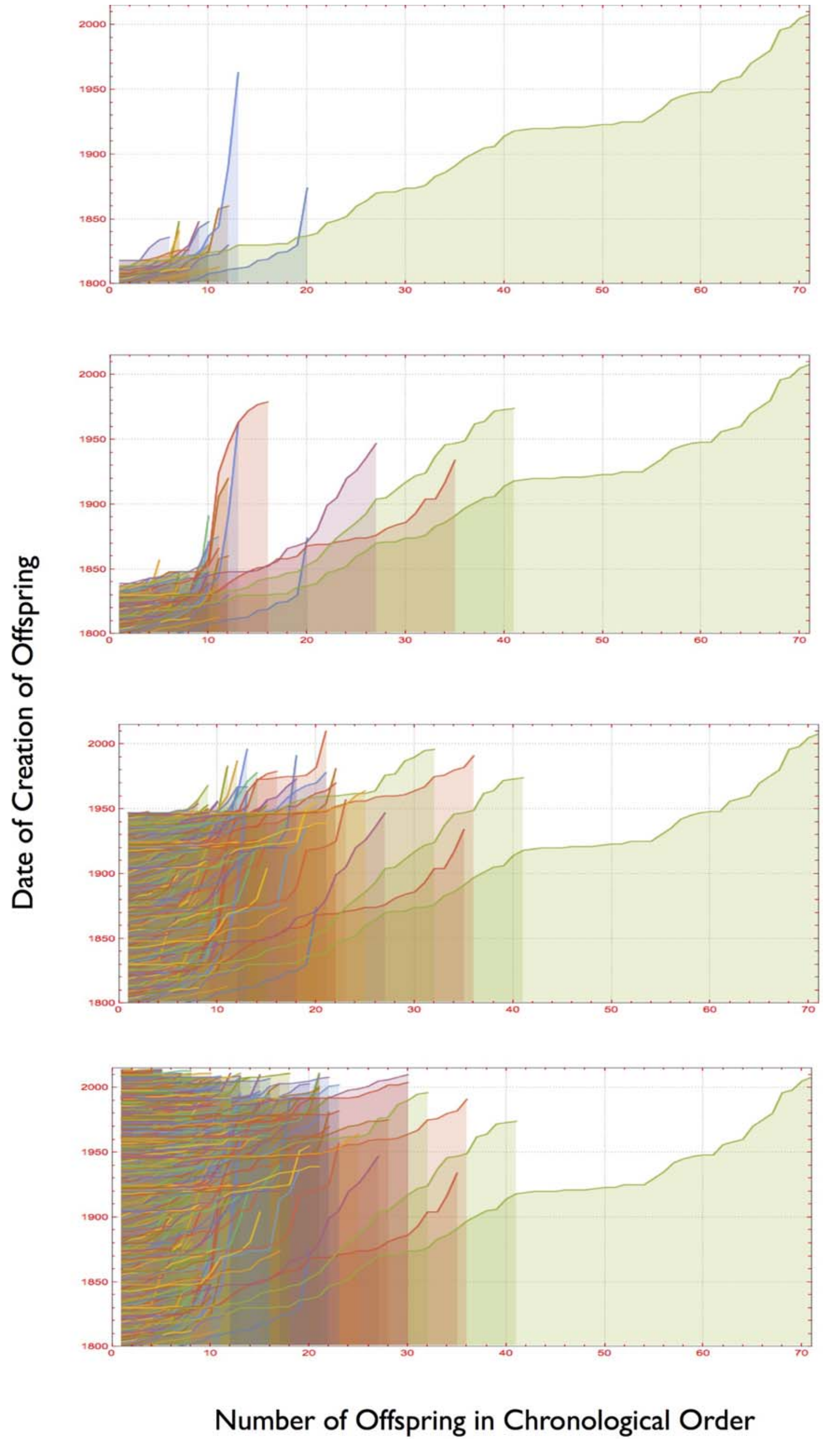

FIG. 6. Fecundity and influence of constitutions. On the *x*-axis are the number of descendant constitutions arranged in chronological order and on the *y*-axis the date of their appearance. In panel A we plot the first 20 constitutions. In panel B the first 100. Panel C shows the first 200. Panel D shows all 591. Most constitutions have few descendants and these appear over a relatively short span of time. Constitutions with many descendants tend to span longer periods of time. Most of the longest-lived constitutions in terms of influence/borrowing were written in the first of the three epochs of constitutional history (as in Figure 4A). [Color figure can be viewed at wileyonlinelibrary.com]

We suggest that the "semantic" interpretation of a given constitution and its practical legal impact is what we mean by the phenotype. We might expect many different genotypes to be neutral, in that their interpretations are equivalent, and that constitutions vary in their "penetrance," that is, their influence on cultural practices.

This approach builds on prior research related to concepts such as "citation backbones" (Gualdi, Yeung, & Zhang,

DOI: 10.1002/asi

2011) in which citations to prior publications form a tree-like structure from which novel papers descend, patent backbones in the automobile industry (Lin, Chen, & Chen, 2011), skewed patterns of borrowing in human designed artifacts (Eldredge, 2011), patterns of word borrowing (Nelson-Sathi et al., 2011), and the evolution of programming languages (Valverde & Solé, 2015).

Reconciling statistical patterns of influence with potential biases and patterns in thinking and writing will bring us closer to frameworks that connect methods of mathematical science with objects of psychological and humanistic interest in the service of new models and theories of cultural transmission and influence. The evolution of the law, with its rich textual and interpretive traditions, provides a nearly ideal model system.

Acknowledgements

DK acknowledges the support of the Feldstein Program on History Regulation and the Law at the Santa Fe Institute.

References

ARTstor. (n.d.). Retrieved from http://www.artstor.org

Blei, D.M. (2012). Probabilistic topic models. Communications of the ACM, 55, 77–84.

Blei, D.M., & Lafferty, J.D. (2006). Dynamic topic models. In *Proceedings of the 23rd international conference on machine learning* (pp. 113–120). New York: ACM. Retrieved from http://doi.acm.org/10.1145/1143844.1143859 doi: 10.1145/1143844.1143859

Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3, 993–1022.

Boli-Bennett, J. (1987). Human rights or state expansion? Cross-national definitions of constitutional rights, 1870–1970. In G. Thomas, J. Meyer, F. Ramirez, & J. Boli (eds.), *Institutional Structure* (pp. 71–91). Newbury Park, CA: Sage.

Boyd, R., & Richerson, P.J. (1996). Why culture is common but cultural evolution is rare. Proceedings of the British Academy, 88, 73–930.

Chang, J., Boyd-Graber, J.L., Gerrish, S., Wang, C., & Blei, D.M. (2009). Reading tea leaves: How humans interpret topic models. Advances in Neural Information Processing Systems 22, 23rd Annual Conference on Neural Information Processing Systems 2009. Proceedings of a meeting held 7–10 December 2009, Vancouver, British Columbia, Canada. Curran Associates, Inc., pp. 288–296.

Christiansen, F.B. (2008). *Theories of population variation in genes and genomes*. Princeton, NJ: Princeton University Press.

Eldredge, N. (2011). Paleontology and cornets: Thoughts on material cultural evolution. Evolution: Education and Outreach, 4, 364–373. Retrieved from https://doi.org/10.1007/s12052-011-0356-z

Elkins, Z., Ginsburg, T., & Melton, J. (2009). *The endurance of national constitutions*. Cambridge, UK: Cambridge University Press.

Elkins, Z., Ginsburg, T., & Simmons, B. (2013). Getting to rights: Constitutions and international law. Harvard International Law Journal, 51, 201–234.

Foti, N., Ginsburg, T., & Rockmore, D. (2014). 'We the Peoples': The global origins of constitutional preambles. George Washington International Law Review, 46, 101–134.

Girvan, M., & Newman, M.E. (2002). Community structure in social and biological networks. Proceedings of the National Academy of Sciences of United States of America, 99, 7821–7826. Retrieved from https://doi.org/10.1073/pnas.122653799

Go, J. (2003). A globalizing constitutionalism? Views from the postcolony 1945–2000. International Sociology, 18, 71–95.

Gomez-Rodriguez, M., Leskovec, J., & Krause, A. (2012). Inferring networks of diffusion and influence. TKDD, 5, 21.

The Google books N-Gram Viewer. (n.d.). Retrieved from http://books.google.com/ngrams

The Google books website. (n.d.). Retrieved from http://books.google.com/

Gualdi, S., Yeung, C.H., & Zhang, Y.-C. (2011). Tracing the Evolution of Physics on the Backbone of Citation Networks. Physical review. E, Statistical, nonlinear, and soft matter physics. 84. 046104.

Hart, D.L., & Clark, A.G. (1997). *Principles of population genetics*. Sunderland, MA: Sinauer Associates.

Henrich, J., & McEalreath, R. (2003). The evolution of cultural evolution. Evolutionary Anthropology, 12, 132–135.

Hughes, J.M., Foti, N.J., Krakauer, D.C., & Rockmore, D.N. (2012). Quantitative patterns of stylistic influence in the evolution of literature. Proceedings of the National Academy of Sciences of United States of America, 109, 7682–7686.

International Music Score Library Project. (n.d.). Retrieved from http://imslp.org/

Jockers, M. (2013). *Macroanalysis: Digital methods and literary history*. Champaign, IL: University of Illinois Press.

Kaiser, D.I. (2009). *Drawing theories apart: The dispersion of Feynman diagrams in postwar physics*. Chicago: University of Chicago Press.

Karlin, S., & Taylor, H.M. (1975). *A first course in stochastic processes*. New York: Academic Press.

Law, D. (2005). Generic constitutional law. Minnesota Law Review, 89, 652.

Law, D., & Versteeg, M. (2011). The evolution and ideology of global constitutionalism. California Law Review, 99, 1163.

Leskovec, J., McGlohon, M., Faloutsos, C., Glance, N.S., & Hurst, M. (2007). Patterns of cascading behavior in large blog graphs. Proceedings of the Seventh SIAM International Conference on Data Mining, April 26-28, 2007, Minneapolis, Minnesota, USA, pp. 551–556.

Lin, Y., Chen, J., & Chen, Y. (2011). Backbone of technology evolution in the modern era automobile industry: An analysis by the patents citation. Journal of Systems Science and Systems Engineering, 20, 416–442.

Mace, R., & Holden, C.J. (2004). A phylogenetic approach to cultural evolution. Trends in Ecology and Evolution, 20, 116–121.

Mesoudi, A., & Whiten, A. (2008). The multiple roles of cultural transmission experiments in understanding human cultural evolution. Philosophical Transactions of the Royal Society B: Biological Sciences, 363, 3489–3501. Retrieved from http://rstb.royalsocietypublishing.org/content/363/1509/3489

Mesoudi, A., Whiten, A., & Laland, K.N. (2006). Towards a unified science of cultural evolution. Behavioral and Brain Sciences, 29, 329–383.

Michel, J.-B., Shen, Y.K., Aiden, A.P., Veres, A., Gray, M.K., Team, T.G.B., ... Aiden, E.L. (2010). Quantitative analysis of culture using millions of digitized books. Science, 331, 176–182.

Moretti, F. (2005). *Graphs, maps, trees*. London: Verso Books.

Nelson-Sathi, S., List, J.-M., Geisler, H., Fangerau, H., Gray, R.D., Martin, W., & Dagan, T. (2011). Networks uncover hidden lexical borrowing in Indo-European language evolution. Proceedings of the Royal Society of London B: Biological Sciences, 278, 1794–1803. Retrieved from http://rspb.royalsocietypublishing.org/content/278/1713/1794

Newman, M. (2006). Modularity and community structure in networks. Proceedings of National Academy of Sciences of United States of America, 103, 8577–8582.

Nowak, M., & Krakauer, D. (1999). The evolution of language. Proceedings of National Academy of Sciences of United States of America, 96, 8028–8033.

Nowak, M., Plotkin, J., & Krakauer, D. (1999). The evolutionary language game. Journal of Theoretical Biology, 200, 147–162.

Pagel, M.D. (2013). *Wired for culture*. New York: W.W. Norton & Company.

Richerson, P.J., & Boyd, R. (2006). *Not by genes alone: How culture transformed human evolution*. Chicago: University of Chicago Press.

Riddell, A. (2014). How to read 22,198 journal articles: Studying the history of German studies with topic models. In M. Erlin & L.

Tatlock (Eds.), *Distant readings: Topologies of German culture in the long nineteenth century* (pp. 91–114). Rochester, NY: Camden House.
Rogers, D.S., & Ehrlich, P.R. (2007). Natural selection and cultural rates of change. Proceedings of the National Academy of Sciences of United States of America, 105, 3416–3420.
Rogers, E.M. (1995). *Diffusion of innovations*. New York: The Free Press.
Ross, N. (2013). Power laws in preferential attachment graphs and stein's method for the negative binomial distribution. Advances in Applied Probability, 45, 876–893.
Sforza, L.L.C., & Feldman, M.W. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton, NJ: Princeton University Press.
Skach, C. (2006). *Borrowing constitutional designs*. Princeton, NJ: Princeton University Press.
Stanford Topic Modeling Toolbox. (n.d.). Retrieved from http://nlp.stanford.edu/software/tmt/tmt-0.4/
Valverde, S., & Solé, R.V. (2015). Correction to 'punctuated equilibrium in the large-scale evolution of programming languages'. Journal of the Royal Society Interface, 13, 117.
van der Hofstad, R. (2017). *Random graphs and complex networks*. Cambridge, UK: Cambridge University Press.
Venter, F. (2013). *Constitutional comparison*. Cambridge, MA: Kluwer.
Wang, X., & McCallum, A. (2006). Topics over time: A non-Markov continuous-time model of topical trends. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 424–433). New York: ACM. Retrieved from http://doi.acm.org/10.1145/1150402.1150450
Watson, A. (1974). *Legal transplants*. Cambridge, UK: Cambridge University Press.
Wimsatt, W.C. (1999). Genes, memes, and cultural heredity. Biology and Philosophy, 14, 279–310.
Wood, J., Tan, P., Das, A., Wang, W., & Arnold, C. (2016). *Source-LDA: Enhancing probabilistic topic models using prior knowledge source*. Retrieved from https://arxiv.org/abs/1606.00577
Youn, H., Strumsky, D., Bettencourt, L., & Lobo, J. (2015). Invention as a combinatorial process: evidence from us patents. Journal of the Royal Society Interface, 12, 106.

DOI: 10.1002/asi